Lecture Notes in Computational Vision and Biomechanics 18

Shuo Li
Jianhua Yao *Editors*

Spinal Imaging and Image Analysis

Springer

Lecture Notes in Computational Vision and Biomechanics

Volume 18

Series editors

João Manuel R.S. Tavares, Porto, Portugal
R.M. Natal Jorge, Porto, Portugal

Editorial Advisory Board

Alejandro Frangi, Sheffield, UK
Chandrajit Bajaj, Austin, USA
Eugenio Oñate, Barcelona, Spain
Francisco Perales, Palma de Mallorca, Spain
Gerhard A. Holzapfel, Stockholm, Sweden
J. Paulo Vilas-Boas, Porto, Portugal
Jeffrey A. Weiss, Salt Lake City, USA
John Middleton, Cardiff, UK
Jose M. García Aznar, Zaragoza, Spain
Perumal Nithiarasu, Swansea, UK
Kumar K. Tamma, Minneapolis, USA
Laurent Cohen, Paris, France
Manuel Doblaré, Zaragoza, Spain
Patrick J. Prendergast, Dublin, Ireland
Rainald Löhner, Fairfax, USA
Roger Kamm, Cambridge, USA
Shuo Li, London, Canada
Thomas J.R. Hughes, Austin, USA
Yongjie Zhang, Pittsburgh, USA

The research related to the analysis of living structures (Biomechanics) has been a source of recent research in several distinct areas of science, for example, Mathematics, Mechanical Engineering, Physics, Informatics, Medicine and Sport. However, for its successful achievement, numerous research topics should be considered, such as image processing and analysis, geometric and numerical modelling, biomechanics, experimental analysis, mechanobiology and enhanced visualization, and their application to real cases must be developed and more investigation is needed. Additionally, enhanced hardware solutions and less invasive devices are demanded.

On the other hand, Image Analysis (Computational Vision) is used for the extraction of high level information from static images or dynamic image sequences. Examples of applications involving image analysis can be the study of motion of structures from image sequences, shape reconstruction from images and medical diagnosis. As a multidisciplinary area, Computational Vision considers techniques and methods from other disciplines, such as Artificial Intelligence, Signal Processing, Mathematics, Physics and Informatics. Despite the many research projects in this area, more robust and efficient methods of Computational Imaging are still demanded in many application domains in Medicine, and their validation in real scenarios is matter of urgency.

These two important and predominant branches of Science are increasingly considered to be strongly connected and related. Hence, the main goal of the LNCV&B book series consists of the provision of a comprehensive forum for discussion on the current state-of-the-art in these fields by emphasizing their connection. The book series covers (but is not limited to):

- Applications of Computational Vision and Biomechanics
- Biometrics and Biomedical Pattern Analysis
- Cellular Imaging and Cellular Mechanics
- Clinical Biomechanics
- Computational Bioimaging and Visualization
- Computational Biology in Biomedical Imaging
- Development of Biomechanical Devices
- Device and Technique Development for Biomedical Imaging
- Digital Geometry Algorithms for Computational Vision and Visualization
- Experimental Biomechanics
- Gait & Posture Mechanics
- Multiscale Analysis in Biomechanics
- Neuromuscular Biomechanics
- Numerical Methods for Living Tissues
- Numerical Simulation
- Software Development on Computational Vision and Biomechanics
- Grid and High Performance Computing for Computational Vision and Biomechanics
- Image-based Geometric Modeling and Mesh Generation
- Image Processing and Analysis
- Image Processing and Visualization in Biofluids
- Image Understanding
- Material Models
- Mechanobiology
- Medical Image Analysis
- Molecular Mechanics
- Multi-Modal Image Systems
- Multiscale Biosensors in Biomedical Imaging
- Multiscale Devices and Biomems for Biomedical Imaging
- Musculoskeletal Biomechanics
- Sport Biomechanics
- Virtual Reality in Biomechanics
- Vision Systems

More information about this series at http://www.springer.com/series/8910

Shuo Li · Jianhua Yao
Editors

Spinal Imaging and Image Analysis

Springer

Editors
Shuo Li
GE Healthcare and University
of Western Ontario
London, ON
Canada

Jianhua Yao
Clinical Center
National Institutes of Health
Bethesda, MD
USA

ISSN 2212-9391 ISSN 2212-9413 (electronic)
Lecture Notes in Computational Vision and Biomechanics
ISBN 978-3-319-12507-7 ISBN 978-3-319-12508-4 (eBook)
DOI 10.1007/978-3-319-12508-4

Library of Congress Control Number: 2014958455

Springer Cham Heidelberg New York Dordrecht London
© Springer International Publishing Switzerland 2015
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

Preface

The spine represents both a vital central axis for the musculoskeletal system and a flexible protective shell surrounding the most important neural pathway in the body, the spinal cord. Spine-related diseases or conditions are common and cause a huge burden of morbidity and cost to society. Examples include degenerative disk disease, spinal stenosis, scoliosis, osteoporosis, herniated disks, fracture/ligamentous injury, infection, tumor, and spondyloarthropathy. Treatment varies with the disease entity and the clinical scenario can be nonspecific.

Spinal imaging via computed tomography (CT), magnetic resonance imaging (MRI), radiography, ultrasound, positron emission tomography (PET), and other radiologic imaging modalities, is essential for noninvasively visualizing and assessing spinal pathology. Computational methods support and enhance the physician's ability to utilize these imaging techniques for diagnosis, noninvasive treatment, and intervention in clinical practice. Algorithms developed in the field of computer vision, computer graphics, signal processing, and machine learning have been adapted to analyze the spinal imaging.

We organize a group of experts in the field of spinal imaging, image analysis, and image guided intervention to contribute their knowledge and insight in this book. The book consists of three parts to cover a broad range of topics encompassing radiological imaging modalities, clinical imaging applications for common spine diseases, image processing, computer-aided diagnosis, quantitative analysis, data reconstruction and visualization, statistical modeling, image-guided spine intervention, and robotic surgery.

Part I of "Clinical Imaging and Applications" focuses on the clinical aspect of this topic. "Imaging of the Spine: A Medical and Physical Perspective" introduces the basic physics of routinely used imaging modalities for visualization of the anatomy and pathology of the spine. The chapter also presents the current paradigm of application in a clinical medical setting. "Arthritis of the Spine" introduces several types of arthritis that commonly affect the spine, including osteoarthritis, degeneration, ankylosing spondylitis, and rheumatoid arthritis. The chapter also

provides examples of imaging tools for diagnosing the spinal arthritis, such as radiography and magnetic resonance imaging. "Osteoporosis" introduces another common spine condition, osteoporosis, which affects most in the elder population. The chapter describes diagnostic tools to assess the osteoporotic fracture risk at the spine based on the clinical risk factors and the measurements of bone mineral density by using dual-energy X-ray absorptiometry (DXA) or quantitative CT (QCT).

Part II of "Image Processing" includes ten chapters covering several state-of-art computer techniques to analyze spine images. "Computer Aided Detection of Bone Metastases in the Thoracolumbar Spine" presents a framework for computer-aided detection of lytic and sclerotic metastatic lesions in the thoracolumbar spine using computed tomography. "Quantitative Monitoring of Bone Formation in Ankylosing Spondylitis Using Computed Tomography" presents a system to quantitatively monitor the bone formation in ankylosing spondylitis in a longitudinal study. "Three-Dimensional Spine Reconstruction from Radiographs" describes algorithms to reconstruct 3D spine structure from multiple 2D radiographs. "Vertebral Column Localization, Labeling, and Segmentation" proposes a framework to automatically locate, segment, and label spine column from CT images. "Automated Determination of the Spine-Based Coordinate System for an Efficient Cross-Sectional Visualization of 3D Spine Images" establishes a spine-based coordinate system for visualization, registration, and planning. "Cross-Modality Vertebrae Localization and Labeling Using Learning-Based Approaches" proposes another approach to localize and label spine column based on machine learning paradigm, and the method can be applied to multiple image modalities. "Articulated Statistical Shape Models of the Spine" presents statistical shape modeling of the spine. The statistical model can be applied in other image processing tasks, such as segmentation and registration. "Reconstruction of 3D Vertebral Models from a Single 2D Lateral Fluoroscopic Image" presents an alternative method to reconstruct the 3D vertebral model from just one single fluoroscopic image using a statistical shape model. "Graphical Model-Based Vertebra Identification from X-Ray Image(s)" describes a method to locate and identify the vertebra on 2D x-ray images based on a graphical model. "Model-Based Segmentation, Reconstruction and Analysis of the Vertebral Body from Spinal CT" is another model-based technique to segment the vertebral body from CTs. A few different approaches are compared.

Part III of "Image Guided Spine Intervention" consists of three chapters that describe the utilization of spine images in surgical planning and procedure. "Toward Virtual Modeling and Templating for Enhanced Spine Surgery Planning" proposes a templating technique for spine surgery planning. "Tracked Ultrasound in Navigated Spine Interventions" presents a tracked ultrasound guided spine surgical navigation system. "Robotic Assistance and Intervention in Spine Surgery" overviews the techniques for robotic assistance in spine intervention.

This is the first book fully dedicated to computational spinal imaging. The goal of this book is to build a bridge between scientists and clinicians in the field of

spinal imaging by introducing state-of-art computational methods in the context of clinical applications. Intended readers include imaging scientists interested in clinical applications and clinicians interested in computing techniques. We hope that with this book we raised attention for this important and interesting field of computational spinal imaging and would like to finally thank all contributors for their efforts in making this book possible.

Contents

Part I
Clinical Imaging and Applications

Imaging of the Spine: A Medical and Physical Perspective

Joseph E. Burns

Abstract This chapter introduces the basic methods of visualization currently used in clinical medical practice for diagnostic imaging of the spine. The different methods of visualization based on physical properties of tissue (imaging modalities) included in this introduction are radiography and fluoroscopy, computed tomography, magnetic resonance imaging, molecular imaging, and ultrasound. The basic physics of each of these techniques for visualization of the anatomy and pathology of the spine is presented, along with current paradigms of application of each in a clinical medical setting. Electronic storage and manipulation of images in a healthcare setting are discussed, with a brief discussion of image noise, resolution, and artifact as a basis of image quality.

1 Introduction

An integral part of modern medical practice, medical imaging has undergone progressive advancements in technology and capability to evaluate the internal anatomy and function of the human body non-invasively. New methodologies and modalities have led to higher standards of medical diagnosis and treatment. Already a major element of disease diagnosis, medical imaging also plays a growing role in treatment efficacy assessment, surgical planning, and medical research.

Medical imaging modalities commonly used in current medical practice include radiographs and computed tomography (X-ray modalities), magnetic resonance imaging, nuclear medicine imaging techniques, and ultrasound. Methods for visualization of the tissues of the human body vary by modality, are based on inherent physical properties of the tissues, and provide information on anatomic and physiologic characteristics. The most common tissue characteristics exploited to create

J.E. Burns (✉)
Department of Radiological Sciences, School of Medicine, University of California, Irvine, CA 92868, USA
e-mail: jburns@uci.edu

© Springer International Publishing Switzerland 2015

S. Li and J. Yao (eds.), *Spinal Imaging and Image Analysis*,
Lecture Notes in Computational Vision and Biomechanics 18,
DOI 10.1007/978-3-319-12508-4_1

these images include X-ray attenuation properties of tissue, magnetic properties of tissue, tissue vascular status, and speed of sound propagation.

Imaging modalities may then be divided into categories based on the physical characteristics used in the visualization. Modalities which are used to visualize anatomic structures based on tissue densities depend on variation in X-ray attenuation properties of tissue within the body to generate anatomic boundaries, and include radiography and fluoroscopy, as well as computed tomography (CT). The nuclear magnetic resonance properties of tissue are used in the process of anatomy visualization by magnetic resonance imaging, which, in effect, forms a map of the magnetic properties of the tissue of interest. Molecular imaging techniques encompass a number of imaging technologies, but have in common the assessment of tissue physiology by tracking of radiolabeled molecules, while additionally providing rudimentary information regarding body anatomy. Ultrasound uses variation in sound propagation speed through organs and reflection at tissue planes to create a map of sound propagation characteristics.

Each of the modalities has a different but somewhat complementary part to play in tissue assessment, evaluating the varying physical properties of tissue. Thus, each of the different modalities may be thought of as having a specific set of optimal functions for anatomic visualization and solving specific diagnostic problems. As a heuristic example, while CT can be optimal for rapid scanning, diagnosis, and accurate assessment of acute spinal vertebral fracture geometry in trauma patients, Magnetic Resonance Imaging (MRI) examination is more time consuming, necessitates co-operation on the part of the patient versus the administration of anesthesia and monitoring, and may make fine scale evaluation of osseous structures more difficult. However, in other circumstances, MRI can be more useful as a tool for the diagnosis of bone contusions, as well as soft tissue injuries of the spinal and para-spinal regions. In the case of bone contusion, for instance, MRI inquiry of the magnetic properties of tissue allows it to highlight reactive bone marrow edema which is not apparent on CT imaging.

An extensive range of spine pathology can be diagnosed and characterized using the modalities individually or in various combinations. In the following sections, modalities currently and commonly used in medical applications are discussed with illustrative clinical examples.

2 Radiographs

The most common modality of spine imaging based on sheer numbers of studies performed is the X-ray generated image, the “radiograph,” which may be thought of as a parallel, in some sense, to digital photography. Current generation X-ray systems produce an X-ray beam via an electronic tube, filter, and collimator, which is then projected onto a detector [1]. Similar to in principle to the optical wavelength photon detector found in digital cameras, the X-ray detector is also designed

to detect photons, but with differentiation that the detector is designed for the detection of X-ray spectral wavelength.

There are a few fundamental and characteristic differentiating factors between the creation of images in photography as compared to radiography. One primary difference, again, is the variation in the characteristic frequency of the radiation being used in the image formation process. A corollary of this follows, in that photons in photography reflect from the surface of the object being recorded, and are absorbed by the detector in the camera unit, creating an image of the subject's surface. X-rays photons, on the other hand, with higher frequency and energy than optical photons, pass through the tissue of the body more easily, with a lower proportion of reflected photons. Thus, the X-ray photons which are not scattered, reflected, or absorbed, pass through the patient and are absorbed by a detector plate placed on the opposite side of the patient from the X-ray source. The image formed as a result of this process is thus a transmission image.

The X-ray photons passing through the patient in this transmission imaging process penetrate through a variety of body tissues, unique to each photon pathway. There are varying X-ray attenuation factors of the internal anatomic tissues of the patient, and structures of higher attenuation (as in the case of bones, for instance) preferentially attenuate the beam, while lower attenuation structures (such as the lung) allow a higher proportion of photons to pass. The result is an image created by the variant pathways of the individual photons and so variant density combinations of tissue through which the photon passed in its beamline to the detector. Each density stack is formed by a unique cumulative superposition of the anatomy encountered, or equivalently a unique total attenuation, illustrated in Fig. 1. The resultant radiographic image consists of multiple boundary shadows created by the internal anatomy tissue planes. A simplified analogy to this phenomenon encountered in everyday life is the shadowing of light on a wall created by intervening structures in a room. This X-ray beam attenuation occurs as a result of five fundamental tissue densities composing the body, and in so doing scales the brightness of the resulting images and delineates the internal anatomic structures. The usual highest naturally-occurring tissue density and correlated highest X-ray beam attenuation is due to calcification or bone. Muscle density follows this, and then in descending order, fluid. Lipoid, or fatty tissue, continues on the decreasing density scale, finally reaching the low end of the density and beam attenuation spectrum with air or gas [1, 2]. Some of these densities are shown in Fig. 2. Metallic appliances such as orthopedic and dental hardware demonstrate higher beam attenuation than any naturally occurring tissue, Fig. 2. Again, resulting final image is thus a map of the "edges" of internal anatomy created by density differences between organs and internal structures of the body created by the transmitted photons.

In areas where no significant different in density difference exists between adjacent normal tissue structures or between normal tissue and pathology, these structure may be not be confidently distinguishable as unique entities, and so may necessitate alternative modalities of visualization. Additionally, since three-dimensional objects are collapsed into two-dimensional data by the imaging process

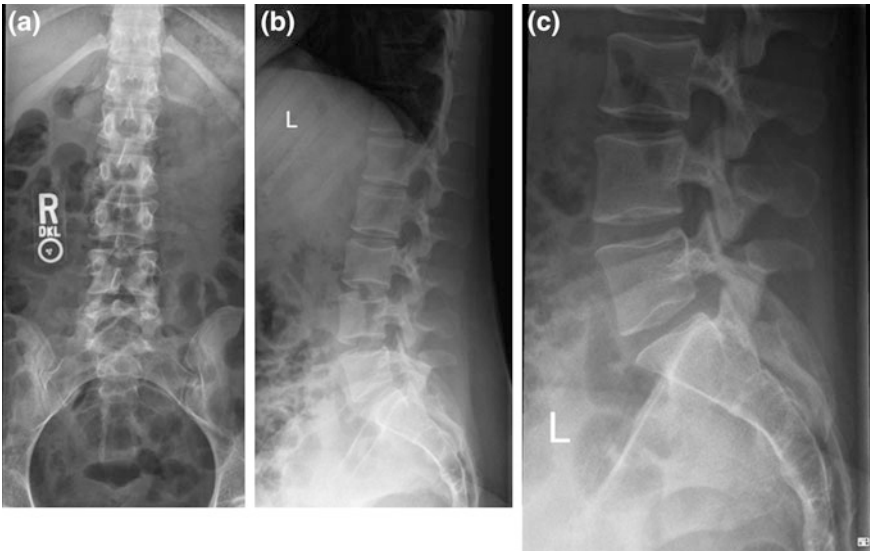


Fig. 1 Illustration of beamline passage through multiple tissue layers and resultant superposed anatomic structures. Images from a radiographic series of the lumbar spine **a** frontal view, **b** lateral view, **c** coned down lateral view. Bowel loops most prominently over-project the spine on the frontal view on the image, due to photon passage through gas filled hollow viscera. Lateral and coned down lateral views demonstrate photon passage through (at different levels) and over-projection of the spine by the lungs, bowel loops, diaphragm, and internal organs (liver and spleen)

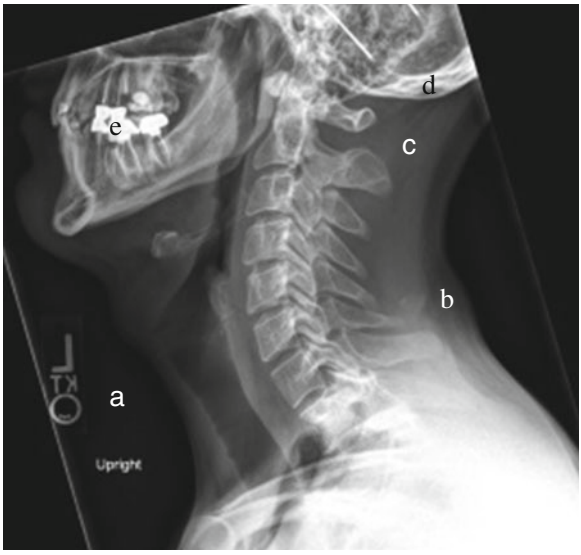


Fig. 2 Tissue density variation as manifested on radiographs. Lateral cervical spine radiograph of a 25 year old male, demonstrating examples of pixel intensity correlation with tissue density, or X-ray beam attenuation. From lesser to greater density: *a* air, *b* fatty tissue, *c* muscle, *d* bone, *e* metallic dental hardware

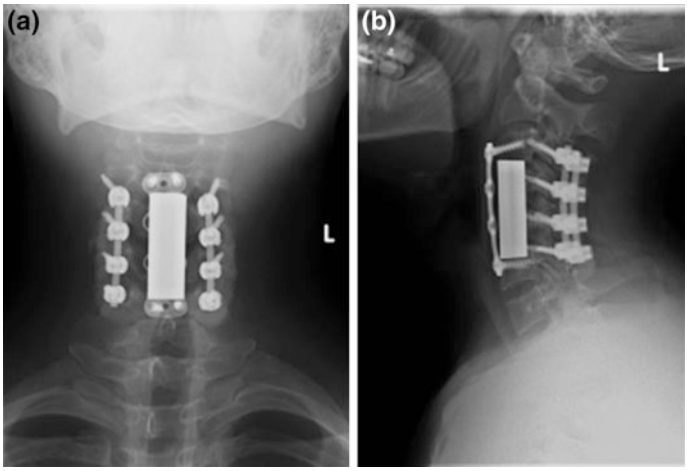


Fig. 3 Post-surgical follow up of spinal fixation procedure on a radiograph. Frontal and lateral view radiographic follow up series of spinal fixation, in a 25 year old male with history of cervical spine fracture. The patient is status-post multilevel cervical vertebral corpectomy, with anterior and posterior instrumented fusion from C3 to C6, and with integrated vertebral body strut spacer placement. Follow-up radiographic series allow for assessment of spinal alignment, as well as for inspection for hardware fracture and loosening

whereby the beam photons pass through numerous anatomic structures on their way to the detector plate, many internal structures will be superimposed on one another in the radiograph. Again, deconvolution of these objects into individually identifiable structures may necessitate alternative modalities of visualization such as cross sectional imaging. Radiographs, which are incomplete data sets, typically allow for a limited number of inferences to be drawn. Multiple orthogonal perspective views of a body structure may be taken as part of a radiographic series, increasing the information content regarding a particular structure, decreasing the level of data degeneracy and so increasing the diagnostic usefulness.

From a clinical medical perspective, radiographs commonly find use as general screening examinations for the spine, as well as for postsurgical follow-up of spinal procedures, as in Fig. 3. Relative drawbacks for radiographic imaging include lower sensitivity for certain classes of subtle pathology such as subtle fractures, low contrast in soft tissues, and patient radiation exposure.

3 Computed Tomography

Computed tomography (CT) is a sophisticated scanning technique for creating cross sectional volumetric X-ray images of body anatomy, which are, in effect, maps of tissue density. As previously described, the beam creates X-rays which penetrate

Fig. 4 Typical CT scanner unit. Patient is placed in alignment with the central axis of the helical trajectory traced by the X-ray beam source and detectors as they rotate about the patient

multiple layers of internal anatomy to detector banks within the scanner. The X-ray beam source and detectors rotate about the patient following a helical trajectory. The beamline is directed toward the patient, who has been placed in alignment with the central axis of the helix, with detectors positioned on the opposite sides of the patient. Thus, the X-ray source projects photons through the patient along a continuous angular progression radial beamline. An example of a CT scanner is shown in Fig. 4.

The extent of X-ray attenuation along the beamlines through the patient is recorded for these radial trajectories, after which complex algorithms for data reconstruction are used to separate structures along the beamlines and construct an X-ray attenuation map of the internal elements of the patient's anatomy [3]. A volume source data set is thus derived from the patient, a three-dimensional entity, as an effective volume map of tissue density. This data is then partitioned into two-dimensional sections, typically following standardized orthogonal body planes, and stored for review by physicians who will be involved in the patient's diagnosis and treatment. An example of this is demonstrated in Fig. 5, with a single image from each series of a CT scan of the spine, reformatted into the standardized planes for clinical interpretation. These anatomic sections have an appearance as if the body had been cross sectioned, with each sectional surface displayed as an image. High resolution sectional images in arbitrary scan planes can be created by

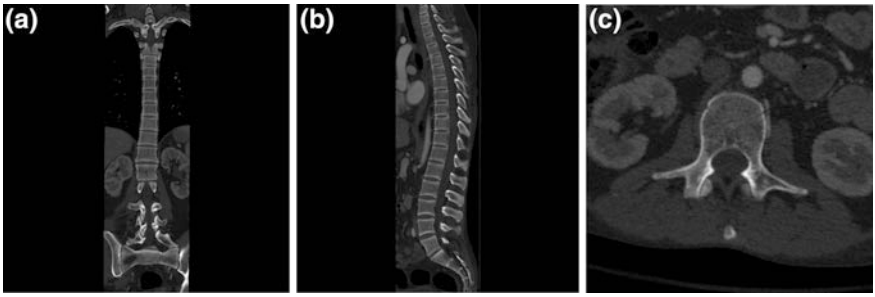


Fig. 5 CT images of the spine in standard planar reformatting. **a** Coronal plane image from a CT study of the spine, with the spine partially visualized on this image, due to normal kyphotic thoracic and lordotic lumbar curvature, **b** sagittal plane image, and **c** axial plane cross section at the level of the mid-abdomen. The axial field of view is kept small in this dedicated spine imaging study to increase the in-plane spatial resolution

current generation CT scanners. However, by convention, standardized orthogonal planes, relative to anatomic positioning of the body are obtained in clinical practice, in axial, coronal, and sagittal orientation relative to the long axis of the body. Additionally, a reconstruction kernel is used in producing the reformatted images, which may have higher spatial resolution, higher noise, and increased edge definition (a "bone" kernel), or improved contrast resolution, lower noise, and decreased edge definition (a "soft tissue" kernel), examples of which are shown in Fig. 6 [1].

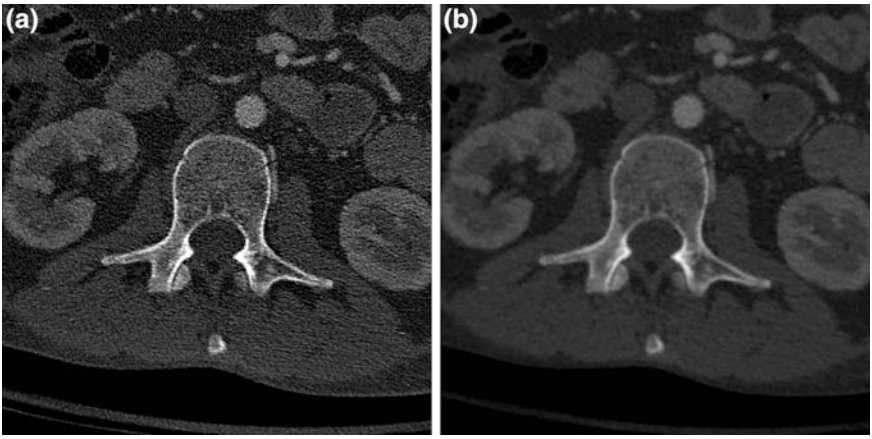


Fig. 6 CT images of the spine demonstrating different reconstruction kernels. Axial plane images from a CT study of the spine, performed with **a** "bone" reconstruction kernel for the body (B70), and **b** "soft tissue" reconstruction kernel (B40). Note difference in edge definition and noise between images

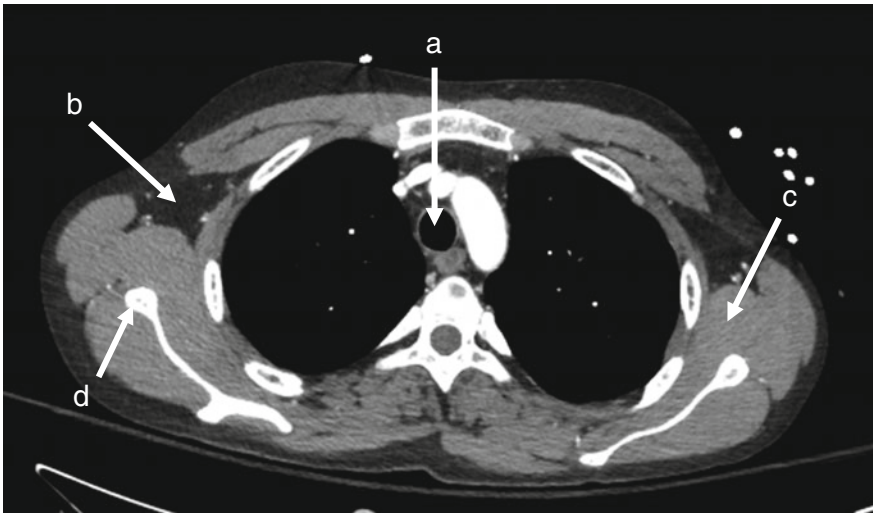


Fig. 7 Demonstration of standardized Hounsfield density units for body tissues on a CT cross section. In the CT axial cross section, measured HU in a sample ROI were **a** −999 within the trachea; **b** −105 in the fatty tissue of the axilla; **c** 45 in the muscle tissue of the shoulder; and **d** 1019 at the cortex of the scapula

As noted, both CT and radiographic images are produced with X-rays, and so represent surrogate maps of body organ density obtained via X-ray attenuation. Following the convention of previous generation radiographs taken on photographic film, images are mapped by grayscale, with dense structures such as bone scaled toward the white or bright end of the scale, and lower density materials scaled toward the dark end of the scale. With the invention of computed tomography, an intrinsic grayscale was created, with the scale subdivided into Hounsfield units (HU) [4, 5]. In the Hounsfield system of units, very low density structures are scaled in negative units, as in the case of air (approximately −1,000 HU) and fat (approximately −100 HU). Water is set at the standard of 0 HU, with muscle tissue scaled at approximately 40 HU and bone scaled at approximately 1,000 HU (Fig. 7).

On the computer monitor, Hounsfield units are scaled to pixel intensity. Medical diagnostic image display systems typically allow 10–12 bit depth, allowing the display of 1,024–4,096 shades in grayscale. Now, as the human eye can only differentiate approximately 30–40 grayscale shades, sets of restricted range brightness setting "windows" are created, centered about the tissue density of interest (Fig. 8). In materials such as bone, these windows isolate and amplify details of the anatomy of interest. The uses of CT imaging in clinical medical practice are multifold, being particularly useful for complex anatomic structures such as the spine, and for assessing certain classes of pathology such as acute fractures in non-osteopenic patients. In cases such as these, pathology on radiographic studies may be obscured or vague conferring a diagnostic advantage on CT (Fig. 9), or more complex

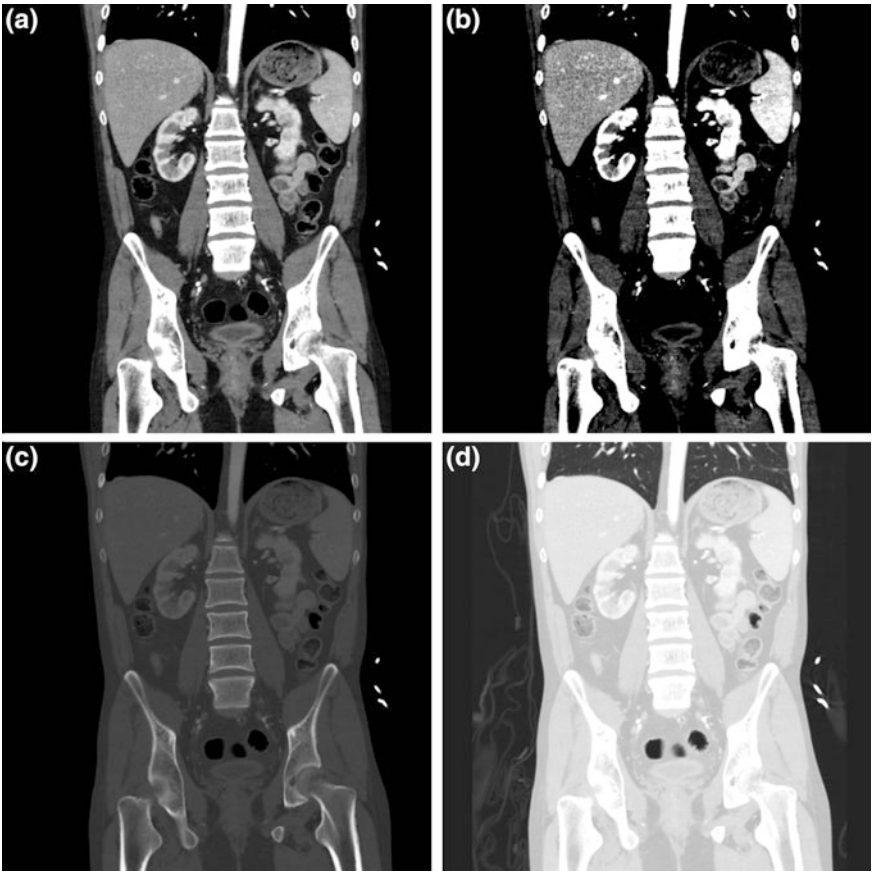


Fig. 8 Abdominal CT image set to different clinical read out windows. Coronal plane image from **a** CT study of the abdomen and pelvis delineating detail of multiple abdominal organs, in a general soft tissue window (here, w = 400, l = 60); **b** "hard contrast" soft tissue window setting (here, w = 150, l = 88) for enhancing visualization of certain abdominal organs, **c** bone window setting for assessment of skeletal structures (w = 2,000, l = 500); and **d** lung window setting (w = 1,500, l = −500) to assess air or gas containing tissues with the body, such as lungs, and hollow visci

assessment of the fracture pattern may be needed for treatment planning, again, conferring a diagnostic advantage on CT (Fig. 10). Additionally, magnetic resonance imaging may be contraindicated in certain patients, and again, CT may be helpful as a diagnostic aid. Drawbacks of CT compared to other modalities include radiation exposure to the patient, higher than for a radiograph, as well as beam hardening or streak artifact due to dense objects such as seen with orthopedic spinal fixation hardware [2].

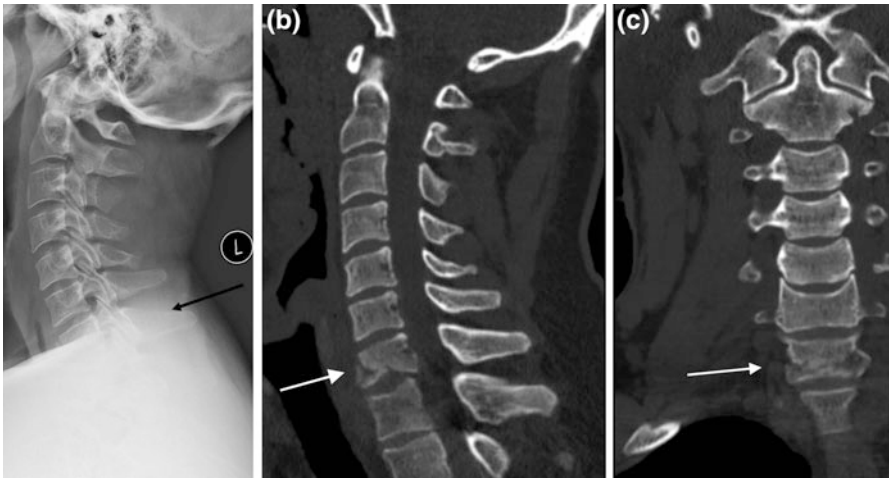


Fig. 9 Diagnostic advantage of CT over radiographs. Patient with history of trauma and neck pain. In the lateral view radiograph of the cervical spine (**a**), patient's injury is not well seen (at level of *black arrow*). Sagittal (**b**) and coronal (**c**) reformatted images from patient's CT scan demonstrates comminuted fracturing of the C7 vertebra (*white arrows*)

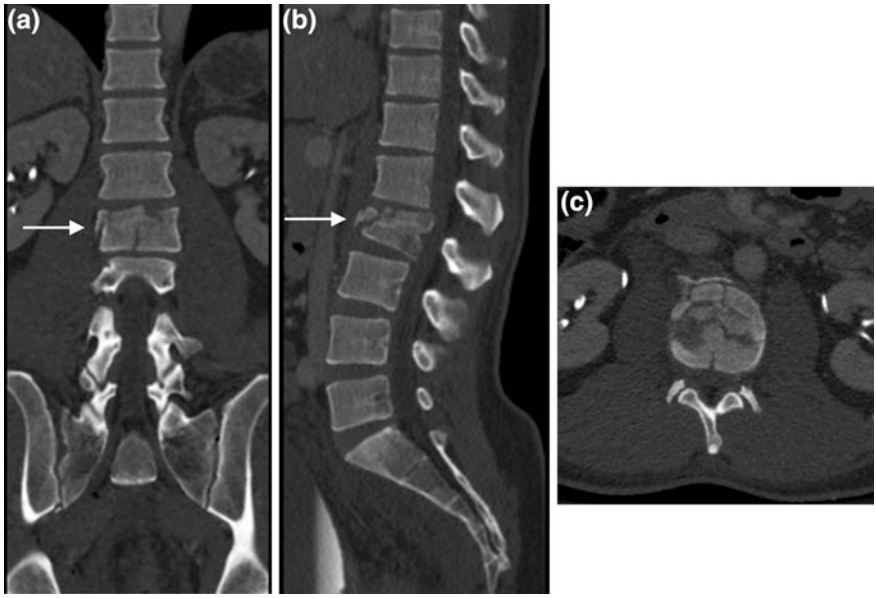


Fig. 10 Spine protocol CT images. Multiplanar reformatted images from a spine protocol CT scan, in **a** sagittal, **b** colonal, and **c** axial planes, facilitate the extraction of detailed information regarding the vertebral fracture pattern and extent allowing a more accurate assessment for treatment planning

4 Magnetic Resonance Imaging (MRI)

Nuclear magnetic resonance (NMR) initially found widespread use in analyzing the spectral characteristics of organic compounds, based on their magnetic properties. In September of 1971, Paul Lauterbur had an idea for the application of three dimensional magnetic field gradients to produce NMR images. The first images of two spatially separate tubes containing water was published in Nature in March of 1973 [6]. This quantum advance of the invention of instruments to examine material characteristics in a spatially distributed manner, led to the design of the magnetic resonance imaging (MRI) scanner. Body tissues have intrinsic and varying magnetic properties. These characteristics are exploited in MRI to create spatially distributed images of the internal anatomy of the body, which represent cross sectional anatomic maps of the magnetic characteristics of the tissues at each spatial point in the regions being imaged.

As heuristic model for magnetic properties of tissue and the basis of MRI, consider a magnetic dipole, or for a more physically conceptual macroscopic model, a bar magnet. Recall that magnetic field lines emanate from the north pole of the bar magnet, with the south pole of the magnet acting as a sink for magnetic field lines. A characteristic quantity associated with the magnetic dipole is the magnetic moment, which is expressed with the magnetic field strength and orientation of the dipole. At the atomic level, the nuclei of atoms constituting the tissue are made up of neutrons and protons. These nucleons individually have small magnetic moments, and when they occur unpaired in the nucleus they give the nucleus a net magnetic moment. The hydrogen atom, with its unpaired proton nucleus is a constituent many of molecules in the body including water and fatty tissues, and is of keystone importance in most current clinical imaging applications.

Now, when a macroscopic "magnetic dipole" is placed into an external magnetic field, it tends toward its lowest energy state, aligning with that field (as in the conceptual macroscopic analogy of a directional compass needle). Energy must be applied to turn the compass needle ("dipole") into a different direction, or higher energy state. Now, as we are on the atomic scale, quantum considerations come into play, and the dipole moment actually precesses about the direction of the applied field. The dipole moment of the proton nucleus will then tend to align with an applied magnetic field within quantum limits. The application of energy is achieved by the application of an electromagnetic energy in the form of a radiofrequency pulse [1]. This energy pulse is absorbed by the tissues, with resultant rotation of the magnetic moments into a higher energy state. The magnetic moments then give off energy as they relax toward lower energy states, at rates related to their local molecular environment, or tissue characteristics. Spatial variances in this relaxation rate by tissue type are detected and used as a basis for creating a spatial map of the magnetic properties of body tissue cross sections, or magnetic resonance images.

Armed with an understanding of the conceptual basis for MR imaging, we now consider the clinical application. MR imaging is performed by placing the patient into the center of a large magnet, typically structured in the overall geometry of a

cylinder (Fig. 11). The magnetic polarity of the applied magnetic field is alternated at characteristic and varying frequencies and pulse settings to emphasize specific properties of the tissues. A single scan utilizing characteristic phase and frequency parameters used to emphasize specific tissue properties is termed a pulse sequence. A group of these sequences, varied in terms of the tissue planes being imaged and physical characteristics of the tissue being emphasized, forms an MRI examination.

A multitude of MRI sequences have now been devised investigate specific tissue characteristics and pathologies. However, there is a fundamental set of sequences ubiquitous in use in medical imaging, known as T1, proton density (PD), and T2, examples of which for the spine are shown in Fig. 12 [7]. Each of these MRI sequences has benefits and drawbacks. The benefits of T1 include increased anatomic detail relative to T2 and ability to assess tissue enhancement with the administration of contrast. T2, on the other hand, is better for assessing edema and has generally shorter imaging times [8]. PD is an intermediate sequence, which seeks to combine T1 and T2 characteristics, with results which are intermediate between the two. Now, on grayscale imaging, certain tissues will show up as high signal intensity (white on the display monitor), and other as low signal intensity

Fig. 11 MRI scanner unit. Patient is placed on a moveable table in alignment with the central axis of the magnet gantry, which then moves into the bore of the magnet during the scanning process

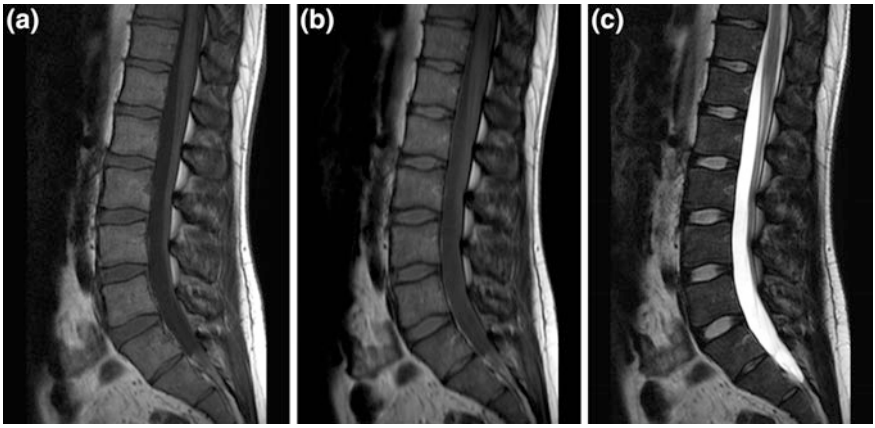


Fig. 12 Spine protocol images in varying sequence weightings. Sagittal images of the lumbar spine imaging sequences demonstrating **a** T1, **b** PD, and **c** T2 imaging

(black on the display monitor). On T1 imaging, fluid in the tissues presents as intermediate to low signal intensity, and fat as high signal intensity. In opposition, fluid on T2 appears as high signal and fat as high signal. Thus, in principle, one could differentiate fluid on a T2 sequence by comparison with a T1 sequence (see Fig. 12). There are exceptions to these signal intensity norms, including proteinaceous fluid in the body which can appear high signal intensity on T1.

Modifications of these basic sequences have been devised to expand their range of clinical utility, including what are termed fat suppression variations. In fat suppression, a process is applied by which adipose tissue, normally high signal intensity on T1 and T2, is turned to low signal intensity. One method by which this is accomplished is called fat saturation. Fat saturation depends on the slight field-dependent variance in precession frequency between the protons of fat and water and uses a frequency selective applied excitation to nullify the fat signal intensity (Fig. 13) [8, 9]. Consider an example of the clinical usage of fat saturation on a standard T2 sequence which demonstrates high signal intensity for both fat and fluid. If the fat signal intensity is now turned low after the application of fat saturation, fluid will now be the principle residual high intensity entity left on the image, amidst an intrinsically dark background of surrounding tissues composed of shades ranging from gray to black. Thus, any structure with fluid or edema (of interest in detecting pathology) will appear prominent, facilitating the detection and characterization of pathological tissue.

However, fat saturation does not always saturate the images with spatial uniformly. The effect of fat saturation depends on the resonant frequency difference between water and fat, and is subject to variable inactivation when non-uniformities occur in the applied magnetic field. This inactivation may occur locally when the patient has had metallic hardware placed, as in the case of spinal fixation hardware. Variable inactivation may also occur near the margins of the magnet or body part

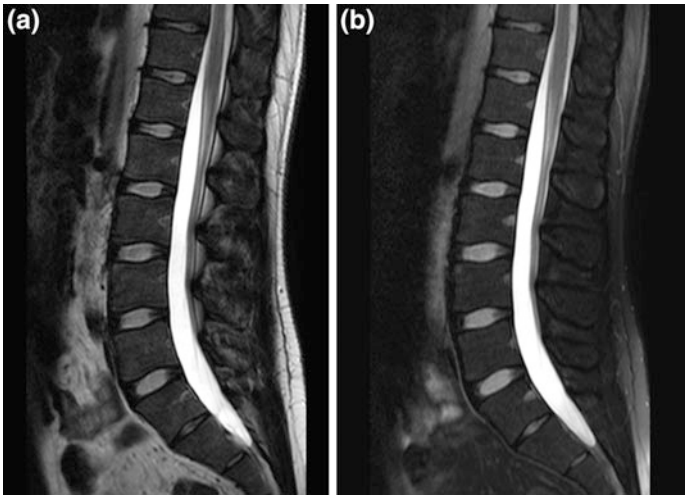


Fig. 13 Sagittal images of the lumbar spine. Sagittal images of the lumbar spine demonstrating **a** T2, and **b** T2 fat saturation imaging. Note mild nonuniformity of fat saturation of the subcutaneous tissues due to edge effects. High signal intensity within the spinal canal is cerebrospinal fluid, surrounding the caudal aspect of the spinal cord and nerve roots. Note also increased signal in the intervertebral discs due to fluid content of the nucleus pulposus

being imaged, both of which can be associated with localized non-uniform alterations in the applied magnetic field in the tissue of interest [10, 11]. In cases such as these, fat saturation is nullified and distorted non-uniformly across the image due to the change in the local field environment, with results that can potentially mimic pathology. An example of this effect in a non-spinal structure (the foot) is shown in Fig. 14. There is an alternative methodology to create an effect similar to fat saturation, termed inversion recovery (IR) imaging. Inversion recovery fat suppression imaging has different physical underpinnings, related to the difference between the longitudinal magnetization relaxation rates of the protons of water and fat after an excitation pulse is applied, and less prone to applied field non-uniform fat suppression [8, 10, 11]. An example of IR imaging used in clinical practice is Short Tau Inversion Recovery (STIR) imaging. In STIR imaging, a specifically timed inversion pulse is applied to suppress the fat signal intensity in the image. A draw-back of IR imaging is lower relative spatial resolution (Fig. 14).

As previously noted, PD sequencing is somewhat of a hybrid between the characteristics of T1 and T2. As a result, PD has higher spatial resolution of the anatomy being imaged when compared to T2 sequencing, but in general demonstrates worse spatial resolution than T1. PD sequencing facilitates the detection and assessment of fluid in anatomic tissues, amplified with the application of fat suppression, due to the increased signal intensity of fluid on PD. PD sequencing thus finds application in spine imaging, combining its relatively high spatial resolution

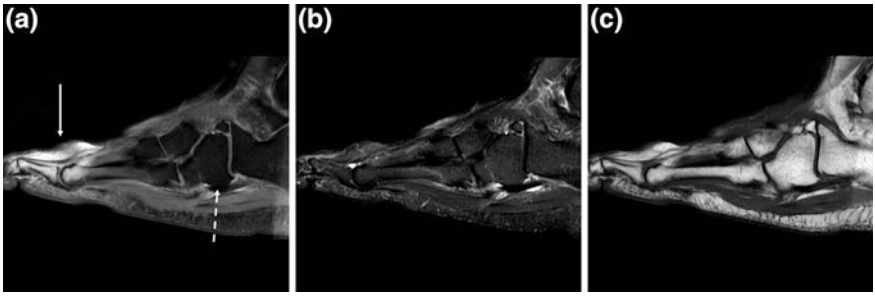


Fig. 14 Sagittal images of the foot demonstrating non-uniform fat saturation. In (**a**) a sagittal PD fat saturation (*FS*) demonstrates increased signal intensity within the metatarsal bone distally, as well as within the phalanges and surrounding soft tissues (*solid arrow*), suggestive of bone marrow and soft tissue edema, and so osteomyelitis with cellulitis. However, in the STIR image (**b**), there is no increased signal in these regions, and additionally, in T1 image (**c**), there is no commensurate decreased signal within the bone (fluid is low signal on T1) to suggest edema. Thus, the increased signal in the toe on the PD FS image is most likely due to non-uniform fat saturation. Note that fat saturation is maintained in the midfoot bone structures on the PD FD image (*dashed arrow*), leading to the term "non-uniform fat saturation."

with fluid detection characteristics to optimize the assessment of fluid associated structures such as the spinal canal.

The ability of T1 sequences to visualize and assess tissue pathology may be enhanced by gadolinium contrast administration. As previously discussed, X-rays interact with CT iodine-based contrast, to enhance the differentiation of tissue types and allow improved assessment of the physiologic processes and anatomic structures being visualized. MRI contrast materials, on the other hand, possess magnetic characteristics allowing them to interact with the magnetic field applied through the scanner. The agents most commonly used are gadolinium chelates, which possess paramagnetic properties (Fig. 15).

There are other forms of specialized MR sequencing which are also of interest in spine imaging. Diffusion tensor imaging (DTI) and Diffusion weighted imaging (DWI) are two examples of these specialized MR series, which measure the diffusion of extracellular fluid molecules through tissue using magnetic properties of the tissue and organ anatomy imaged. This extracellular diffusion of fluid molecules indirectly indicates the motion of water in tissue on the molecular level. DWI is usually performed in association with apparent diffusion coefficient (ADC) imaging. ADC is a tensor characterizing the diffusion mobility, or diffusion magnitude [12]. Combined ADC and DWI imaging clarify the etiology of increased DWI signal as due to restricted diffusion or "T2 shine through" (increased T2 signal due to fluid in the tissues). In clinical practice, DWI is used in the assessment of active demyelinating lesions in the spine, where foci of high signal intensity indicate restricted diffusion. Other uses of DWI in spinal imaging include assessment of cord infarcts, as well as increasing sensitivity for detection of osseous lesions. In combination with ADC, DWI can help differentiate acute traumatic cord injury

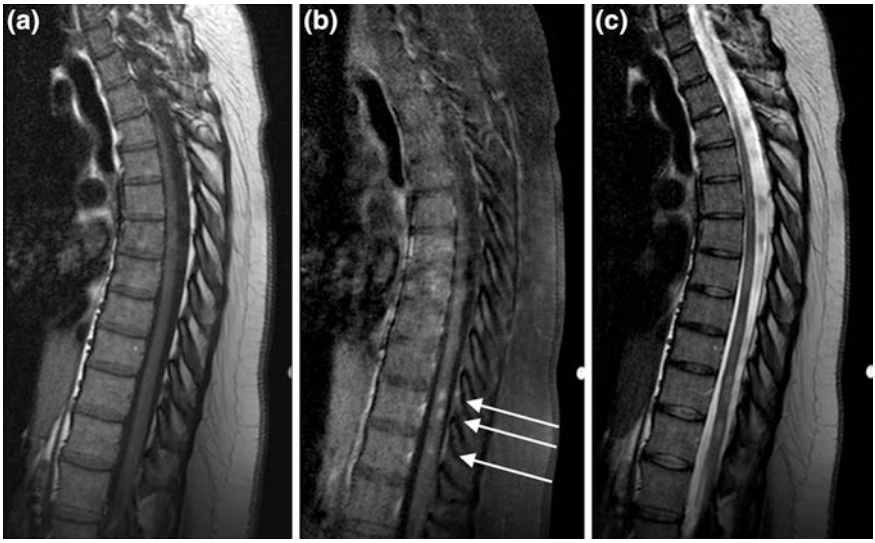


Fig. 15 T1 sagittal images of the lumbar spine before and after contrast enhancement. Sagittal plane images from an MRI study of the spine demonstrating pathologic enhancement and so increased conspicuity of multiple sclerosis lesions (*arrows*) in the spinal cord of a 52 year old patient. Images obtained include **a** T1, **b** T1 post intravascular contrast administration, and **c** T2 images

(high signal DWI, decreased ADC) from myelomalacia (intermediate to high DWI, increased ADC). The combination of fluid diffusion on the microscopic scale and anisotropy of neuronal tissue fiber tracts allow DTI to create images for assessment of white matter tracts in the spine with sensitivity at a microscopic length scale [13, 14]. Diffusion tractography is an active area of research, analyzing varied diseases of the spinal cord such as traumatic spinal cord injury, intramedullary tumors, and myelopathy.

MR imaging has some advantages in assessment of tissues as compared to ultrasound and CT, including a higher degree of soft tissue contrast, optimizing the assessment of soft tissue anatomy. Another advantage of MRI compared to CT imaging is that CT imaging exposes the patient to ionizing radiation whereas MRI does not. One disadvantage of MRI is its relative increased scan time compared to CT imaging. At the current time, a typical MR imaging sequence in clinical practice can last for 3–6 min, with 4–12 sequences per examination being common. During the time of the MRI examination, the patient must lie very still for optimal visualization, difficult under normal circumstances, and more so for patients in pain or distress. In comparison, CT scanning can typically be performed through a body region within a span of seconds. In the past, imaging of moving internal structures of the body (such as bowel and the heart) with MR was suboptimal, but there are current efforts in the development of adaptive sequences. Additionally, MR imaging in certain patients with pacemakers, neurostimulators, or cerebral

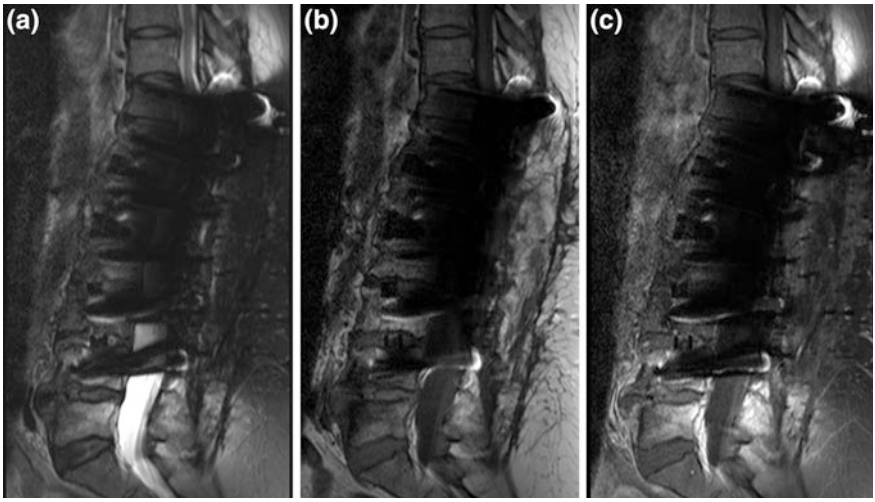


Fig. 16 Susceptibility artifact on sagittal images of the lumbar spine in a patient with spinal fixation hardware. Sagittal plane images from an MRI study of the lumbar spine demonstrating significant susceptibility artifact in this patient with metallic spinal fixation hardware, with associated image distortion and obscuration of the field of view (compare with Fig. 11, lumbar spine images from a different patient). **a** T2, **b** precontrast T1, and **c** post intravascular contrast administration images

aneurysm coils, may be contraindicated. Finally, both CT and MRI suffer from artifact obscuring or distorting pathology in patients with metallic hardware, referred to as beam hardening or streak artifact in CT scanning and magnetic susceptibility artifact in MRI [1, 15]. In MRI, the susceptibility artifact generally arises from immediate apposition of two materials of different magnetic susceptibilities, causing a distortion in the local magnetic field, and resulting in inhomogeneities in the local magnetic field. An example is shown in Fig. 16, with susceptibility artifact about metallic spinal fixation hardware.

5 Molecular Imaging (Nuclear Medicine)

The physical basis of molecular imaging (also referred to as nuclear medicine imaging) is the intravascular injection of radiolabeled tracer molecules which emit the elements of radioactive decay, followed by detection of the emitted waves and particles to assess local regions of radiotracer accumulation within the body. The radiotracers are created by complexing physiologic molecules, which are known to preferentially localize to prespecified target organs and tissues of abnormal physiology, to radionuclides. These radionuclides are typically formed by short half life gamma photon or particle emitting radioisotopes. An example of a radionuclide in

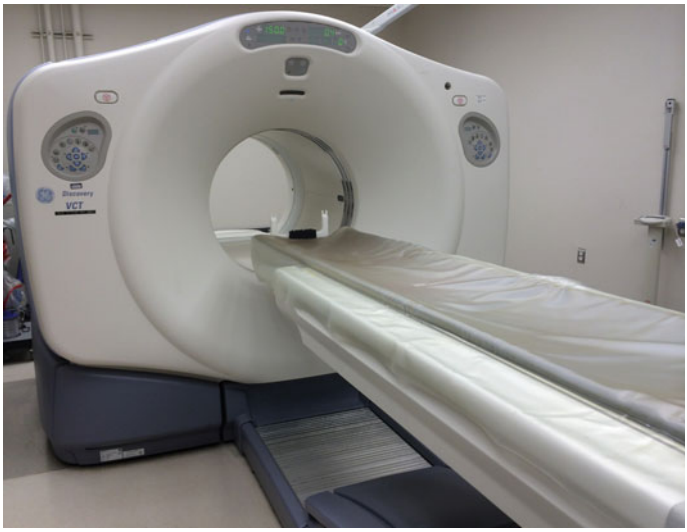


Fig. 17 PET/CT scanner unit. Patient is placed on movable table in alignment with the central axis of the cylindrical scanner bore, with aligned sequential PET and CT scanners (note depth of machine bore)

common usage in clinical practice is Technetium-99m (99m-Tc), which emits 140 keV gamma ray photons [1]. Radiotracers may either emit photons directly, or emit particles which then decay and release photons. In either case, these photons are then detected by varying geometry detector configurations, and are used to produce cross sectional and volumetric images. Molecular imaging detectors commonly used in clinical medical practice include gamma ray detectors used in Positron Emission Tomography (PET) and Single Photon Emission Computed Tomography (SPECT) scanners (Fig. 17). The soft tissues of the body have a low attenuation coefficient for gamma rays, so the gamma photons pass through and escape the body, and are detected by the scanner outside the patient. Emitted photons from the normal regions of physiologic radiotracer distribution are used to create a rough projection of the body anatomy. Foci of abnormal radiotracer accumulation may be identified superimposed over the expected normal distribution, indicating regions of pathologic processes. After imaging scan completion, the complexed radiotracer molecules decay toward a stable state where they are no longer radioactive, and in general as well, are physiologically excreted from the body.

An advantage of molecular imaging is the ability to integrate information on the spatial distribution of physiologic processes within the body tissues with quantitative data from these processes. However, nuclear medicine imaging is lower in spatial resolution as compared to CT and MRI, a disadvantage. Other disadvantages of nuclear medicine imaging include the requirement for highly specialized equipment or services to obtain radionucides for imaging, and ionizing radiation

exposure. To overcome the spatial resolution limitations of molecular imaging, multi-modality hybrid scanning instruments are now in clinical usage, and include PET/CT, SPECT/CT, and PET/MRI scanners. High spatial resolution anatomic data from MRI or CT scanners are integrated synergistically with molecular imaging data from physiologic processes within the same scan body volume, to create a more complete picture of the pathology of the tissue and body organ of interest.

6 SPECT/CT

Data regarding the spatial architectural features of body anatomy such as size, density, boundary features, and texture are obtained via CT scanning, allowing potential diagnosis and (partial) characterization of the pathologic tissue. As such, CT scanning helps to localize regions of organ pathology within the body, and assess crossing of pathology through anatomic planes.

However, as noted, the information obtained from the CT scan regarding the pathology is in general incomplete, in the sense that in a number of cases, pathology of differing cellular etiologies may be indistinguishable on the basis of their CT appearance, or CT may be limited in its ability to even visualize the pathology. In some cases, metabolic data from the tissues of interest may allow pathologic distinction for more specific diagnoses. In the case of SPECT scanning, a 3-dimensional map of the distribution of radiotracer in the body is generated. However, while SPECT provides information regarding physiologic processes, the data provided is relatively low resolution with blurring of spatial landmarks, limiting accurate localization within the body anatomy. Photons emitted as part of the SPECT imaging process are subject to scattering and absorption as they pass through the body tissues to reach the detector, leading to attenuation artifacts. The resulting distortion of the SPECT data can lead to imaging appearances of irregular signal attenuation and spatial misrepresentation. Diagnostic accuracy of these scans can be improved using co-registered CT maps of the anatomy obtained as part of the SPECT/CT imaging process, with the additional benefit of correction for photon attenuation on the SPECT portion of the scan. Clinical medical uses of SPECT/CT in imaging the spine may include staging of malignant disease, anatomic localization and assessment of infection (Fig. 18) [16, 17].

7 PET/CT

The most common radiotracer currently used in PET/CT scanning is fluorine-18-fluorodeoxyglucose (18F-FDG), a glucose analog which is metabolized by the body. The 18F-FDG radiotracer complex passes across the body tissue cell membranes, and is thereafter entrapped within the cells. This radiotracer is preferentially

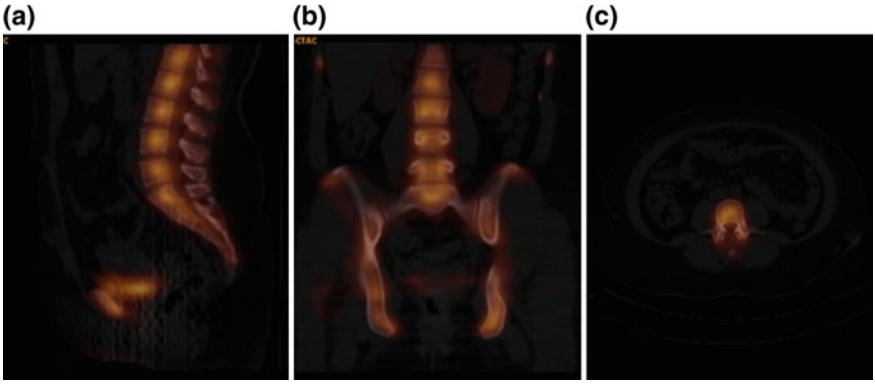


Fig. 18 SPECT/CT images of the spine. Sagittal (**a**), coronal (**b**), and axial (**c**) plane fused CT (*grayscale*) and SPECT (color over-projection) images from a SPECT/CT study of the spine obtained for assessment of possible metastatic bone lesions in this patient with a history of metastatic cancer and back pain (no metastatic lesions are apparent on these images)

taken up by cells undergoing increased glycolytic enzyme activity and increased glycolysis, seen in pathologic cells such as tumor cells, and results in localized foci of abnormally increased radiotracer accumulation within the tissues. The radionuclide portion of the molecule is fluorine-18, which emits positrons as part of its decay process. The emitted positrons then encounter local neighborhood electrons and undergo annihilation, shortly after being produced. The annihilation of the electron-positron pair leads to release of two oppositely directed gamma rays, which are then absorbed and detected in the PET/CT scanner [16, 18]. The two oppositely directed gamma rays are detected along a coincidence line, resulting in improved spatial localization data relative to SPECT, and so improved spatial resolution on the generated images. These coincidence detections are used to localize the pathologic process.

PET/CT is in widespread use in the detection, staging, and monitoring of cancer, and has also shown promise in infection imaging. The roles of PET/CT continue to expand with the development of new targeted radiotracer molecules. Figure 19 shows example images from a patient obtaining a PET/CT scan, extending from the head to the thighs, without evidence of active spinal metastatic disease. Images obtained from a whole body (head through thighs) PET/CT scan for a 52 year old female patient with a history of metastatic lung cancer as shown in Fig. 20.

8 Ultrasound

Sound waves of frequency range 2–10 MHz are used in clinical medical ultrasound applications, and are high frequency compared with the spectrum human hearing ranging 15–20 kHz [1]. The velocity of sound propagation varies depending on the

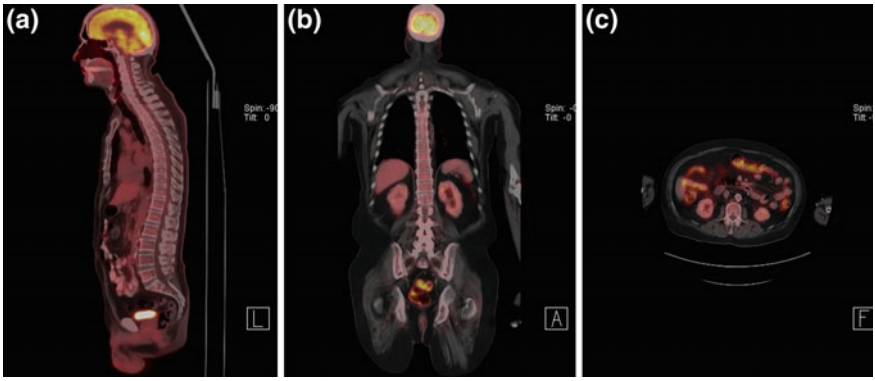


Fig. 19 Example of PET/CT images obtained extending from the head to the mid-thighs. Fused CT and PET images in **a** sagittal, **b** coronal, and **c** axial projections. CT portion of the scan is in grayscale, with color over-projection of physiologic (PET) activity. Areas of *light yellow* to *orange* represent areas of increased radiotracer uptake, and areas of *red*, lower uptake. Note that there is a normal physiologic distribution of radiotracer, with increased uptake in the brain, and increased activity also seen in the bowel and bladder, not to be mistaken for pathologic uptake

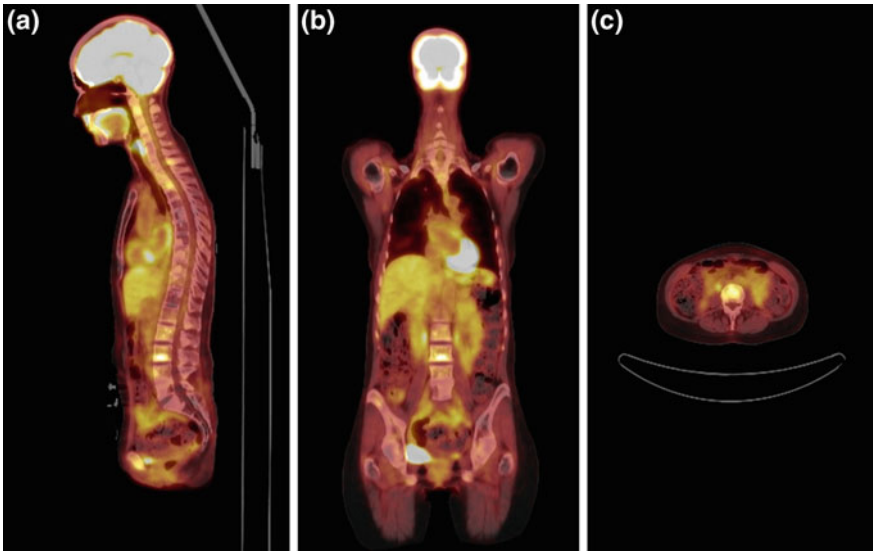


Fig. 20 PET/CT images obtained extending from the head to the mid-thighs. Fused CT and PET images in **a** sagittal, **b** coronal, and **c** axial projections. Areas of *light yellow* to *orange* again represent areas of increased radiotracer uptake, and areas of red, lower uptake. Note that there is a normal increased and expected physiologic uptake of radiotracer, in the brain and heart, with increased activity also noted in the liver, kidneys. This case, however, also demonstrates foci of abnormally increased radiotracer uptake in the spine, diagnosed as metastatic disease to bone. (compare Fig. 19)

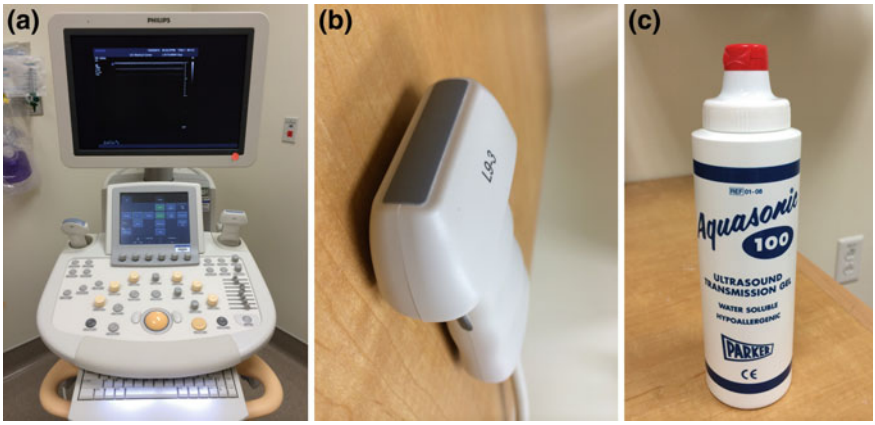


Fig. 21 Abdominal CT images. **a** Portable ultrasound scanner. **b** Ultrasound scanning probe (source and receiver), and **c** conducting gel to facilitate sound wave transmission from the probe into the body

physical characteristics material through which the sound waves are travelling, and this property is used to form images in the ultrasound examination. To perform an ultrasound examination, an ultrasound probe which acts as a source of high-frequency sound waves is put in contact with the body surface (Fig. 21). Air gaps intervening between the probe tip and the body surface are relatively poor conductors of ultrasound waves, so a sound conducting gel between the probe tip (source/receiver) and the skin to facilitate transmission of the sound waves into the body tissue. As the sound waves propagating through the tissues encounter transition points, or boundaries, between different internal body structures, some of the sound is reflected back to the probe, and some continues to propagate. The high-frequency sound waves that are reflected back are then detected by the same ultrasound probe which acted as the source. The ultrasound scanner uses the amplitude and return time of the waves which were reflected to construct a rendering of the body anatomy encountered by the sound waves as they propagated through within the body. Real-time visualization of movement of anatomic structures is also possible with ultrasound, as well as the movement of fluid. Using the Doppler effect, ultrasound can measure fluid movement direction and speed, including variation with time, through cine sequences [20]. At the current time, ultrasound is primarily used as a targeted modality, with each image used to look at specific tissue structures with a relatively small field of view (as opposed to CT, which can be used visualize the entire cross-section of the body on each image) Benefits of ultrasound are that it does not expose the patient to ionizing radiation, it is portable, can be used for noninvasive visualization of dynamic internal processes of the body. However, drawbacks (at the current time) can be significant, and include significant dependence of image quality on the operator performing the study, image resolution which is relatively low, and for our purposes in visualization of spinal structures, limited penetration of osseous

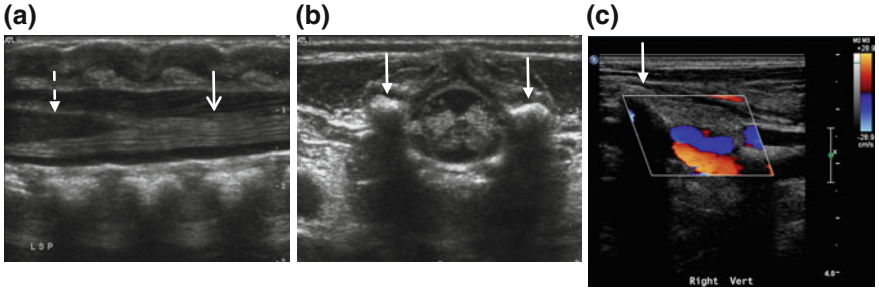


Fig. 22 Ultrasound images. Ultrasound images of the spine. Sagittal **a** and axial **b** image of the lower spine showing the bone (*solid arrows*), spinal canal, caudal aspect of the spinal cord (*dashed arrow*), and the cauda equina nerve roots (*open arrow*), and **c** Doppler ultrasound longitudinal image at the level of the cervical spine showing the vertebral bony spinal elements (*shadowing*) and color indicating flow in the vertebral artery

structures. Figure 21 shows a typical ultrasound scanner. Examples of ultrasound images are shown in Fig. 22, including images to examine lower spine in a newborn infant, and color Doppler to show blood flow in a vertebra artery segment in the cervical spine.

9 CAD Computer Systems and Image Quality

Once formed on the medical imaging scanners, these electronic images of varying modalities are securely transmitted to a computer network system for storage, display, manipulation, and analysis. These medical image data systems consist of multiple components, and are termed picture archiving and communication systems (PACS), and include software subroutine libraries for image storage and handling [21]. A simplified view of the main components of the PACS is that of a central server, with element components of archiving capability, viewing workstations, image processing workstations, and ancillary equipment for study distribution through removable media, typically via CD/DVD burning or printing of study images to photographic film. A current generation diagnostic workstation is shown in Fig. 23.

Medical imaging examinations typically consist of multiple images, grouped into series based on some common imaging characteristic, such the reconstruction plane, or MR imaging sequence. The images in these studies are stored electronically in Digital Imaging and Communications in Medicine (DICOM) format on the PACS, a standardized file format developed by National Electrical Manufacturers Association (NEMA) and the American College of Radiology (ACR) for to facilitate communication of images between equipment of different manufacturer origins [22–24]. The DICOM software integration standard specifies the representation and network transmission of image data, as well as of associated metadata and

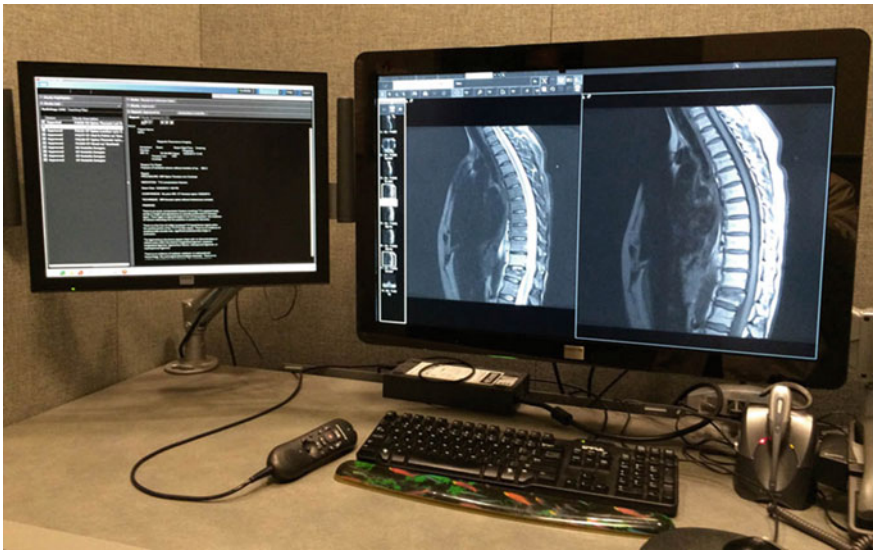


Fig. 23 Clinical viewing workstation for diagnostic interpretation of studies. Diagnostic workstation displays patient's DICOM images with associated relevant clinical medical history submitted by the ordering physician, as well as prior (interpreted and archived) images and reports. Basic real time manipulation of the images may be performed at the workstation. Handset (*bottom left*) and headset (*bottom right*) are seen, for dictation of study report through voice recognition system. Sample MRI examination of the spine is shown

informational objects. Diagnostic interpretative data, rendered by the physician reviewing the study, is stored on the PACS in association with the image data in voice or text format.

In the past, medical images such as radiographs were obtained on sheets of photographic film, with data regarding pathology then retained in analog format, both spatially, and in terms of detected energy translated into gray levels in the photographic emulsion. Current generation imaging scanners quantize diagnostic information in the detection process, and generate digital images as output, which are stored as pixel matrices. As noted previously, grayscale presentation of data is most common, with bit depths of 10 or 12 typically found on the diagnostic workstations. Color imaging is used for specialized diagnostic visualization, such as multimodal imaging (PET/CT) or Doppler ultrasound, where two data sets are superimposed. Matrix size generally varies by modality, typically can range from 256 × 256 to 1,024 × 1,024 for CT and 1,024 × 1,024 to 2,048 × 2,048 for radiography.

Other integrated computer systems in the typical diagnostic imaging department and hospital are termed the Radiology Information System (RIS) for storage, editing, and transmission of patient radiological data, and the Hospital Information System (HIS), a multifunctional information management system more broadly incorporating multilevel patient clinical information as well as other functions.

10 Image Quality

Image quality is a function of the interplay between the image noise, contrast, and spatial resolution. For purposes of brevity and illustration, we will limit our discussion here to CT and MR imaging modalities.

In an ideal medical image, tissue which is homogeneous in density or signal intensity would have a uniform gray-scale appearance. However, "real" images contain noise, arising from a number sources including thermally based radiofrequency energy emission in MRI, scattering of X-ray photons in CT, and due to detector noise. This noise is represented as fluctuations of pixel intensity units about some intensity average for a homogeneous tissue. The ability to assess the object or structure of interest (displayed in the image by detection of the signal from the object of interest) may be limited by the noise, which degrades the image quality and also the object representation. Thus, the magnitude of the noise relative to the signal of the image is an important concept in medical image interpretation and analysis. The power of the image signal is derived from the square of the average pixel intensity. The image noise is determined from the standard deviation of the pixel intensities in a region of the image with homogeneous signal intensity. The ratio of the power of the image signal to the power of the noise is the signal-to-noise ratio (SNR) [25].

For a particular imaging modality, the signal-to-noise ratio may be improved by adjustment of scanning parameters for a particular machine, or by the use of a machine with different inherent technology. For instance, the SNR for magnetic resonance imaging may be increased by using a machine with a higher intrinsic magnetic field, which may be accomplished by using a machine with a 3.0 T field strength versus 1.5 T. SNR may also be increased in a particular machine by adjustment of scanning parameters, such as decreasing the matrix size, increasing the section thickness, decreasing the RF bandwidth, and increasing the number of image acquisitions [26, 27]. In CT imaging, the SNR may be increased by increasing the mA, kVp, voxel size, and scan time [26].

The image contrast in medical images may be thought of as the difference in pixel intensity between different anatomic structures. In CT imaging, the contrast is in general increased by administering iodine-based contrast media (for soft tissues), and decreasing the kVp. MR imaging contrast is connected to the magnetic properties of the body tissues being imaged with the particular MR sequence being used, such as T1, versus PD, versus T2. As with CT imaging, the contrast of the MR images may be increased by the administration of contrast material, in this case, gadolinium based materials.

The spatial resolution of an image is related to voxel size, and is determined by the ability to discriminate adjacent anatomic structures. When performing a CT scan of a patient, the in-plane spatial resolution (axial plane, perpendicular to the long axis of the scanner) may be increased by decreasing the field of view and

increasing the matrix size. Through-plane (sagittal and coronal plane) resolution may be increased by decreasing the section thickness. In MR imaging, the matrix size and the field of view also affect spatial resolution, in the same manner as with CT imaging. Additionally, increasing the number of phase encoding steps and applying stronger gradients can increase spatial resolution.

The contrast resolution and spatial resolution of an image are linked to the signal-to-noise ratio. A contrast oriented quantity related to the SNR is the contrast-to-noise ratio (CNR). The CNR is determined by the difference between the signal of interest (foreground) and intensity of the surrounding structures (background), divided by the noise power [1, 25]. The contrast-to-noise ratio gives a measure of the ability to distinguish a structure of interest from surrounding structures.

Finally, image quality may be degraded by various modality based artifacts. As noted in prior sections, common to both CT and MR imaging, metallic surgical hardware can cause artifact, termed susceptibility artifact on MRI and streak artifact on CT. Other types of artifact on CT include partial volume artifact, caused by the finite spatial size of voxels in tissue with high spatial frequency variation in X-ray beam attenuation properties, and patient motion artifact. MR imaging is more sensitive to motion given the longer time interval for scanning of most MRI sequences relative to CT, and results in ghost images if the patient moves. Additionally, motion artifact may result from internal or physiologic motion, such as cerebrospinal fluid (CSF) pulsation flow artifact in the spinal canal due to flow of the CSF, and respiratory motion artifact. If the imaging field of view is smaller than the structure being imaged, aliasing artifact may occur, with portions of the structure anatomy outside the field of view are mapped to the opposite side of the image.

11 Conclusion

Medical imaging is a key component of many areas of modern clinical medical practice and biomedical research. Physicians are able to diagnose and form treatment plans for numerous pathologies on the basis of data provided by radiographic, CT, MRI, molecular imaging, and ultrasound studies. Additionally, new areas of research are emerging based on data provided by imaging, and conversely, imaging is used as an investigative tool in a variety research areas. The medical imaging field has been and is continuing to make great progress both in expanding clinical medical applications and at the frontiers of research. Linked with advanced image processing and visualization techniques, it has the potential to continue to open frontiers in the development new and advanced diagnostic and treatment models.

References

1. Bushberg JT (2002) The essential physics of medical imaging, 2nd edn., vol. xvi. Lippincott Williams & Wilkins, Philadelphia
2. Brant WE, Helms CA (eds) (1999) Fundamentals of diagnostic radiology, 2nd edn. Lippincott Williams & Wilkins, Philadelphia
3. Herman GT (2009) Fundamentals of computerized tomography: image reconstruction from projections, vol xi, 2nd edn. Springer, New York (Adv pattern recognition Dordrecht)
4. Gunderman R (2006) Essential radiology: clinical presentation pathophysiology imaging, 2nd edn. Thieme, New York
5. Brooks RA (1977) A quantitative theory of the Hounsfield unit and its application to dual energy scanning. J Comput Assist Tomogr 1(4):487–493
6. Lauterbur PC (1973) Image formation by induced local interactions: examples of employing nuclear magnetic resonance. Nature 242:190–191
7. Fullerton GD (1982) Basic concepts for nuclear magnetic resonance imaging. Magn Reson Imaging 1:39–55
8. Kaplan PA et al (2001) Musculoskeletal MRI. Saunders, Philadelphia
9. Stoller DW (2007) Magnetic resonance imaging in orthopedics and sports medicine, vol 1. Amirsys, Salt Lake City
10. Berquist TH (2012) MRI of the musculoskeletal system, 6th edn. Lippincott Williams & Wilkins, Philadelphia
11. Hargreaves BA et.al (2011) Metal-induced artifacts in MRI. 197(3):547-555
12. Schaefer PW (2000) Diffusion-weighted MR imaging of the brain. Radiology 217(2):331–345
13. Dong Q (2004) Clinical applications of diffusion tensor imaging. J Magn Reson Imaging 19 (1):6–18
14. Alexander AL (2007) Diffusion tensor imaging of the brain. Neurotherapeutics 4(3):316–329
15. Ross JS (2010) Diagnostic imaging spine, 2nd edn. Amirsys, Salt Lake City
16. Mettler FA (2006) Essentials of nuclear medicine imaging, 5th edn. Elsevier, Philadelphia
17. Mariani G et al (2010) A review on the clinical uses of SPECT/CT. Eur J Nucl Med Mol Imaging 37(10):1959–1985
18. Kapoor V, McCook BM, Torok FS (2004) An introduction to PET-CT imaging. Radiographics 24:523–543
19. Alavi A (ed) (2010) Musculoskeletal PET imaging. PET Clinics 5(3):247–390 (ScienceDirect)
20. Rumack CM (2010) Diagnostic ultrasound, 4th edn. Elsevier/Mosby, Philadelphia
21. Choplin R (1992) Picture archiving and communication systems: an overview. Radiographics 12(1):127–129
22. http://medical.nema.org/
23. Kahn CE Jr et al (2011) Informatics in radiology: an information model of the DICOM standard. Radiographics 31(1):295–304
24. Bidgood WD Jr et al (1997) Understanding and using DICOM, the data interchange standard for biomedical imaging. J Am Med Inform Assoc 4(3):199–212
25. Dhawan AP (2011) Medical image analysis, 2nd edn. IEEE/Wiley, Hoboken
26. Huda W (2009) Review of radiologic physics, 3rd edn. Lippincott Williams & Wilkins, Philadelphia
27. Allisy-Roberts P, Williams J (2007) Farr's physics for medical imaging. W.B. Saunders Company, Philadelphia

Arthritis of the Spine

Runsheng Wang and Michael M. Ward

Abstract Arthritis is the common term used to describe pathological changes of joints and adjoining parts of the bone. Several types of arthritis commonly affect the spine. Osteoarthritis, a non-inflammatory type of arthritis, most often affects the cervical spine and the lumbar spine. Neck pain, limited neck and head motion, low back pain, and limited flexibility of the low back can result from progressive joint damage. Degeneration of the intervertebral disk may accompany cervical and lumbar osteoarthritis, and can cause either nerve root or spinal cord compression. Ankylosing spondylitis is the most common inflammatory arthritis that principally affects the spine rather than other joints, and is characterized by slow development of bony fusion among the adjacent vertebrae. Rheumatoid arthritis, the most common type of inflammatory arthritis, affects mostly the limb joints but can also affect the cervical spine, causing neck pain and headache. Cervical spine arthritis also often occurs in children with juvenile idiopathic arthritis. Radiography is an essential diagnostic tool in the evaluation of patients with spinal arthritis, but provides limited information on the posterior spinal structures. Magnetic resonance imaging can be useful for defining abnormalities in the posterior spinal joints, the nerve roots, and the spinal cord.

This work was supported by the Intramural Research Program, National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health.

R. Wang · M.M. Ward (✉)
Intramural Research Program, National Institute of Arthritis and Musculoskeletal and Skin Diseases/National Institutes of Health, Building 10 CRC, Room 4-1339, 10 Center Drive, Bethesda, MD 20892, USA
e-mail: wardm1@mail.nih.gov

© Springer International Publishing Switzerland 2015

S. Li and J. Yao (eds.), *Spinal Imaging and Image Analysis*,
Lecture Notes in Computational Vision and Biomechanics 18,
DOI 10.1007/978-3-319-12508-4_2

1 Overview

1.1 Spine Anatomy and Classification of Arthritis

Vertebrae are the bony structure of the spine. The anterior parts of the vertebrae, the vertebral bodies, are interconnected by the intervertebral disks. The outer layer of the intervertebral disk is a dense fibrous tissue, called the annulus fibrosis. The annulus fibrosis and its adjacent vertebral body form the discovertebral joint. The posterior part of the vertebra consists of the neural arch. At each vertebral level, the inferior processes of one vertebra articulate with the superior processes of the vertebra immediately below it, forming a facet joint (also known as the zygoapophyseal joint) on both the right and left sides. The neural arch forms the spinal canal, containing the spinal cord, spinal nerves, spinal membranes, and spinal fluid. There are openings on the side at each vertebra level, called the intervertebral foramen. The spinal nerves, originating from the spinal cord, exit the spinal canal through the foramen. The vertebral bodies are lined with ligaments on the front and back that provide stability. At the neural arch, the ligamentum flavum connects vertebrae to each other.

Arthritis is the common term used to describe pathological changes to joints and their associated structures, including bones, cartilage, and ligaments. Depending on whether the immune system is involved or not, arthritis can be categorized into two main subtypes: inflammatory and non-inflammatory. Several types of arthritis can affect the spine. Osteoarthritis, a degenerative process often associated with aging, is the most common type of non-inflammatory arthritis. The cervical spine and the lumbar spine are often affected in osteoarthritis. When the spine is affected by osteoarthritis, it is termed spondylosis. Degenerative changes occur in the intervertebral disks as well, causing degenerative disk disease. Structural changes of the vertebrae and the disks may lead to compression on the spinal cord or the spinal nerve roots, causing a condition known as spinal stenosis. Rheumatoid arthritis is the most common type of inflammatory arthritis. Although it primarily targets joints of the extremities, rheumatoid arthritis may affect the cervical spine. In children, juvenile idiopathic arthritis may damage the cervical spine in a similar fashion. Ankylosing spondylitis is the prototypic inflammatory arthritis that primarily involves the spine and the sacroiliac joints, with a prominent feature of slow development of bony fusion among the adjacent vertebrae.

1.2 Uses of Imaging in Diagnosis, Prognosis, Treatment, and Assessment of Treatment Response

Plain radiography is the essential diagnostic tool for spinal arthritis, not only because it can demonstrate much of the relevant pathological changes in the spinal structures, but also because it is widely available and inexpensive. However, plain

radiography provides limited information about the posterior spinal structures and the intervertebral disk pathology. Magnetic resonance imaging (MRI) and computer tomography (CT), as three-dimensional imaging tests, provide superior structural information and a better resolution. MRI is ideal for visualizing pathology of the intervertebral disk, neural structures such as the spinal cord, and is often the spinal imaging test of choice for patients with neurological symptoms. MRI is also useful to depict inflammatory changes in bones and soft tissues. Therefore it has gained interest as a method to assess the treatment response in inflammatory spine diseases. CT is most often used in patients who have contraindications to MRI. CT is also an ideal imaging modality for bony structures. CT myelography remains the gold standard for diagnosing the cause of nerve root compression, differentiating osteophytes from disk pathology. Spinal CT is also used in research settings to assess the progression of bony pathology in ankylosing spondylitis.

2 Osteoarthritis

2.1 Cervical and Lumbar Spondylosis

2.1.1 Definition and Occurrence

Spondylosis refers to degenerative arthritis of the spine, including osteoarthritis of the discovertebral and facet joints, and degenerative changes of related soft tissues, including surrounding ligaments and muscles.

The cause of spondylosis remains unclear. A widely cited hypothesis states that degenerative changes begin with the loss of water content in the annulus fibrosis [1]. The annulus gradually becomes drier and weaker, and eventually the disk content leaks out, resulting in intervertebral disk protrusion and narrowing of the disk space. This subsequently leads to increased mechanical stress at the discovertebral joints, the facet joints and the spinal ligaments, causing both bony overgrowth and ligament thickening. Bony growths at the front and side of the vertebral bodies are commonly seen. These so-called marginal osteophytes originate from the end plate of the vertebral body. At the microscopic level, the cartilage endplate degenerates and is replaced by bony proliferation; over time, it becomes hard and protrudes into the intervertebral disk and the edge of the vertebral body [2]. Similar changes occur at the facet joints, with overgrowth of bone (osteophytes) and narrowing of the joint spaces. The spinal ligaments, especially the ligamentum flavum, become thickened and may eventually calcify. These degenerative changes are most commonly found at the fifth cervical, eighth thoracic and third lumbar spinal levels, possibly due to greater spinal flexibility in these areas [3]. Progression of degenerative changes may lead to compression of the adjacent structures, particularly on the spinal cord and/or the spinal nerve roots, causing spinal stenosis. This condition is discussed in detail below.

Degenerative changes of the spine are found to be present as early as age of 15, but symptoms only develop in much older individuals [4]. In a community-based study in the United States, facet joint osteoarthritis was found on CT scan in 36 % of people younger than 45 years old, in 67 % who were 45–64 years old, and in 89 % who were older than 65 years [5]. A large epidemiologic study in Japan reported the prevalence of radiographic lumbar spondylosis as 75.8 % in people older than 60; however, only 28.8 % of these people had symptoms of low back pain [6]. Aging and trauma are the main risk factors for developing spondylosis. No associations have been established with other conditions, such as lifestyle, height, obesity, physical activity, smoking and alcohol use.

2.1.2 Clinical Manifestations

Patients may present with a wide spectrum of symptoms. The majority of patients with spondylosis do not have any symptoms, even with advanced changes on radiographs.

In symptomatic patients, pain is the most common complaint. It may present as acute episodes, or may be chronic. In some patients, pain is caused by osteoarthritis of the facet joints, called facet joint syndrome. In lumbar facet joint syndrome, pain travels down to the buttock and the back of the thighs, and typically improves with bending forward and worsens with bending to the affected side. In cervical facet joint syndrome, patients often complain of neck pain traveling along the spine, the shoulder blades and the back of the head.

Limited motion of the neck or the back may occur, especially when trying to extend the back or raise the head to look up. Osteophytes at the cervical facet joints sometimes compress the arteries and decrease the blood supply to the brain, causing dizziness.

Patients with spondylosis may have concurrent degenerative disk disease, or may progress to develop spinal stenosis. These conditions often present with neurologic symptoms, such weakness of the legs or the arms, numbness, or urinary and/or bowel dysfunction. These conditions will be discussed in detail in the following sections.

2.1.3 Treatment

Conservative management is the mainstay treatment for patients who do not have neurologic symptoms. Patient education on the natural history of spondylosis, self-care options and coping techniques is the first step. A long-term follow-up of patients with neck pain found that in 79 % patients, the pain resolved after 15 years without surgical intervention [7]. Immobilization of the cervical spine with a soft collar is often used, however its effectiveness is not proven. In patients with back pain, maintaining daily activity, instead of rest, is beneficial. Physical therapy, including mechanical traction and manipulation, are sometimes used. Exercise,

stretching, and muscle strengthening are also recommended. Heating pads or blankets may provide local relief. In general, evidence for these measures is not based on controlled trials.

Pain management often includes the use of medications. Acetaminophen and nonsteroidal anti-inflammatory drugs (NSAIDs) such as ibuprofen or naproxen are the first line treatments. In patients with severe pain who do not improve sufficiently with these treatments, opioids can be considered. Muscle relaxants may provide relief in acute pain episodes. However, they are often associated with side effects such as dizziness, so long term use is not recommended. Local injections with corticosteroids or anesthetics are sometimes used, with variable results.

Surgery is indicated for patients with progressive nerve symptoms and compression of the spinal cord or the spinal nerve roots. For neck pain or back pain without nerve compression, surgery is not recommended due to lack of effectiveness.

2.1.4 Imaging

Plain radiographs of the cervical spine and lumbar spine are typically adequate to reveal spondylosis. Because the correlation between symptoms and radiographic changes is poor, radiographs have limited usefulness in the evaluation of neck pain or back pain. In the absence of systemic symptoms such as fever or weight loss, history of trauma, or progressive nerve symptoms, radiographs are typically not obtained until after 6–8 weeks of conservative management.

However, radiographs may still provide valuable information. Osteophytes, narrowing of the intervertebral disk spaces, narrowing of the facet joints, sclerosis (increased radiographic density) of the facet joints and the endplates of the vertebral bodies, and narrowing of the neural foramen are common findings in spondylosis (Figs. 1 and 2). Radiographs are also useful to assess the alignment of the spine and to exclude other diagnoses. The Kellgren/Lawrence system was developed to classify the degree of osteoarthritic change in the spine, including the facet joints. Lateral views of cervical spine and lumbar spine are obtained for grading. Five features are considered in the Kellgren/Lawrence system: osteophytes, ossicles near the joints, narrowing of joint spaces with subchondral sclerosis, pseudocysts, and altered bone shape (Table 1). Radiographic changes are classified into five grades (0–4), with a grade of 2 or higher as the conventional standard of diagnosis [9].

In patients with progressive neurologic symptoms, or in patients with persistent pain and severe radiographic spondylosis, MRI is the imaging test of choice. It provides a better resolution for structural changes, and is ideal for visualization of the spinal cord, the intervertebral disk, and the soft tissues.

CT is superior for detection of bony changes, especially small osteophytes or erosions arising from the lateral edge of the vertebral body and the facet joints (Fig. 3). However, because of the exposure to radiation, it is only used in patients

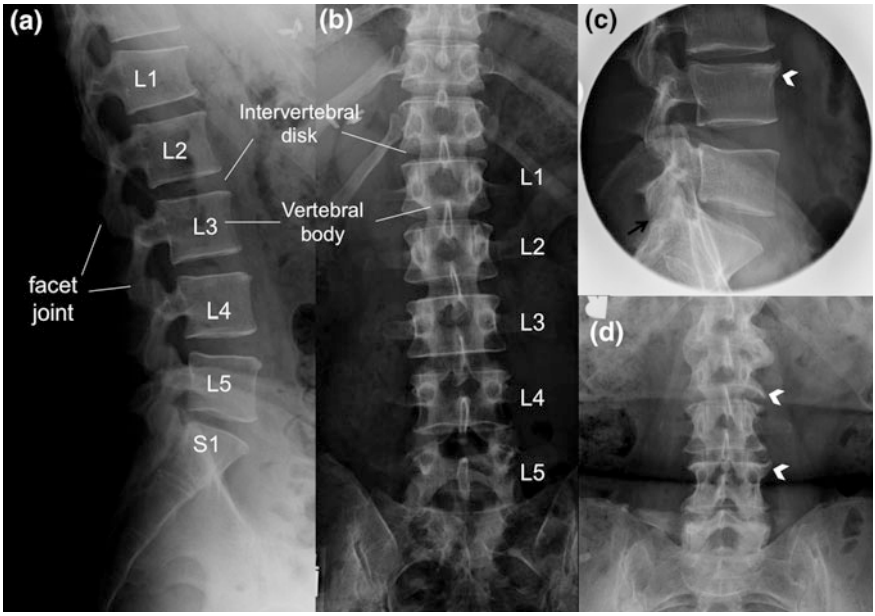


Fig. 1 Plain radiograph of a normal lumbar spine and of lumbar spondylosis. **a** Lateral view of a normal lumbar spine. **b** Anteroposterior view of a normal lumbar spine. **c** Lateral view of a patient with lumbar spondylosis. Marginal osteophyte (*arrowhead*) and narrowing of a facet joint (*arrow*) are present. **d** Anteroposterior view of a patient with lumbar spondylosis. Notice the marginal osteophytes (*arrowhead*) at multiple levels

with contraindications to MRI, and when establishing a firm diagnosis is needed to guide treatment.

As degenerative changes are common in older people and are not always symptomatic, the interpretation of MRI results (or CT with myelogram) should be cautious and correlated with the clinical findings for diagnosis and further management.

2.2 Degenerative Disk Disease

2.2.1 Definition and Occurrence

Degenerative disk disease is a group of conditions caused by wear and tear changes of the intervertebral disks. It is usually a part of the aging process; however, in rare conditions, accelerated degeneration occurs, causing juvenile degenerative disk disease.

The intervertebral disk is composed of the gelatinous nucleus pulposus in the center, surrounded by the annulus fibrosis, which is composed of layers of dense,

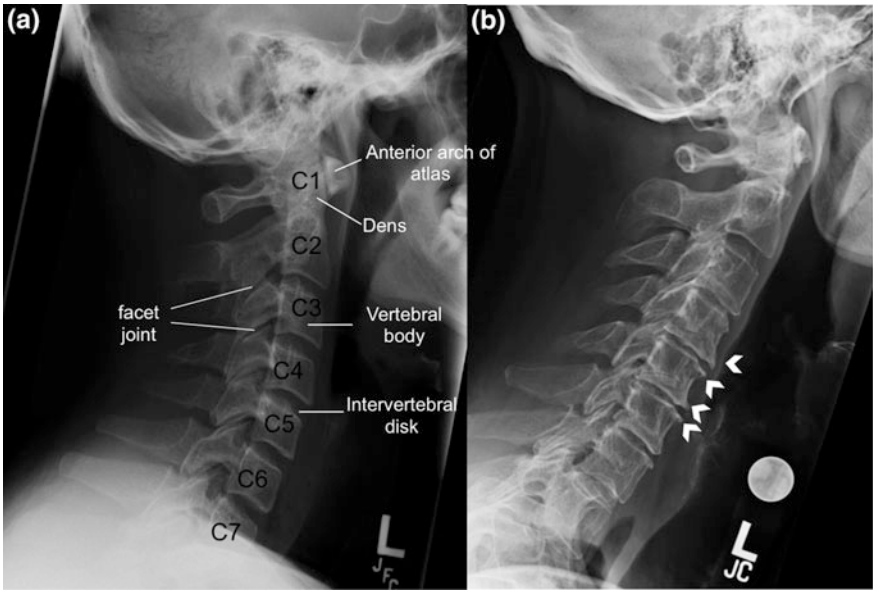


Fig. 2 Plain lateral radiographs of a normal cervical spine (**a**) and of a patient with cervical spondylosis (**b**). Marginal osteophytes are denoted by *arrowheads*

Table 1 Kellgren Lawrence grading system: osteoarthritis is divided into five grades as follows [8]

Grade 0	None
Grade 1	Doubtful
Grade 2	Minimal
Grade 3	Moderate
Grade 4	Severe

fibrotic tissue. With aging, the disk undergoes three phases of degenerative changes [10]. In phase I, or the dysfunctional phase, microtrauma from repetitive use causes small tears and fissures in the annulus fibrosis, associated with pain. Meanwhile, the nucleus pulposus loses water content. MRI study often reveals disk bulging without herniation and tears in the annulus. In phase II, or the unstable phase, more tears occur and lead to disk disruption, resorption, and loss of the disk space. Local inflammation may follow if the herniated disk compresses the spinal nerve root. Cartilage degeneration and malalignment can develop in the facet joints. Clinically, patients often present with spine instability and symptoms related to nerve irritation. Phase III is the stabilization phase. With disk resorption and disk space narrowing, mechanical stress leads to fibrosis of the disk and degenerative changes at the vertebral endplates.

The prevalence of degenerative disk disease is difficult to estimate. In a MRI study in 239 asymptomatic individuals with a mean age of 39 years, degenerative cervical disk disease progressed in 81 % of the study subjects over 10 years [11].

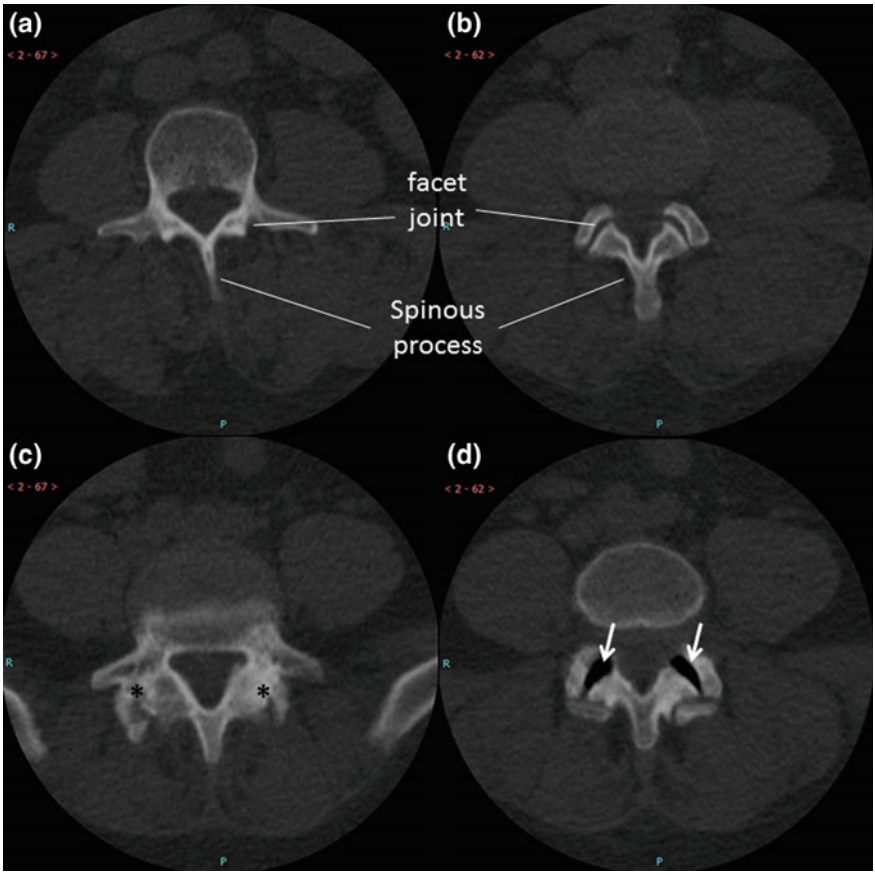


Fig. 3 Facet joint osteoarthritis on computed tomography. **a**, **b** Axial view of a normal lumbar spine. **c**, **d** Axial view of a lumbar spine with facet joint osteoarthritis. Hypertrophy (*asterisks*) and erosions (*arrows*) are common findings in facet joint osteoarthritis

In patients with cervical radiculopathy, disk protrusion was identified as the cause in 21.9 % of patients [12]. In a study of cervical MRI scans of patients undergoing throat surgery who had no neck pain, cervical disc protrusion or herniation was incidentally seen in 20 % of patients aged 45–54, and 57 % of those older than 64 [13]. Lumbar degenerative disk disease affects young to middle aged people as well, with a peak incidence at age 40 years. A recent study using MRI of the whole spine of 975 individuals found degenerative disk disease in 71 % of men and 77 % of women younger than age 50 years, and in more than 90 % of men and women older than age 50 years [14].

2.2.2 Clinical Manifestations

Degenerative disk disease may present a wide spectrum of symptoms, and MRI studies have shown that the degree of disk degeneration does not correlate with the symptoms [15, 16]. The majority of patients with degenerative changes on MRI remain asymptomatic for years [11].

Common symptoms include pain and nerve dysfunction such as radiculopathy. Radiculopathy is caused by the compression of a spinal nerve root from a laterally herniated disk. It is one of the most common causes of acute pain of the neck or the back. In cervical disk herniation, pain usually affects the arms, shoulders, the region between the shoulder blades, or the rib cage, and can mimic chest pain. Persistent compression will lead to numbness, tingling, and weakness of the arms or the hands, in the areas supplied by the compressed spinal nerve. In one series, 70 % of patients with cervical radiculopathy were found to have lesions at the disk between the sixth and seventh cervical vertebrae (C6–C7), and 20 % of patients had lesions at the C5–C6 intervertebral disk [17].

In lumbar disk herniation, pain is usually in the low back, travelling to the buttocks or down the leg to below the knee. Pain is usually worsened with bending forward, coughing, or sneezing. Numbness, tingling, and weakness of the leg may occur. The straight leg test is a physical exam test used to assess lumbar disk herniation, with specificity of 89 % and sensitivity of 52 % [18].

A large central disk herniation may cause compression of the spinal cord, with neurological symptoms compatible with myelopathy or cauda equina syndrome in the lower lumbar spine. Detailed clinical manifestation will be discussed in the section on spinal stenosis.

Disk pain is caused by irritation on the annulus fibrosis. It comprises neck pain or back pain, extending along the spine, without associated neurologic symptoms. A diagnosis of discogenic pain is based a fluoroscopic provocative test, and will be discussed in detail in the imaging section.

2.2.3 Treatment

No treatment is needed for asymptomatic patients. For patients with pain along the spine without nerve symptoms, conservative management is appropriate. Most patients have a benign course. In a study with 19 years of follow up, 75 % of patients had only one episode of pain or mild recurrent symptoms [7]. Pain management with NSAIDs, muscle relaxant during acute episodes, exercise, and cervical or lumbar traction, are commonly used.

In patients with progressive neurologic symptoms, MRI is indicated. If compression of the spinal nerves or nerve roots by a herniated disk is found to be the cause of symptoms, local injection of corticosteroids is often used, especially in cases of lumbar disk herniation. Without treatment, persistent compression may

lead to permanent damage of the nerve, and lead to irreversible nerve symptoms. If compression of the spinal cord is present with clinical symptoms of myelopathy, definitive surgery is indicated.

2.2.4 Imaging

Plain radiographs of the spine have limited use in diagnosing degenerative disk disease, although some radiographic findings indicate degenerative disk disease. These findings include narrowing or loss of the disk height (Fig. 4), sclerotic changes of the vertebral endplates, and, in later stages, the presence of osteophytes and sclerosis of the facet joints. "Vacuum phenomenon" is considered a specific radiographic indicator of disk degeneration. With degeneration of the disk, gases transpired from the circulation accumulate in the space that the nucleus pulposus once occupied, causing the intervertebral disk space to appear radiolucent (Fig. 4). In general, in the absence of trauma, radiographs are not always needed.

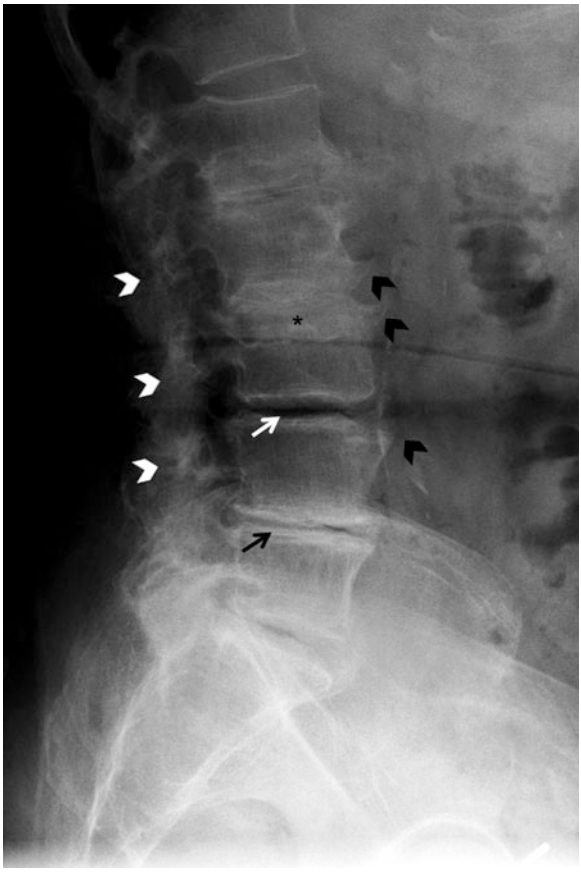


Fig. 4 Severe degenerative disk disease and lumbar spondylosis on plain radiograph. Lateral view of a lumbar spine. Narrowing of intervertebral disk space (*black arrow*), complete loss of disk space (*asterisk*), and vacuum phenomenon (*white arrow*) are characteristic features of degenerative disk disease. Osteophytes originating from vertebral bodies (*black arrowheads*) and facet joint narrowing (*white arrowheads*) are present, signifying associated spondylosis

Provocative discography may be used to confirm that a degenerative disk is the source of neck pain or back pain in difficult cases [19]. Under fluoroscopy, a diseased disk is injected at a certain pressure to see if this procedure reproduces the patient's pain. If the injection to an adjacent normal disk does not reproduce the pain, the test is confirmatory.

MRI is the standard imaging modality for detecting disk disease. It is indicated in patients who have progressive neurological symptoms despite conservative management, or in patients who plan to undergo surgery. On MRI, a degenerated disk has decreased intensity on T2 weighted images, due to loss of water content and glycosaminoglycans [20]. Bulging of the annulus (Fig. 5a), herniation of the disk contents (Fig. 5b), and loss of intervertebral disk height can be demonstrated on MRI. Early changes of disk degeneration, such as tears of the annulus fibrosis, can be seen as high intensity zone lesions [21, 22].

CT scan can depict degeneration, bulging and herniation of the disk, but with much less detail than MRI. CT can also show sclerotic changes of the vertebral endplate and loss of the disk height, which are commonly seen in degenerative disk disease, but these findings most often can be readily seen on plain radiographs. Clinically, CT is used in patients with contraindications to MRI.

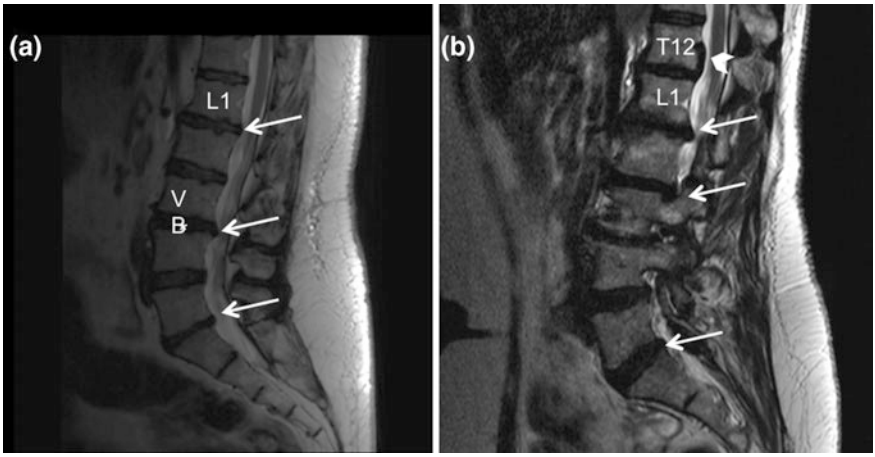


Fig. 5 Degenerative disk disease by magnetic resonance imaging. Sagittal view of a lumbar spine, T2 weighted images. **a** Disk bulging at multiple levels, most prominent at L1–L2, L3–L4 and L5–S1, and indenting the spinal canal (*arrows*). VB indicates vertebral body; *Asterisk* indicates intervertebral disk. **b** Disk herniation at the T12–L1 level, with migration of disk material posterior to the T12 vertebral body (*arrowhead*). Disk bulging (*arrow*) is present at L1–L2 and L2–L3 as well

2.3 Spinal Stenosis

2.3.1 Definition and Occurrence

Spinal stenosis is a condition of narrowing of the central spinal canal, causing compression on the structures within the canal, mainly the spinal cord and spinal nerve roots, with associated nerve dysfunction. The spinal cord extends from the base of the brain and ends at the level of the first and second lumbar vertebrae (L1–L2). Spinal nerves branch off the spinal cord and course alongside it before exiting the spinal canal. Below the L1–L2 level, the spinal nerves form a bundle called the cauda equina. Anatomically, when compression happens above the L1–L2 level, both the spinal cord and the nerve roots can be affected, while below the L1–L2 level, compression of the nerve roots alone is seen. Both direct mechanical compression and secondary changes due to lack of blood supply contribute to damage of the spinal cord and nerve roots. When the spinal cord is affected, it is called myelopathy, and when the spinal nerve roots are involved, it is termed radiculopathy.

Spondylosis is the most common cause of spinal stenosis in people older than 60 [23]. Osteophytes of the facet joints, disk bulging, and calcification and overgrowth of the posterior longitudinal ligament and ligamentum flavum can slowly encroach the spinal canal, and eventually lead to compression of the spinal cord or nerve roots. Spondylolisthesis, a condition in which the one vertebra slips relative its neighboring vertebra, can be a cause of lumbar spinal stenosis, especially at the L4–L5 level.

Conditions other than degenerative changes can cause spinal stenosis, including tumors and post-operative scar tissue. Inflammatory conditions, such as rheumatoid arthritis, may lead to overgrowth of the synovium at the facet joints, with compression of the spinal cord in severe cases. Some people are born with a narrow spinal canal and are susceptible to spinal stenosis with even minor changes in spine anatomy or mild degrees of degenerative disease. Spina bifida is another congenital cause of spinal stenosis.

Spinal stenosis usually has an insidious onset. It can be an incidental finding on radiologic study in asymptomatic individuals. It occurs in 20–30 % of people older than 60 years of age [24]. In a study of 187 individuals, the prevalence of lumbar spinal stenosis increased with age, affecting 2.1 % of people aged 40–49, 6.1 % of people aged 50–59, and 16.3 % of people older than 60 years [25].

2.3.2 Clinical Manifestations

In patients with cervical spinal stenosis, neck pain or pain in the area below the shoulder blades is frequently reported. When the narrowing and compression damage the cord, neurological symptoms develop. Clumsiness and weakness are common, and both the arms and legs may be affected. Many patients experience

loss of sensation, often associated with numbness and tingling. When the spinal cord is compressed, a sensory plane can be detected, separating the body into areas with normal sensation above the plane and areas without normal sensation below the plane. If a nerve root is affected, the sensory change is often distributed in the skin area supplied by the compressed spinal nerve, or in other words, in a dermatomal pattern. Patients may also experience difficulty with urination or having bowel movements. On physical exam, patients are found to have an abnormal gait early in the course of disease, indicating weakness of the legs. Neck motion is often limited. Lhermitte's sign is a characteristic finding, and can be induced by bending the neck. When this sign is present, patients experience a sensation of an electric shock in the neck, shooting down to the arms and along the spine.

Rarely, cervical spinal stenosis presents acutely. This can happen after minor injury or whiplash injury. These patients often have pre-existing degenerative changes, and a minor disturbance then leads to worsening and onset of nerve symptoms, with rapid progression of weakness, sensory changes, and bladder or bowel dysfunction.

The common clinical presentation of lumbar spinal stenosis was well characterized in a cohort of 68 patients [26]. Back pain, often travelling down the legs, numbness, and weakness of the legs are common. A prominent feature is neurogenic claudication, with worsening of symptoms on walking or standing, and relief when sitting or bending forward. On exam, patients are often found to have a wide-based gait. Weakness and sensory changes are distributed in one or more spinal nerve areas, indicating radiculopathy. Cauda equina syndrome is a rare complication of lumbar spinal stenosis, with weakness of both legs associated with urinary dysfunction. If spinal stenosis occurs higher in the spine than the L1–L2 level, damage of the spinal cord will cause myelopathy, with presentation similar to that of cervical spinal stenosis, but involving the legs.

2.3.3 Treatment

Conservative management is the mainstay treatment for spinal stenosis. In patients with cervical spinal stenosis, immobilization with a soft collar or a brace is often recommended. Activities such as action sports or intense neck movements should be avoided. Prevention of whiplash injury during motor vehicle accident is important. For patients with lumbar spinal stenosis, although evidence is lacking, exercise is recommended with a goal to strengthen muscles and to maintain correct posture. Pain control with acetaminophen and NSAIDs is commonly used, and can be escalated to opioids if needed. Epidural injection of corticosteroids is used in lumbar spinal stenosis, but with limited evidence supporting its effectiveness.

In some cases, compression can be relieved by surgery. However, the indications for surgery and its timing have not been well studied. Commonly, surgery is considered in patients with progressive nerve symptoms or moderate to severe symptoms with difficulty performing daily tasks [26]. In patients with spinal

stenosis but without neurologic symptoms, surgery can be deferred with close monitoring [26, 27].

Acute nerve symptoms may be the first presentation in some patients, and is a medical emergency. Immediate MRI is indicated for diagnosis and assessment of severity. Neurosurgery or orthopedic evaluation for potential surgical intervention is essential. Treatment with high dose intravenous corticosteroids to decrease acute inflammatory changes in the spinal cord may improve outcomes [28].

2.3.4 Imaging

The diagnosis of spinal stenosis is based on imaging and a compatible clinical presentation. Plain radiographs have limited utility for this condition. It is used in cases of neck pain or back pain without neurologic symptoms to exclude other conditions. In patients with nerve symptoms, MRI is the study of choice, while CT with myelography is used in patients with contraindications to MRI. Direct compression of the spinal cord can be visualized on MRI, and it may or may not be associated with a signal change in the spinal cord. Findings may be present in one or multiple vertebral levels.

Measurement of the anteroposterior diameter of the spinal canal or the intraspinal canal area has been suggested as radiologic diagnostic criteria of spinal stenosis [24], and for assessment of myelopathy [29, 30], however it has not been routinely used in clinical practice. More importantly, radiologic spinal stenosis is an incidental finding in 6–7 % of asymptomatic individuals, and its prevalence increases to 20–30 % in people older than 60 years [24].

Abnormal MRI signal in the spinal cord can be a useful marker of myelopathy (Fig. 6). Hyperintense signal on T2-weighted imaging, hypointense signal on T1-weighted imaging, and hyperintense signal on diffusion-weighted imaging (DWI) have been evaluated for their correlation with clinical findings, and DWI has a better correlation [31, 32].

2.4 Diffuse Idiopathic Skeletal Hyperostosis

2.4.1 Definition and Occurrence

Diffuse idiopathic skeletal hyperostosis (DISH) is a non-inflammatory condition characterized by calcification and ossification of ligaments, with a predilection for the spine. It most commonly affects the anterior longitudinal spinal ligament, particularly in the thoracic spine. Large flowing osteophytes with an appearance of ‘candle wax dripping down the spine’ is the typical finding in this condition. Thickening, calcification, and ossification may also involve peripheral ligaments, especially at sites of the tendon insertions. Unlike spondylosis, in which the primary pathologic target is cartilage, the discovertebral joints and the facet joints are

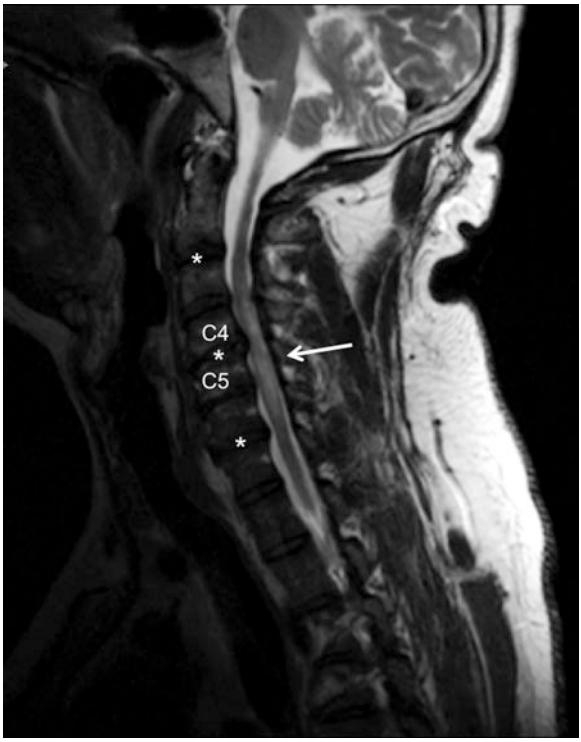


Fig. 6 Cervical spinal stenosis with spinal cord compression by magnetic resonance imaging. Sagittal view of a cervical spine, T2 weighted images. At the C4–C5 level, intervertebral disk bulge compresses the spinal cord, associated with hyperintense signal at the corresponding level (*arrow*), indicating structural damage of the spinal cord. *Asterisks mark* intervertebral disks

usually intact in DISH. Lack of sacroiliac joint involvement and inflammation distinguishes DISH from ankylosing spondylitis [33, 34]. The cause of DISH remains unclear.

DISH is rare in people younger than 40 years old. It generally affects people older than 50, with a prevalence of 15 % in women and 25 % in men. This prevalence increases to 26–28 % in those over 80 years [35].

2.4.2 Clinical Manifestation

DISH is largely asymptomatic. Patients may report pain in the spine and legs, morning stiffness, and limited spine flexibility. Pain in the upper back is common, and is often associated with limited chest expansion. The cervical spine and lumbar spine may also be involved. Although rare, in severe cases, large calcifications may impinge on the airway to cause difficulty or pain with swallowing, hoarseness, or high-pitched sounds from the throat with breathing. In the peripheral joints, calcification of the ligaments and entheses (sites where tendon attach to the bone) cause local pain, and can limit movement of the affected joints. Some patients have tenderness and nodules of the entheses.

2.4.3 Treatment

Pain control with acetaminophen and NSAIDs is the mainstay of treatment. Physical therapy and exercise may relieve some symptoms and improve function. Surgery is needed if compression is present and causing symptoms.

2.4.4 Imaging

The current accepted diagnostic criteria for DISH is based on plain radiography of the thoracic spine [36]. Large, flowing right-sided ossification over the thoracic spine is typical, extending over at least four vertebral bodies (Fig. 7). Preservation of intervertebral disk heights and absence of facet joint and sacroiliac joint involvement are also required for diagnosis.

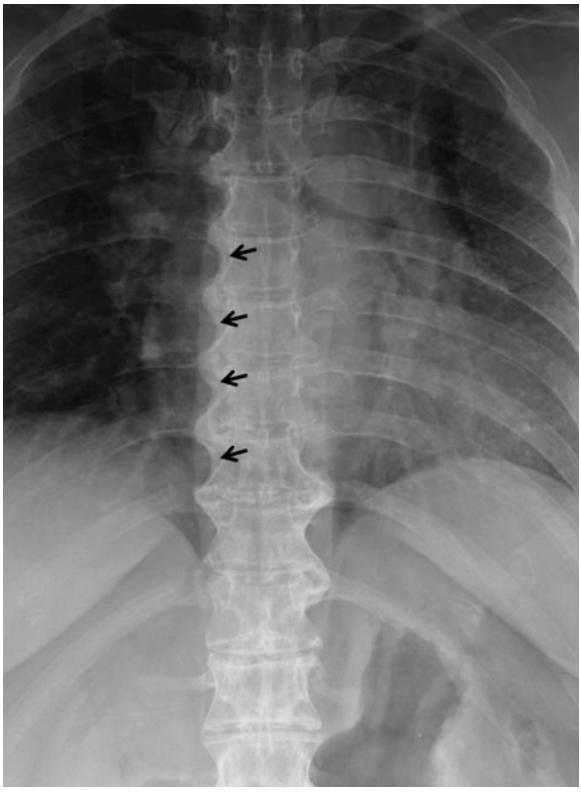


Fig. 7 Plain radiograph of the thoracic spine in a patient with diffuse idiopathic skeletal hyperostosis. Anteroposterior view of a thoracic spine. Radiolucent areas (*arrow*) indicate space between ossified longitudinal ligament and the vertebral bodies. The changes are predominantly located on the right side of the thoracic spine

3 Inflammatory Arthritis

3.1 Ankylosing Spondylitis

3.1.1 Definition and Occurrence

Ankylosing Spondylitis (AS, from the Greek *ankylos*, fused; *spondylos*, vertebrae; *-itis*, inflammation) is the prototypic disease of the seronegative spondyloarthritis family, a group of inflammatory spinal arthritis that includes AS, psoriatic arthritis, reactive arthritis, spondyloarthritis associated with inflammatory bowel disease, and undifferentiated spondyloarthritis. In contrast to rheumatoid arthritis, patients with seronegative spondyloarthritis usually do not produce autoantibodies, such as rheumatoid factor or anti-cyclic citrullinated protein (anti-CCP) antibody, and therefore are termed "seronegative." AS primarily involves the axial skeleton, including the spine and sacroiliac joints, with features of chronic inflammation and new bone formation. The sacroiliac joints are the connections between the lower end of the spine (sacrum) and the pelvis.

Genetic factors are important in the susceptibility to AS. Early study of AS in 1970s discovered an association with a gene called human leukocyte antigen B27 (HLA-B27) [37]. It is estimated that 85–90 % of patients with AS have HLA-B27, compared to under 10 % of the general population [38]. Among those who have HLA-B27, AS is more common among those with a close relative who also has AS than in those without any close relative with AS [39]. This indicates other genetic factors are involved in the pathogenesis of AS. Recent genome-wide association studies have advanced our understanding of the genetic basis of AS. More than 20 genes, e.g. ERAP1, IL-23R, KIF21B, etc., and a few intergenic regions, e.g. 2p15 on chromosome 2, are now identified to be associated with AS [40–42]. Substantial evidence suggests environmental factors trigger the onset of AS in people with certain genetic background, a theory well supported by the study of HLA-B27 transgenic rats. These rats develop arthritis and gut inflammation, resembling human HLA-B27 associated diseases. Interestingly, they are protected from the disease if raised in germ-free conditions [43].

Chronic inflammation of the entheses and new bone formation are two cardinal features of AS. Entheses are the sites where tendons or ligaments insert into bones. Immunohistologic staining of entheses from the sacroiliac joints [44] and the foot ligaments [45] of patients with AS showed inflammatory cell infiltration at these sites. In addition to enthesitis, inflammation occurs in bone (osteitis) and synovium (synovitis), and can cause pain and swelling.

New bone formation is a slow and insidious process, with new skeletal tissues formed in connection with, but extending outside the original bone [46]. Bony growths originating from the ligament insertions of the spine are called syndesmophytes, while those originating from the entheses in the extremities are called enthesophytes. Growth of syndesmophytes starts at the thoracolumbar junction, gradually involves other spinal areas, and may eventually lead to bridging of

vertebral bodies and complete fusion (ankylosis) of the spine, causing significant loss of mobility. The same process happens at the sacroiliac joints, causing ankylosis. Several molecular pathways have been proposed to be involved in this process, including bone morphogenic proteins (BMPs), Wnt protein, hedgehog protein and fibroblast growth factors (FGFs) [46]. The relationship between chronic inflammation and new bone formation remains unclear.

The prevalence of AS ranges from 0.1 to 1 % of the population, depending on the ethnic groups studied. Caucasians and Native Americans have the highest prevalence, while AS is rare in Africans. Men are 3 times more likely to have AS than women, and often have more severe disease. AS tends to run in families, with an estimated heritability of more than 90 % [47]. It usually begins during adolescence or early adulthood, and is life-long.

3.1.2 Clinical Manifestations

Inflammatory back pain is the most common symptom in patients with AS. Patients often describe pain in their lower back or the buttock, worse after rest, especially during the second half of the night. They often report waking up in significant pain and stiffness in the morning. Exercise and NSAIDs improve the back pain. The symptoms usually fluctuate, and are often associated with fatigue.

With progression of the disease and the growth of syndesmophytes, ankylosis of the spine becomes a more prominent feature in the disease presentation. In advanced stages of AS, patients may develop a stooped posture, have decreased movement of their spine, and significant loss of function. Fusion of the cervical spine leads to a forward flexion of the head, and patients may have difficulty raising their head to look straight ahead. Involvement of the thoracic spine and the chest wall may make it difficult for patients to expand their chest and take deep breaths, affecting the function of the lungs. In severe cases, patients may develop roundback, which prohibits them from sleeping on their back. Lumbar spine involvement may make it difficult for patient to bend forward and reach the floor. Complete fusion of the spine makes patients more susceptible to trauma and spine fractures. The extent of spine fusion varies greatly among patients, and progression to complete fusion is not inevitable.

Complaints in the limbs are common, mainly due to enthesitis and arthritis. Inflammation of the entheses causes intermittent pain and swelling at various tendon insertion sites, for example, at the heel, where the Achilles tendon attaches, or at the bottom of the foot. Hips, shoulders and collarbone joints are frequently involved with pain, stiffness, and sometimes swelling, indicating ongoing inflammation. Over time, inflammation may causes damage, with erosion of the bone and loss of cartilage and the joint space. Joint movements may become limited, for example, with flexion contracture of the hips.

Paradoxically, despite the propensity to add extra abnormal bone to the spine, patients with AS often develop osteoporosis, a condition of decreased bone density, with increased risk of fracture. Organs other than musculoskeletal system are affected in AS. Common manifestations include uveitis (inflammation of eyes),

aortitis (inflammation of the aorta, the largest artery in the body, originating from heart), and colitis (inflammation of the large bowel). Inflammation can also lead to secondary amyloidosis (a process of protein deposition in internal organs), usually associated with kidney dysfunction.

Laboratory tests have limited use in diagnosing AS. Patients may not have elevated blood markers of inflammation, even if they are actively having inflammatory symptoms. HLA-B27 is not required for diagnosis, and absence of HLA-B27 does not rule out the diagnosis of AS. However, in the appropriate clinical setting, HLA-B27 may suggest the diagnosis.

The modified New York criteria have been used for 30 years for the classification of AS. By these criteria, patients need have a characteristic clinical presentation and characteristic radiographic changes in the sacroiliac joints. Inflammatory back pain, limited motion of lumbar spine and limited chest wall expansion comprise the clinical components; at least one of these features is required for classifying a person as having AS by these criteria. Radiographic changes of the sacroiliac joints will be discussed in the Imaging section.

3.1.3 Treatment

The goal of the treatment is to control inflammation and pain, reduce new bone formation, and improve or maintain function. This is achieved through a combination of medications and non-pharmacologic modalities.

NSAIDs are the first line therapy for pain control and to decrease inflammation. If one NSAID is not effective or causes side effects, usually another NSAID from a different class can be tried. After an adequate trial of NSAIDs, if patients still have symptoms suggesting active inflammation, anti-tumor necrosis factor (anti-TNF) agents are usually considered as the next step. TNF is a pro-inflammatory cytokine. Anti-TNF agents are effective in decreasing pain, stiffness, fatigue, and joint swelling in AS, and in improving patient's function. With a tolerable side effect profile, anti-TNF agents are a mainstay treatment for AS. Discovery of the association between AS and the interleukin-23 pathway brings new treatment options. While these medications have shown effectiveness in controlling active inflammation, whether they can reduce new bone formation in AS remains unclear.

Physical therapy and exercise are essential in the treatment of AS. Patients usually experience a significant reduction of symptoms after exercise, and it helps them to maintain function. Stretching exercises, such as yoga, may increase spinal mobility, and deep-breathing may increase chest wall expansion and prevent the loss of lung function. Postural training is important and patients should avoid a flexed position for a prolonged period of time.

Patients with advanced AS may need corrective surgeries for complications associated with AS. In patients with severe hip involvement, total hip replacement often provides pain relief and functional improvement. In patients with complete fusion of the spine, the risk of spinal fracture is increased; surgical stabilization is needed if spinal fracture occurs.

3.1.4 Imaging

Plain radiography is important in the diagnosis of AS and in the exclusion of other diagnoses, particularly in patients with advanced disease. Usually an anteroposterior (AP) view of the pelvis is obtained for evaluation of the structural changes of the sacroiliac joints. Erosions, sclerosis and ankylosis of the sacroiliac joints are the common findings in AS. These changes are graded as 0–4, from normal to the most advanced disease (Fig. 8 and Table 2). Presence of bilateral grade 2 changes or unilateral grade 3 or 4 changes is required for classifying AS by the modified New York criteria. Structures other than the sacroiliac joints can be assessed by pelvis X-ray. Erosions and loss of the joint space of the hips, and calcification along the tendon insertions are seen in patients with AS.

However, pelvis radiographs have limitations. They have low sensitivity and specificity for bony changes early in the course of AS, and cannot show inflammation in the bone or joints. In patients with a short duration of symptoms, MRI of the pelvis and the lumbar spine is often used to detect early disease. Active

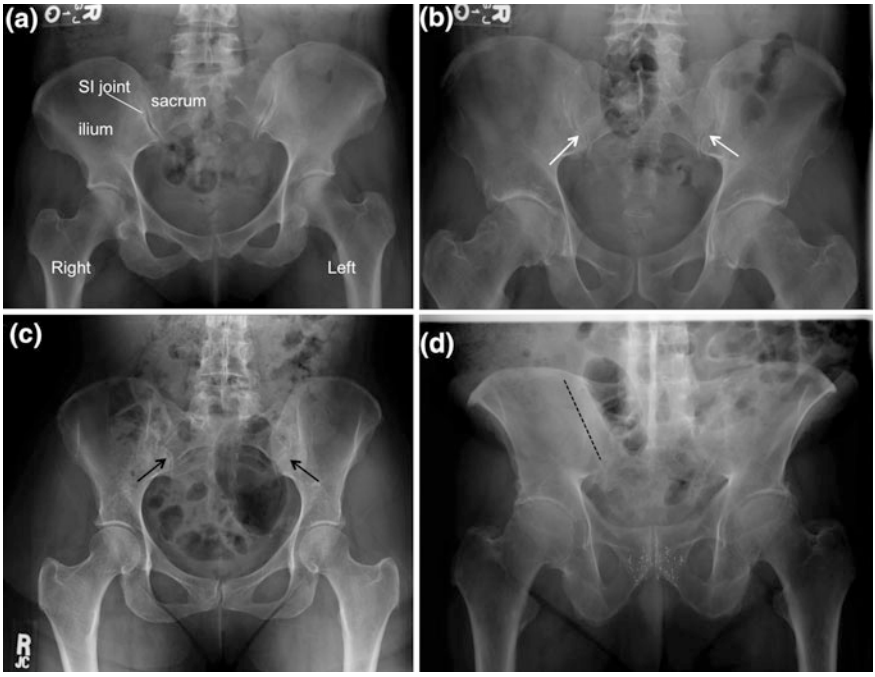


Fig. 8 Radiographic grading of sacroiliac joint involvement in ankylosing spondylitis. **a** *Right* sacroiliac (SI) joint grade 0 (normal); *left* sacroiliac joint grade 1 (suspicious for changes). **b** *Right* sacroiliac joint grade 2, *left* sacroiliac joint grade 2 (small localized narrowing, indicated by *white arrows*). **c** *Right* sacroiliac joint grade 3, *left* sacroiliac joint grade 3 [partially fused with residual joint space (*black arrows*)]. **d** *Right* sacroiliac joint grade 4, *left* sacroiliac joint grade 4 (complete fusion). *Dotted line* indicates the location of the fused *right* sacroiliac joint

Table 2 Radiographic grading of sacroiliac joint changes [48]

Grade 0	Normal
Grade 1	Suspicious changes
Grade 2	Minimum abnormality (small localized areas with erosion or sclerosis, without alteration in the joint width)
Grade 3	Unequivocal abnormality (moderate or advanced sacroiliitis with erosions, evidence of sclerosis, widening, narrowing, or partial ankylosis)
Grade 4	Severe abnormality (total ankylosis)

inflammatory lesions are best visualized in T2 weighted fat-saturated sequence or short tau inversion recovery (STIR) sequence as hyperintense signals. When the hyperintense signal appears in the sacrum or iliac bone, it represents bone marrow edema or osteitis, which is thought to represent inflammation of the bone. It may appear in other locations, such as tendon insertions or synovium, indicating enthesitis or synovitis. Active sacroiliitis is defined by the presence of bone marrow edema/osteitis, and is very suggestive of AS or a condition in the spondyloarthritis family. Erosions, sclerosis, and fatty change can also be detected by MRI. They are chronic sequelae of active inflammation, but their utility in diagnosing AS is not clear at this time.

Spine radiographs are useful to exclude other conditions, and are also useful to assess disease progression in patients with an established diagnosis of AS. Anteroposterior and lateral views of the cervical and lumbosacral spine are usually obtained for evaluation. The vertebral bodies are best visualized on the lateral views. Erosions, sclerosis, or squaring of the vertebral bodies are early findings; with disease progression, syndesmophytes may develop, bridging syndesmophytes form between the neighboring vertebras, and may eventually lead to a completely fused spine, or a 'bamboo spine' (Fig. 9). The modified Stoke Ankylosing Spondylitis Spinal Score (mSASSS) is a scoring system used to assess the extent of this process by examining changes at the anterior corners of the cervical and lumbar vertebrae on lateral radiographs. Structural changes of the facet joints can be visualized on AP view, with sclerosis and loss of the joint space being the most common findings.

The treatment goal for AS is to reduce inflammation and to reduce new bone formation. To assess treatment response objectively, imaging modalities to visualize inflammation and that are sensitive to bone growth are ideal. Plain radiography has several disadvantages for these purposes. First, as two-dimensional imaging modality, it has poor visualization of syndesmophytes due to overlying shadows. Second, scoring systems based on plain radiograph are semi-quantitative, and therefore tend to be insensitive to change. Third, as mentioned earlier, it does not detect inflammatory changes. Three-dimensional imaging modalities may potentially address these issues.

MRI of the spine is considered the "gold standard" for visualizing inflammation. Similar to the changes seen in the sacroiliac joints, hyperintense signal on T2 fat saturated sequence or STIR sequence depicts inflammation in the spine (Fig. 10b). Structural changes, such as erosions, can be detected as well (Fig. 10a). Studies

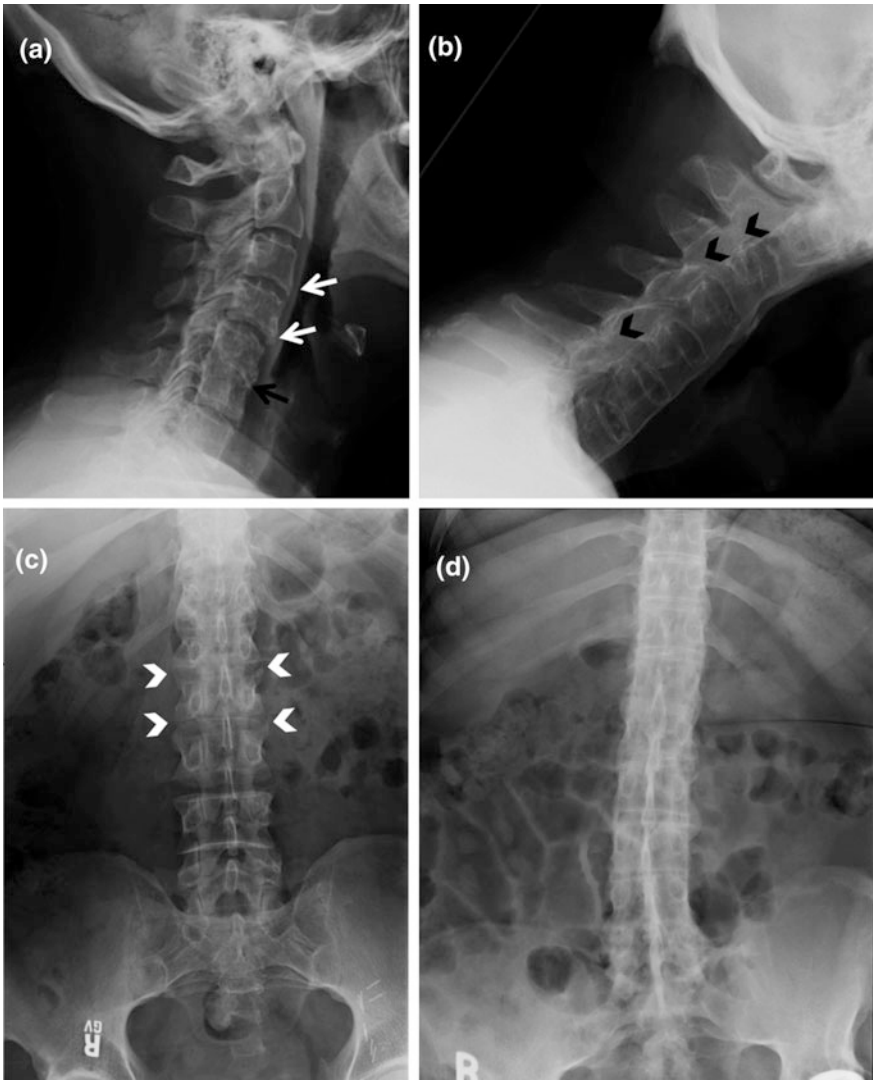


Fig. 9 Plain radiographs of the spine in patients with ankylosing spondylitis. **a** Lateral view of cervical spine. Syndesmophytes (*white arrows*) that projected vertically from the edge of the vertebral body. There is loss of intervertebral disk space (*black arrow*). **b** Lateral view of a completely fused cervical spine. Notice the 'bamboo' shape of vertebral column and complete loss of facet joints (*black arrowheads*). **c** Anteroposterior view of lumbar spine showing syndesmophytes (*white arrowheads*), some of which are almost bridging. **d** Anteroposterior view of a completely fused lumbar spine (*bamboo spine*). The linear vertical density in the center of the spine is formed by calcification of the ligaments between adjacent spinous processes

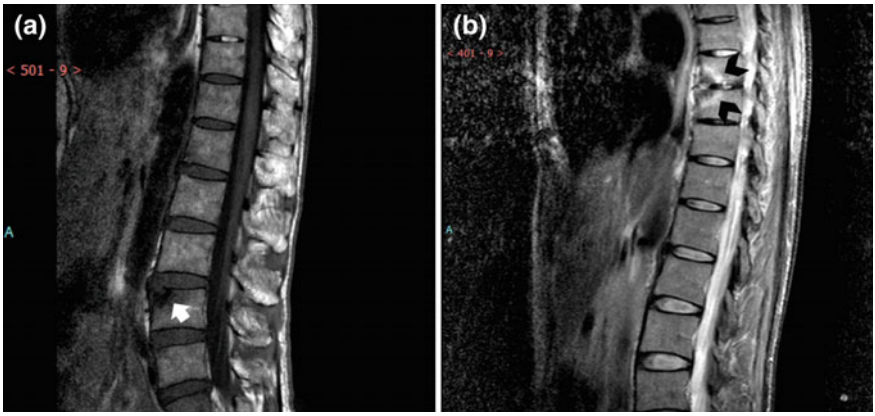


Fig. 10 Spine abnormalities on magnetic resonance imaging in a patient with ankylosing spondylitis. Sagittal view of lumbar spine. **a** Erosion of the anterior portion of a lumbar vertebral body (*arrow*) is shown as hypointense signal in T1 weighted image. **b** Erosions and osteitis (*arrowheads*) on adjacent vertebral bodies is shown as hyperintense signal on a STIR weighted image

have shown that anti-TNF therapy can decrease inflammation as detected by spinal MRI. Therefore, MRI scoring systems have been developed to assess the inflammatory signals and their changes with treatment [49]. It has also been reported that the combination of bone marrow edema and fatty deposits on MRI may predict the development of new syndesmophytes [50]. However, its utility for evaluating structural damage is still under investigation.

CT provides an accurate and sensitive assessment of changes of bony structures, so it is ideal to detect new bone formation. CT of the lower thoracic spine has been used to quantitate the volume of syndesmophytes and changes in their size over 2 years, and demonstrated good validity. Compared to MRI or plain radiograph, CT was more sensitive to change [51]. At this time, its application is limited to research due to its radiation exposure.

Spondyloarthritis associated with inflammatory bowel disease, also called enteropathic arthritis, develops in 20 % of patients who have Crohn's disease or ulcerative colitis, conditions that involve bowel inflammation. Many have spine involvement with the same Pathogenetic process of AS.

3.2 Psoriatic Arthritis

3.2.1 Definition and Occurrence

Psoriatic arthritis is a chronic inflammatory arthritis in patients with psoriasis that involves the spine and peripheral joints. In 60–80 % of cases, the skin rash of psoriasis precedes the development of arthritis; in 15 % patients, arthritis is the presenting symptom; occasionally, psoriasis and arthritis develop concurrently [52].

Prominent features in psoriatic arthritis are synovitis, enthesitis, erosions and new bone formation. Biopsies of the joint tissue revealed increased white blood cells and prominent blood vessels, and 47 % patients develop erosions in bones within 2 years of diagnosis [53]. Enthesitis and bone marrow edema on MRI demonstrate its similarity with other seronegative spondyloarthritis.

Psoriatic arthritis develops in 4–30 % patients with psoriasis [52, 54], with a prevalence of 0.1–0.2 % of the general population [54]. It affects men and women equally, most commonly developing in people aged 30–55 years.

3.2.2 Clinical Manifestations

Five clinical patterns have been described in psoriatic arthritis: asymmetric oligoarthritis, symmetric polyarthritis, distal interphalangeal (DIP) arthropathy, arthritis mutilans, and spondylitis with or without sacroiliitis [55]. Patients may have features of more than one pattern, and their presentations may change during the course of the disease.

Asymmetric oligoarthritis is the most common pattern, and involves inflammation in fewer than 5 joints. Large joints, such as knees or hips, are affected most often. The symmetric polyarticular pattern resembles rheumatoid arthritis, and mainly involves small joints such as fingers, hands and wrists. DIP arthropathy, affecting the finger joints closest to the nails, is a characteristic of psoriatic arthritis, being rarely seen in rheumatoid arthritis. Patients with peripheral joint involvement often complain pain and swelling of these joints, associated with morning stiffness. Synovitis, or inflammation of lining of the joint, is the underlying pathology. Arthritis mutilans is a rare destructive condition caused by absorption of the finger bones, and is also characteristic of psoriatic arthritis.

Spine involvement is less common than limb arthritis in psoriatic arthritis. Back pain, buttock pain, stiffness, and fatigue are the main complaints in these patients. Involvement of the sacroiliac joints is not always present in psoriatic arthritis, or may only affect the right or left side, as opposed to both sacroiliac joints in AS. Another commonly affected site is the cervical spine. As seen in rheumatoid arthritis, inflammation and erosions can cause atlantoaxial (C1–C2) subluxation, which can lead to cervical myelopathy.

Enthesitis is often present in patients with psoriatic arthritis. Patients may report pain and sometimes swelling at the heel or the bottom of the foot. A few features help to distinguish psoriatic arthritis from AS. Psoriatic skin and nail changes are seen in most patients, providing the major diagnostic clue. Sausage-shaped swelling of a finger or toe is a characteristic manifestation of psoriatic arthritis [56]. Ultrasound and MRI studies show that inflammation of the tendon sheath (tenosynovitis) is the cause of this type of finger or toe swelling.

Laboratory tests are non-diagnostic. Elevated blood markers of inflammation may be present [57]. As one of the seronegative spondyloarthritis, rheumatoid factor and anti-CCP antibody are often absent. HLA-B27 is present in some patients, particularly those with spine involvement.

3.2.3 Treatment

Treatment of psoriatic arthritis is varied because of the diversity of clinical presentations. For mild arthritis, NSAIDs are the first line treatment to control symptoms. For patients with peripheral arthritis affecting more than 3 joints, a disease modifying anti-rheumatic drug (DMARD) is often considered, such as leflunomide, sulfasalazine, cyclosporine, or methotrexate [58, 59]. Methotrexate is the first choice of many rheumatologists.

Anti-TNF agents have shown effectiveness in controlling acute inflammation and preventing bony erosions in psoriatic arthritis. Alefacept, a fusion protein targeting lymphocyte function antigen 3 (LFA3), and ustekinumab, an interleukin 12/23 inhibitor, are also effective.

Conventional DMARDs and newer biologics may have some effect on other manifestations, such as enthesitis and spondylitis. However, responses of these manifestations have not been well studied.

3.2.4 Imaging

Plain radiography remains the standard for diagnosing psoriatic arthritis and monitoring its progression. A characteristic finding of psoriatic arthritis is the co-existence of erosions and new bone formation, most prominent at the finger joints. Absorption and lysis of the finger bones may lead to typical 'pencil-in-cup' appearance on radiographs. Fusion of hand bones, fluffiness of the bony cortex, and calcification of entheses are evidence of new bone formation. The sacroiliac joints are occasionally involved in psoriatic arthritis, and erosion, sclerosis and ankylosis of these joints are common findings. Dynamic imaging of the cervical spine, with flexion and extension of the head, may reveal instability of the cervical spine. Spinal radiographs may depict syndesmophytes, which tend to originate from the mid-part of the vertebral body rather than the vertebral corner, and spine involvement is often discontinuous.

Radiographic progression of psoriatic arthritis is slow. Radiographic scoring systems have been adapted from rheumatoid arthritis but modified to include the distal interphalangeal joints, and are used to assess joint erosions and disease progression in clinical trials.

MRI has been used to assess enthesitis in psoriatic arthritis, and led to new understanding of its pathogenesis. Inflammation of the entheses and associated bone marrow edema are the most common MRI finding. It has been proposed to use MRI of the spine and sacroiliac joints as a more sensitive way to assess spinal involvement in psoriatic arthritis; however, at present, it is still limited to research settings.

Musculoskeletal ultrasound is a sensitive way to detect inflammation, and has been used to assess the response of tenosynovitis, synovitis, and enthesitis to treatment.

3.3 Reactive Arthritis

Reactive arthritis is a form of seronegative inflammatory arthritis that develops after an infection somewhere in the body outside of the joints. Some gastrointestinal and urinary infections are considered causal, including those due to the bacteria *Chlamydia trachomatis*, *Yersinia*, *Salmonella*, *Shigella*, *Campylobacter* [60], *Escherichia coli*, *Clostridium difficile* and *Chlamydia pneumonia* [61].

Reactive arthritis is uncommon, and affects men and women equally. Symptoms of arthritis develop several days or weeks after the initial infection [60]. A preceding infection is not always identified, even in patients with a typical presentation. Often, pain and swelling develops in a few joints, particularly in the knees or ankles. Enthesitis is not uncommon. Patients may have signs of eye inflammation, urinary symptoms, and skin rashes. In most patients, the arthritis subsides after 6 months, however, in a small proportion, it may become chronic.

Spine involvement usually manifests as inflammatory back pain. Twenty-five percent of patients develop radiographic changes of the sacroiliac joints, usually affecting only one side. Extensive spine fusion is very uncommon. Patients with spine involvement often have HLA-B27 [62].

Treatment of reactive arthritis is mainly symptomatic. NSAIDs are used to control acute inflammation. Limited evidence supports the use of antibiotics when Chlamydia is the cause [63]. In chronic reactive arthritis, methotrexate, sulfasalazine, and biologics have been used with various responses.

In acute reactive arthritis, plain radiographs of the affected joints are mainly used to exclude other diagnoses. In patients with spinal involvement, radiographic changes of the sacroiliac joint can be seen after some duration of symptoms, with more severe changes on one side compared to the other. Syndesmophytes may develop that are often bulky, asymmetric and extend laterally [64] (Fig. 11). MRI of the spine and entheses is potentially useful for the assessment of inflammation and monitoring responses to treatment, but has not been evaluated extensively.

3.4 Undifferentiated Spondyloarthritis

Undifferentiated spondyloarthritis refers to spondyloarthritis that does not fulfill criteria for AS, psoriatic arthritis, reactive arthritis, or enteropathic spondyloarthritis [65]. Most patients are young men with inflammatory low back pain who may have HLA-B27. Some may have arthritis in peripheral joints or enthesitis. The major

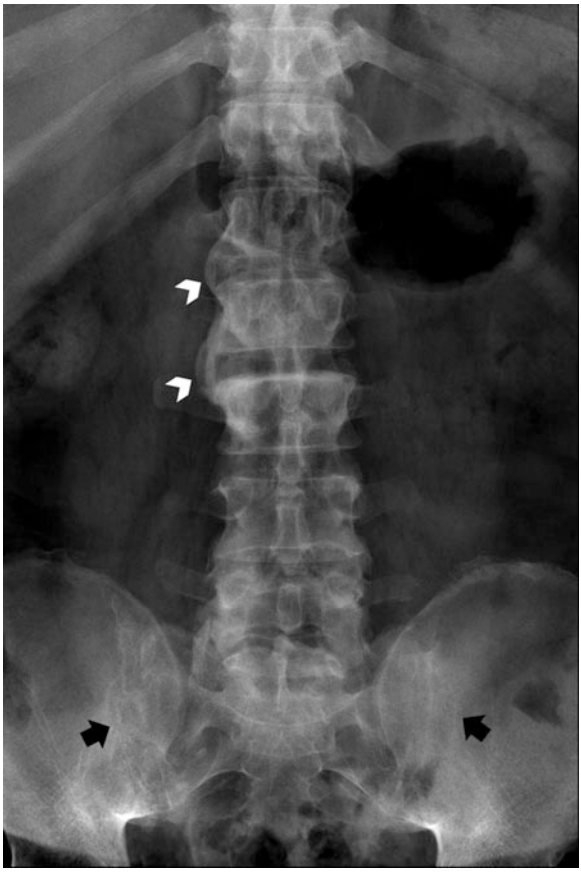


Fig. 11 Plain radiograph of the lumbar spine in a patient with reactive arthritis. Notice the asymmetric distribution of the syndesmophytes (*arrowheads*), affected primarily the right side. This patient also has grade 3 changes in both sacroiliac joints (*arrows*)

difference between undifferentiated spondyloarthritis and AS is the presence or degree of the sacroiliac joint changes on plain radiography. In patients with undifferentiated spondyloarthritis, the sacroiliac joint changes are absent or very mild, and do not meet the radiographic requirement for AS. MRI of the sacroiliac joints in the undifferentiated patients may reveal active inflammatory lesions similar to those present in AS. Patients with undifferentiated spondyloarthritis may evolve into a more specific type of spondyloarthritis, with AS being the most common. Alternatively, these patients may persist without differentiating to a more defined disease, or it may resolve completely after several years [66].

3.5 Rheumatoid Arthritis

3.5.1 Definition and Occurrence

Rheumatoid arthritis (RA) is the most common type of inflammatory arthritis in adults, affecting 0.5–1 % of the population. RA typically begins in middle-age or old-age, and women are three times more likely to have RA than men. The cause of RA is unknown. There are a number of gene variations that are associated with an increased risk of RA, but RA is not strictly hereditary. Current theory holds that RA develops as a consequence of exposure to an environmental trigger in a genetically-susceptible person [67]. Whether the trigger is the same for all patients is unknown. Smoking has been identified as a risk factor for RA, and smokers who have particular variants of the HLA-DR gene are at greatly increased risk. For most patients, RA is a life-long disease. While there is currently no cure for RA, medications can improve and control symptoms, and remission is possible with treatment. A small proportion of patients may have their RA spontaneously go into remission.

RA is an autoimmune disease. Autoimmune diseases are a category of diseases characterized by immune reactions against the body's own tissues. While the primary roles of the immune system are to provide protection from infections and to seek and destroy cells that may progress to tumors, these immune responses become subverted in autoimmune diseases. In autoimmune diseases, the immune system senses certain normal proteins or cells as foreign or abnormal. In RA, the immune system generates inflammatory cells that target the lining tissue of the joint (synovium), the cartilage that caps the end of bones and forms the gliding surface of joints, and components of the immune system itself [68]. The immune system also begins to make antibodies against normal proteins, which can inactivate them. In RA, two of these so-called autoantibodies are commonly made. Rheumatoid factor is an antibody to immunoglobulins, which are proteins that provide immune protection against viruses and bacteria. Antibodies to citrullinated proteins bind specific proteins found in the connective tissue between cells. Both of these antibodies can be measured in clinical laboratories, and are used to aid in the diagnosis of RA. About 80 % of patients with RA have either rheumatoid factor or antibodies to citrullinated proteins detectable in their blood.

Inflammation develops as a consequence of these autoimmune reactions. In the joints, this inflammation causes joint swelling due to fluid accumulation in the joint space, infiltration of the synovium by white blood cells and expansion of blood vessels, and over time, proliferation of the synovial cells. Persistent inflammation can, over weeks to months, lead to loss of mineralization of the surrounding bone, wearing away of the joint cartilage, and eventually erosion of the bone surfaces at the margins of the joints. Persistent joint swelling can also stretch and weaken surrounding ligaments and tendons, resulting in shifting of the joints out of normal alignment. RA is therefore known as a deforming arthritis. This shifting places the joints at mechanical disadvantage, and which along with swelling, can cause weakness.

3.5.2 Clinical Manifestations

RA affects multiple joints simultaneously, with pain, stiffness, and swelling [69]. These symptoms in turn cause problems in using the joints to accomplish movements and tasks, such as walking or getting dressed. RA primarily affects peripheral joints, and less commonly the spine. Small joints of the fingers and hands, and wrists are affected in almost all patients. Knees, ankles, and small joints of the feet and toes are also commonly affected. While other joints are less commonly involved in RA, any synovial joint may be affected. Without treatment, the joint pain and swelling tends to persist and can last weeks or months. Even with treatment, symptoms may at times wax and wane, with "flares" of worsening joint inflammation occurring episodically. A feeling of stiffness, or restricted ease of movement, in and around the joints is common, particularly in the morning or after periods of inactivity. Fatigue is also common during periods of active inflammation, and joint pain may interfere with sleep. Patients often experience depression as a consequence of chronic pain and concern about their future health.

Chronic joint inflammation that leads to cartilage, bone, and ligament damage can result in joint deformities. Common deformities include fixed flexion of the fingers, sideways drifting of the fingers at the knuckles, and inward deviation of the knees and ankles. Muscle weakness may result from both these deformities and from disuse of painful joints. Loss of cartilage can also lead to limited range of motion of the joints, which in severe cases can fuse and become immobile.

The cervical spine is involved in up to 80 % of patients with RA, although symptoms related to the cervical spine may be present in less than one-half of patients [70]. Cervical spine problems are more common later in the course of RA than at the onset. The main symptoms are neck pain, headache at the back of the head, and less commonly, numbness of the arms, hands, or legs. Rarely, instability of the cervical spine as a result of inflammation can cause the vertebrae to impinge on nerve roots or even the spinal cord, causing radiculopathy or myelopathy. Depending on the location of the impingement, serious neurological complications may occur. If the spinal cord is impinged, paralysis may result. If the brainstem is impinged, sudden death may occur. These problems may be provoked by movements that flex or extend the neck, and so raise particular concerns about whiplash injuries in automobile accidents. Inadvertent injuries may also occur during the placement of breathing tubes prior to general anesthesia, which requires the head and neck to be extended. Impingement of major blood vessels at the base of the skull may also occur and cause dizziness, weakness, and vision changes. The Ranawat classification system is commonly used to grade the degree of neurological damage in patients with RA-related cervical spine disease. Class I indicates no neurological deficits. Class II indicates subjective weakness and numbness. Class IIIA represents objective weakness and signs of spinal cord compression but with preserved ability to walk, while Class IIIB represents weakness and signs of cord compression with inability to walk.

Three types of cervical spine involvement are commonly recognized, which are distinguished by the specific areas of the cervical spine that are involved: atlanto-axial

subluxation, atlanto-axial impaction, and subaxial subluxation. Atlanto-axial subluxation is the separation of joint between the atlantis (the common name of C1, the first cervical vertebra) and the axis (the common name of C2, the second cervical vertebra). Inflammation of this joint leads to loosening of the surrounding ligaments, which can permit a dynamic separation of this joint with flexion of the head. With extensive subluxation, the superior part of C2 (known as the dens) can compress the spinal cord when the head is flexed. Atlanto-axial impaction results when bone and cartilage loss between the base of the skull and C1, and between C1 and C2, leads to the superior migration of C2 relative to the skull. In severe cases, part of C2 can penetrate the spinal cord opening at the base of the skull and compress the brainstem. Because the brainstem controls vital functions such as respiration, compression may result in death. Subaxial subluxation is the malalignment of vertebrae below C2, due to chronic erosive joint inflammation and ligament instability. Atlanto-axial subluxation is the most common cervical spine abnormality, occurring in up to 50 %. Atlanto-axial impaction occurs in up to 40 %, while subaxial subluxation occurs in 10–20 %. The thoracic spine and lumbar spine are typically not affected by RA.

Although RA primarily affects the joints, other parts of the body may be affected by inflammation due to RA, including the lungs or lung linings, the heart lining, the outer surface of the eye, and blood-forming elements in the bone marrow. Vasculitis, or inflammation of the blood vessels, may also occur.

The diagnosis of RA is based on a compatible clinical presentation, inflammation in many small joints on both the right and left sides, and the presence of autoantibodies (either rheumatoid factor or antibodies to citrullinated proteins). Blood tests indicating systemic inflammation, such as the C-reactive protein level, are also often elevated. Radiographs that show bone erosions in typical locations can also be helpful, but because these lesions take time to develop, they are often not present at the start of symptoms.

3.5.3 Treatment

The goal of RA treatment is prompt and complete control of joint inflammation, which will lessen symptoms, improve quality of life, and decrease the likelihood of chronic joint damage and associated disability [71]. Medications therefore occupy the central focus in RA treatment. While analgesics and NSAIDs such as naproxen and ibuprofen can help lessen joint pain, they provide only temporary symptom benefit. Corticosteroids can also be beneficial in controlling joint inflammation, but side effects preclude their chronic use. Appropriate treatment requires the long-term use of one or more “disease-modifying” medications, which over time provide for more sustained control of inflammation and the potential to decrease the development of joint damage [72]. Methotrexate, taken weekly in low doses, is the most commonly used disease-modifying medication, based on evidence of sustained efficacy and generally good tolerability. Hydroxychloroquine, sulfasalazine, and leflunomide are other disease-modifying medications that can be used alone or in

conjunction with methotrexate. Biologic medications, which are antibodies developed to block key mediators of inflammation such as tumor necrosis factor-alpha or interleukin-6, are also effective in controlling inflammation and slowing joint damage. Biologics are often added when conventional disease-modifying medications have proven to be insufficient at controlling joint pain and swelling. For most patients, treatment is needed for years or decades, although slow tapering of medications is often possible as the inflammation comes under control.

Physical therapy and occupational therapy can help improve joint function and range of motion. Joint replacement surgery or joint fusion surgery is indicated when dysfunction or persistent pain of damaged joints limits the patient's ability to do daily activities. These surgeries are very effective in relieving pain and restoring functional ability. Treatment of cervical spine involvement includes traction to help relieve pressure on the impinged nerves or spinal cord, surgery to decompress the area by removing excess synovial tissue, and surgical fusion of the vertebrae to stabilize regions of subluxation. Cervical spine surgery is often effective in providing at least partial pain relief and preventing worsening of the neurological problems. However, existing neurological damage may not reverse with surgery. Recovery from quadriparesis is uncommon and the survival of these patients is low [73].

3.5.4 Imaging

Plain radiographs of peripheral joints are very useful in the diagnosis of RA, as well as helping to distinguish other types of arthritis that may have clinical features that mimic RA [74]. In early RA, radiographs may be normal or show only prominent shadows of the joint linings or excess joint fluid. Osteopenia next to the joints may also be visible. Tell-tale bone erosions at the margins of the joints are the most specific radiographic sign of RA, and occur in up to 60 % of patients. Radiographs can also demonstrate joint space narrowing due to cartilage loss, bony fusion, and bone malalignment or subluxations. In the cervical spine, radiographs can adequately show each of the three main types of involvement, although films taken with both neck flexion and extension may be needed to reveal dynamic atlanto-axial subluxation (Fig. 12). Radiographs are useful in planning surgical approaches and evaluating the results of surgical corrections. Development of new bone erosions and progressive joint space narrowing on radiographs is used in clinical trials to test the efficacy of medications and in clinical practice to monitor patient's responses to medications.

Diagnostic ultrasound has been increasingly used to detect thickening of the joint linings and other signs of joint inflammation, such as enhanced Power Doppler signals, which may help in diagnosis [75]. Improvement or resolution of these features can also be used to assess remission and response to treatment. Erosions of bone are also visible on ultrasound, but the time needed for examination may limit its use for assessing progression of erosions. Ultrasound can also be used to guide the injection of medications into a joint.

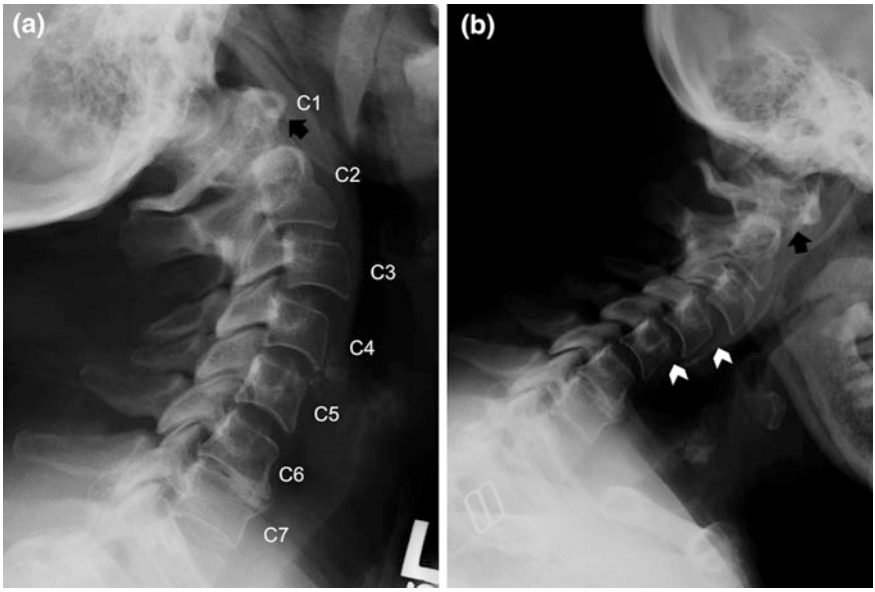


Fig. 12 Plain radiographs of the cervical spine showing atlanto-axial subluxation and subaxial subluxation in a patient with rheumatoid arthritis. The lateral views of the cervical spine were taken with neck extension (**a**) and flexion (**b**). In the flexion position, a separation of anterior arch of C1 from the dens of C2 (*arrow*) is revealed. This is not observed in the extension position (*arrow*), indicating the presence of dynamic atlantoaxial subluxation. Subaxial subluxation (*arrowheads*) is also present at C3–C4 and C4–C5

MRI can also demonstrate synovial thickening and excess joint fluid related to RA, as well as bone erosions. In addition, bone marrow edema on MRI may indicate inflammation not otherwise appreciated. However because of its expense and the fact that most information can be obtained by other modalities, MRI is not often used in clinical practice for imaging the peripheral joints in RA. However, MRI is valuable in imaging the cervical spine, particularly in detailing areas of spinal cord or nerve root compression. Computed tomography is not often used, because the necessary structural information can in most instances be obtained by radiography or MRI.

3.6 Juvenile Idiopathic Arthritis

Juvenile idiopathic arthritis is the most common type of arthritis in children, but is rare, affecting 10,000–60,000 children in the United States. Among several subtypes, three are most common: pauci-articular, polyarticular, and systemic-onset subtypes. The pauci-articular subtype, which typically affects girls under the age of 5, presents with inflammation in 4 or fewer joints, most often the knees, ankles, or

elbows. Over time, the arthritis tends to resolve, although in some patients, it may persist and affect additional joints. The polyarticular subtype affects older girls, involves 5 or more joints at onset, and mimics adult RA. Patients with the systemic-onset subtype have not only arthritis but fever, skin rashes, blood cell abnormalities, and liver inflammation. Treatment of juvenile idiopathic arthritis generally follows that of adult RA.

Cervical spine inflammation occurs in up to 70 % of patients with juvenile idiopathic arthritis, and affects patients with each subtype. Neck pain and limited range of motion are the most common associated symptoms. While atlanto-axial subluxation and subaxial subluxation occur, the most common cervical spine abnormality in juvenile idiopathic arthritis is fusion of the facet joints [76]. This fusion often extends for the entire length of the cervical spine, with consequent inability to freely move the head and neck.

Radiographs are useful to detect cervical spine involvement, and MRI may be helpful to identify areas of potential neurological impingement.

References

1. Lestini WF, Wiesel SW (1989) The pathogenesis of cervical spondylosis. Clin Orthop Relat Res 239:69–93
2. Aoki J et al (1987) End plate of the discovertebral joint: degenerative change in the elderly adult. Radiology 164(2):411–414
3. Moskowitz RW, Holderbaum D (2001) Clinical and laboratory findings in osteoarthritis. In: Koopman WJ (ed) Arthritis and allied conditions. Williams & Wilkins, Baltimore, pp 2216–2245
4. Li J et al (2011) Prevalence of facet joint degeneration in association with intervertebral joint degeneration in a sample of organ donors. J Orthop Res 29(8):1267–1274
5. Gellhorn AC, Katz JN, Suri P (2013) Osteoarthritis of the spine: the facet joints. Nat Rev Rheumatol 9(4):216–224
6. Muraki S et al (2009) Prevalence of radiographic lumbar spondylosis and its association with low back pain in elderly subjects of population-based cohorts: the ROAD study. Ann Rheum Dis 68(9):1401–1406
7. Saal JS, Saal JA, Yurth EF (1996) Nonoperative management of herniated cervical intervertebral disc with radiculopathy. Spine (Phila Pa 1976) 21(16):1877–1883
8. Kellgren JH, Lawrence JS (1957) Radiological assessment of osteo-arthrosis. Ann Rheum Dis 16(4):494–502
9. Kellgren JH, Lawrence J (eds) (1963) The epidemiology of chronic rheumatism: atlas of standard radiographs of arthritis. Blackwell Scientific, Oxford
10. Kirkaldy-Willis WH (1988) The pathology and pathogenesis of low back pain. Managing low back pain. Churchill Livingston, New York, p 49
11. Okada E et al (2009) Aging of the cervical spine in healthy volunteers: a 10-year longitudinal magnetic resonance imaging study. Spine (Phila Pa 1976) 34(7):706–712
12. Radhakrishnan K et al (1994) Epidemiology of cervical radiculopathy. A population-based study from Rochester, Minnesota, 1976 through 1990. Brain 117(Pt 2):325–335
13. Teresi LM et al (1987) Asymptomatic degenerative disk disease and spondylosis of the cervical spine: MR imaging. Radiology 164(1):83–88

14. Teraguchi M et al (2013) Prevalence and distribution of intervertebral disc degeneration over the entire spine in a population-based cohort: the Wakayama Spine Study. Osteoarthritis Cartilage
15. Schellhas KP et al (1996) Cervical discogenic pain. Prospective correlation of magnetic resonance imaging and discography in asymptomatic subjects and pain sufferers. Spine (Phila Pa 1976) 21(3):300–311 (discussion 311–312)
16. Boden SD et al (1990) Abnormal magnetic-resonance scans of the cervical spine in asymptomatic subjects. A prospective investigation. J Bone Joint Surg Am 72(8):1178–1184
17. Yoss RE et al (1957) Significance of symptoms and signs in localization of involved root in cervical disk protrusion. Neurology 7(10):673–683
18. Majlesi J et al (2008) The sensitivity and specificity of the Slump and the Straight Leg Raising tests in patients with lumbar disc herniation. J Clin Rheumatol 14(2):87–91
19. Bogduk N, Aprill C (1993) On the nature of neck pain, discography and cervical zygapophysial joint blocks. Pain 54(2):213–217
20. Panagiotacopulos ND et al (1987) Water content in human intervertebral discs. Part I. Measurement by magnetic resonance imaging. Spine (Phila Pa 1976) 12(9):912–917
21. Smith BM et al (1998) Interobserver reliability of detecting lumbar intervertebral disc high-intensity zone on magnetic resonance imaging and association of high-intensity zone with pain and anular disruption. Spine (Phila Pa 1976) 23(19): 2074–2080
22. Aprill C, Bogduk N (1992) High-intensity zone: a diagnostic sign of painful lumbar disc on magnetic resonance imaging. Br J Radiol 65(773):361–369
23. Atlas SJ, Delitto A (2006) Spinal stenosis: surgical versus nonsurgical treatment. Clin Orthop Relat Res 443:198–207
24. Andreisek G, Hodler J, Steurer J (2011) Uncertainties in the diagnosis of lumbar spinal stenosis. Radiology 261(3):681–684
25. Kalichman L et al (2010) Computed tomography-evaluated features of spinal degeneration: prevalence, intercorrelation, and association with self-reported low back pain. Spine J 10 (3):200–208
26. Hall S et al (1985) Lumbar spinal stenosis. Clinical features, diagnostic procedures, and results of surgical treatment in 68 patients. Ann Intern Med 103(2):271–275
27. Bednarik J et al (2011) Are subjects with spondylotic cervical cord encroachment at increased risk of cervical spinal cord injury after minor trauma? J Neurol Neurosurg Psychiatry 82 (7):779–781
28. Bracken MB et al (1990) A randomized, controlled trial of methylprednisolone or naloxone in the treatment of acute spinal-cord injury. Results of the second national acute spinal cord injury study. N Engl J Med 322(20):1405–1411
29. Kadanka Z et al (2005) Predictive factors for mild forms of spondylotic cervical myelopathy treated conservatively or surgically. Eur J Neurol 12(1):16–24
30. Karpova A et al (2013) Reliability of quantitative magnetic resonance imaging methods in the assessment of spinal canal stenosis and cord compression in cervical myelopathy. Spine (Phila Pa 1976) 38(3):245–252
31. Kerkovsky M et al (2012) Magnetic resonance diffusion tensor imaging in patients with cervical spondylotic spinal cord compression: correlations between clinical and electrophysiological findings. Spine (Phila Pa 1976) 37(1):48–56
32. Uda T et al (2013) Assessment of cervical spondylotic myelopathy using diffusion tensor magnetic resonance imaging parameter at 3.0 tesla. Spine (Phila Pa 1976) 38(5):407–414
33. Resnick D, Niwayama G (1976) Radiographic and pathologic features of spinal involvement in diffuse idiopathic skeletal hyperostosis (DISH). Radiology 119(3):559–568
34. Utsinger PD (1985) Diffuse idiopathic skeletal hyperostosis. Clin Rheum Dis 11(2):325–351
35. Weinfeld RM et al (1997) The prevalence of diffuse idiopathic skeletal hyperostosis (DISH) in two large American Midwest metropolitan hospital populations. Skeletal Radiol 26 (4):222–225
36. Resnick D, Niwayama G (1998) Diagnosis of bone and joint disorder, 2nd edn

37. Schlosstein L et al (1973) High association of an HL-A antigen, W27, with ankylosing spondylitis. N Engl J Med 288(14):704–706
38. Khan MA (1997) A worldwide overview: the epidemiology of HLA-B27 and associated spondyloarthritides. In: Calin A, Taurog J (eds) Spondyloarthritides. Oxford University Press, New York, pp 17–26
39. van der Linden SM et al (1984) The risk of developing ankylosing spondylitis in HLA-B27 positive individuals. A comparison of relatives of spondylitis patients with the general population. Arthritis Rheum 27(3):241–249
40. Spondyloarthritis Australo-Anglo-American (2010) C., et al., Genome-wide association study of ankylosing spondylitis identifies non-MHC susceptibility loci. Nat Genet 42(2):123–127
41. Evans DM et al (2011) Interaction between ERAP1 and HLA-B27 in ankylosing spondylitis implicates peptide handling in the mechanism for HLA-B27 in disease susceptibility. Nat Genet 43(8):761–767
42. International Genetics of Ankylosing Spondylitis Consortium (2013) Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. Nat Genet 45(7):730–738
43. Taurog JD et al (1994) The germfree state prevents development of gut and joint inflammatory disease in HLA-B27 transgenic rats. J Exp Med 180(6):2359–2364
44. Braun J et al (1995) Use of immunohistologic and in situ hybridization techniques in the examination of sacroiliac joint biopsy specimens from patients with ankylosing spondylitis. Arthritis Rheum 38(4):499–505
45. McGonagle D et al (2002) Histological assessment of the early enthesitis lesion in spondyloarthropathy. Ann Rheum Dis 61(6):534–537
46. Lories RJ, Schett G (2012) Pathophysiology of new bone formation and ankylosis in spondyloarthritis. Rheum Dis Clin North Am 38(3):555–567
47. Brown MA et al (1997) Susceptibility to ankylosing spondylitis in twins: the role of genes, HLA, and the environment. Arthritis Rheum 40(10):1823–1828
48. van der Linden S, Valkenburg HA, Cats A (1984) Evaluation of diagnostic criteria for ankylosing spondylitis. A proposal for modification of the New York criteria. Arthritis Rheum 27(4):361–368
49. Braun J, Baraliakos X (2011) Imaging of axial spondyloarthritis including ankylosing spondylitis. Ann Rheum Dis 70(Suppl 1):i97–i103
50. Maksymowych WP et al (2013) Suppression of inflammation and effects on new bone formation in ankylosing spondylitis: evidence for a window of opportunity in disease modification. Ann Rheum Dis 72(1):23–28
51. Tan S et al (2013) Quantitative syndesmophyte measurement in ankylosing spondylitis using CT: longitudinal validity and sensitivity to change over 2 years. Ann Rheum Dis
52. Gladman DD et al (1987) Psoriatic arthritis (PSA)—an analysis of 220 patients. Q J Med 62 (238):127–141
53. Gladman DD et al (2011) Do patients with psoriatic arthritis who present early fare better than those presenting later in the disease? Ann Rheum Dis 70(12):2152–2154
54. Madland TM et al (2005) Prevalence, disease manifestations, and treatment of psoriatic arthritis in Western Norway. J Rheumatol 32(10):1918–1922
55. Moll JM, Wright V (1973) Familial occurrence of psoriatic arthritis. Ann Rheum Dis 32 (3):181–201
56. Kane D et al (2003) A prospective, clinical and radiological study of early psoriatic arthritis: an early synovitis clinic experience. Rheumatology (Oxford) 42(12):1460–1468
57. Gladman DD, Farewell VT, Nadeau C (1995) Clinical indicators of progression in psoriatic arthritis: multivariate relative risk model. J Rheumatol 22(4):675–679
58. Kaltwasser JP et al (2004) Efficacy and safety of leflunomide in the treatment of psoriatic arthritis and psoriasis: a multinational, double-blind, randomized, placebo-controlled clinical trial. Arthritis Rheum 50(6):1939–1950

59. Clegg DO et al (1996) Comparison of sulfasalazine and placebo in the treatment of reactive arthritis (Reiter's syndrome). A Department of Veterans Affairs Cooperative Study. Arthritis Rheum 39(12):2021–2027
60. Braun J et al (2000) On the difficulties of establishing a consensus on the definition of and diagnostic investigations for reactive arthritis. Results and discussion of a questionnaire prepared for the 4th international workshop on reactive arthritis, Berlin, Germany, July 3–6, 1999. J Rheumatol 27(9):2185–2192
61. Morris D, Inman RD (2012) Reactive arthritis: developments and challenges in diagnosis and treatment. Curr Rheumatol Rep 14(5):390–394
62. Hannu T et al (2006) Reactive arthritis or post-infectious arthritis? Best Pract Res Clin Rheumatol 20(3):419–433
63. Carter JD et al (2010) Combination antibiotics as a treatment for chronic Chlamydia-induced reactive arthritis: a double-blind, placebo-controlled, prospective trial. Arthritis Rheum 62 (5):1298–1307
64. Lam GK, Bingham CO (2007) Reactive arthritis. In: Imboden JB, Stone JH, Hellman DB (eds) Current rheumatology diagnosis and treatment. McGraw-Hill, New York, pp 183–186
65. Zochling J, Brandt J, Braun J (2005) The current concept of spondyloarthritis with special emphasis on undifferentiated spondyloarthritis. Rheumatology (Oxford) 44(12):1483–1491
66. Sampaio-Barros PD et al (2010) Undifferentiated spondyloarthritis: a longterm followup. J Rheumatol 37(6):1195–1199
67. Viatte S, Plant D, Raychaudhuri S (2013) Genetics and epigenetics of rheumatoid arthritis. Nat Rev Rheumatol 9(3):141–153
68. Schett G et al (2011) Structural damage in rheumatoid arthritis, psoriatic arthritis, and ankylosing spondylitis: traditional views, novel insights gained from TNF blockade, and concepts for the future. Arthritis Res Ther 13(Suppl 1):S4
69. Aletaha D et al (2010) 2010 Rheumatoid arthritis classification criteria: an American College of Rheumatology/European League against rheumatism collaborative initiative. Arthritis Rheum 62(9):2569–2581
70. Bouchaud-Chabot A, Liote F (2002) Cervical spine involvement in rheumatoid arthritis. A Rev Joint Bone Spine 69(2):141–154
71. Gramling A, O'Dell JR (2012) Initial management of rheumatoid arthritis. Rheum Dis Clin North Am 38(2):311–325
72. Singh JA et al (2012) 2012 update of the 2008 American College of Rheumatology recommendations for the use of disease-modifying antirheumatic drugs and biologic agents in the treatment of rheumatoid arthritis. Arthritis Care Res (Hoboken) 64(5):625–639
73. Wolfs JF et al (2009) Neurologic outcome of surgical and conservative treatment of rheumatoid cervical spine subluxation: a systematic review. Arthritis Rheum 61(12):1743–1752
74. Jacobson JA et al (2008) Radiographic evaluation of arthritis: inflammatory conditions. Radiology 248(2):378–389
75. McQueen FM, Ostergaard M (2007) Established rheumatoid arthritis - new imaging modalities. Best Pract Res Clin Rheumatol 21(5):841–856
76. Espada G et al (1988) Radiologic review: the cervical spine in juvenile rheumatoid arthritis. Semin Arthritis Rheum 17(3):185–195

Osteoporosis

Thomas Baum, Dimitrios C. Karampinos, Stefan Ruschke, Hans Liebl, Peter B. Noël and Jan S. Bauer

Abstract Osteoporosis is defined as a skeletal disorder characterized by compromised bone strength predisposing an individual to an increased risk for fracture. Osteoporotic fractures, in particular spine fractures, are associated with a high mortality and generate immense financial costs. Osteoporotic vertebral fractures frequently occur in absence of a specific trauma and may be asymptomatic. Since a prevalent vertebral fracture increases the risk of a subsequent fracture, the diagnosis of osteoporotic vertebral fractures is highly important to initiate appropriate therapy. Computer-assisted diagnostic tools for spine radiographs, dual-energy X-ray absorptiometry (DXA) and multi-detector computed tomography (MDCT) images have been developed to support radiologists to correctly diagnose and report osteoporotic vertebral fractures. The assessment of fracture risk at the spine has traditionally relied on the measurements of bone mineral density (BMD) by using DXA. However, BMD values of subjects with versus without osteoporotic fractures overlap. Bone strength reflects the integration of BMD and bone quality. The latter can be partly determined by measurements of bone microstructure. High-resolution MDCT allows for the assessment of trabecular bone microstructure at the spine.

T. Baum (✉) · D.C. Karampinos · S. Ruschke · H. Liebl · P.B. Noël
Department of Radiology, Klinikum rechts der Isar, Technische Universität München, Munich, Germany
e-mail: thbaum@gmx.de

D.C. Karampinos
e-mail: dimitrios.karampinos@tum.de

S. Ruschke
e-mail: stefan.ruschke@tum.de

H. Liebl
e-mail: lieblhans@gmail.com

P.B. Noël
e-mail: peter.noel@tum.de

J.S. Bauer
Section of Neuroradiology, Klinikum rechts der Isar, Technische Universität München, Munich, Germany
e-mail: jsb@tum.de

© Springer International Publishing Switzerland 2015

S. Li and J. Yao (eds.), *Spinal Imaging and Image Analysis*,
Lecture Notes in Computational Vision and Biomechanics 18,
DOI 10.1007/978-3-319-12508-4_3

MDCT-based trabecular bone microstructure parameters and finite element models have shown to improve the prediction of bone strength beyond DXA-based BMD and revealed pharmacotherapy effects, which were partly not captured by BMD. Furthermore, recent studies demonstrated that quantitative magnetic resonance imaging (MRI) including proton single-voxel magnetic resonance spectroscopy (^{1}H-MRS) and chemical shift-based water-fat imaging techniques quantifying bone marrow fat content at the spine may provide complementary information for diagnosing osteoporosis and assessing vertebral fracture risk.

1 Introduction

This chapter focuses on osteoporosis imaging at the spine and is structured into five parts: After a background section, imaging techniques and post-processing methods are outlined to correctly diagnose osteoporotic vertebral fractures. Subsequently, bone mineral density (BMD) measurements, which have traditionally been used for the assessment of osteoporosis, are presented. Measurements of bone microstructure and bone marrow fat content at the spine, which have been already used or have been emerging for predicting osteoporosis-related fracture risk and evaluating therapy response beyond BMD, are discussed in the last sections.

2 Background

Osteoporosis is defined as a skeletal disorder characterized by compromised bone strength predisposing an individual to an increased risk for fracture [1]. Bone strength primarily reflects the integration of bone mineral density (BMD) and bone quality including bone microstructure, turnover, and damage accumulation (e.g. microfractures). Osteoporotic subjects show a loss of BMD and deterioration of bone quality (Fig. 1). Osteoporosis is classified as either primary or secondary. Primary osteoporosis results from the cumulative bone loss due to ageing and the corresponding changes of sex hormones. It is further divided in type I and type II osteoporosis. Type I osteoporosis affects women after menopause, while type II osteoporosis can be found in elderly men as well as women and is also named senile osteoporosis. Secondary osteoporosis results from medications (e.g. glucocorticoids) or other conditions (e.g. hypogonadism). The most common form is the primary, type I (postmenopausal) osteoporosis.

The clinically most important fracture sites are the radius, hip and spine. It has been demonstrated that osteoporotic vertebral and hip fractures are associated with a reduced quality of life [2, 3]. Furthermore, vertebral and hip fractures are associated with an increased mortality [4–6]. In 2010, 22 million women and 5.5 million men were estimated to have osteoporosis in the European Union [7]. The number of

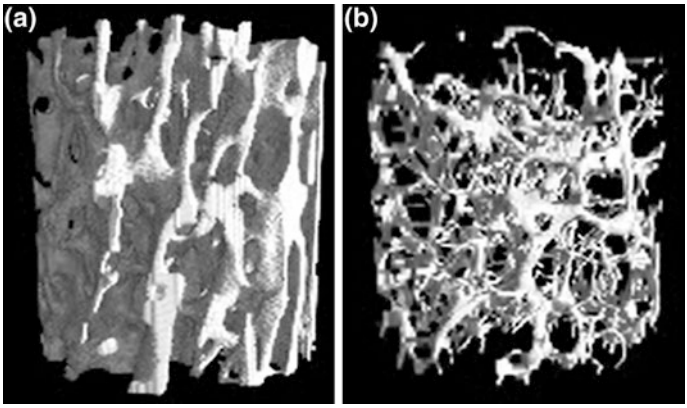


Fig. 1 Trabecular bone specimens of T10 from a normal (**a**) and an osteoporotic (**b**) subject. 3D reconstructions of micro-CT scans with a spatial resolution of 26 μm^3. Note the bone loss and rarefication in (**b**) compared to (**a**)

incident osteoporotic fractures amounted to 3.5 million, comprising 620,000 hip fractures, 520,000 vertebral fractures, 560,000 forearm fractures, and 1,800,000 other fractures. The economic burden of prevalent and incident osteoporotic fractures was estimated at €37 billion. Prevalent and incident osteoporotic fractures also accounted for 1,180,000 quality-adjusted life years lost during 2010. Due to the aging population, the prevalence of osteoporosis and consecutively the incidence of osteoporotic fractures is expected to increase [8]. In the European Union, the costs are expected to increase on average by 25 % in 2025 [7]. Similar projections have been reported for the United States [9]. Therefore, osteoporosis is classified as a public health problem.

The World Health Organisation (WHO) based the diagnosis of osteoporosis on the measurement of BMD at the spine and hip using dual-energy X-ray absorptiometry (DXA) [10]. Subjects with BMD values 2.5 standard deviations below the mean of the reference population consisting of healthy young adult women are classified as osteoporotic (T-score <-2.5), and subjects with BMD values ranging from 2.5 to 1.0 standard deviations below the mean of the reference population are classified as osteopenic (T-score between -2.5 and -1.0). However, T-scores and BMD values of subjects with and without osteoporotic fractures overlap [11, 12]. Schuit et al. [11] assessed in a prospective study baseline BMD and incidence of non-vertebral fractures during follow-up in 7,806 men and women aged 55 years and older. They reported that only 44 % of all non-vertebral fractures occurred in women with a T-score below -2.5. In men, this percentage was even lower (21 %). Similar findings were reported for incident osteoporotic vertebral fractures [12]. Thus, the BMD thresholds for the pharmacological intervention to prevent fractures are often inadequate. Subjects at high risk for osteoporotic fractures may not be identified and the necessary pharmacological treatment is not initiated. This is particularly regrettable, since the current osteoporosis medications including

bisphosphonates and anti-receptor activator of NF-κB ligand (RANKL) have demonstrated to efficiently reduce the incidence of osteoporotic fractures [13, 14]. Therefore, the Fracture Risk Assessment Tool (FRAX) has been introduced which uses easily obtainable clinical risk factors to estimate a 10-year fracture probability in order to provide a better clinical guidance for treatment decisions [15, 16]. Age, sex, weight, height, personal history of fracture, parental history of hip fracture, smoking status, glucocorticoid intake, rheumatoid arthritis, secondary osteoporosis, alcohol use, and femoral BMD value are inquired for this purpose. Bone turnover markers may be useful for monitoring osteoporosis treatment, e.g. annual infusion of zoledronic acid reduced bone turnover markers and explained much of the observed fracture risk reduction [17]. However, measurements of bone turnover markers are not included in algorithms for fracture risk prediction at the moment due to the lack of data [18].

The most important role of osteoporosis imaging at the spine is the correct assessment of vertebral fracture status and BMD measurements, which are outlined in the following two sections. FRAX does not directly assess bone strength and quality, parameters in particular important for monitoring drug effects. Therefore, imaging techniques measuring bone microstructure and bone marrow fat content, which have been already used or have been emerging to predict bone strength and assess bone quality, are presented in the last sections of this chapter.

3 Diagnosis of Osteoporotic Vertebral Fractures

A prevalent osteoporotic fracture increases the risk of a subsequent fracture, independent on BMD [19–22]. Therefore, the correct diagnosis and reporting of prevalent osteoporotic vertebral fractures by radiologists is highly important to initiate appropriate therapy. Osteoporotic vertebral fractures most commonly occur between thoracic vertebra 5 (T5) and lumbar vertebra 5 (L5).

Several scoring methods for osteoporotic vertebral fractures have been introduced [23]. Quantitative morphometry (QM) obtains ratios from direct vertebral body height measurements to define osteoporotic vertebral fractures. Semiquantitative (SQ) methods are based on the visual grading of fractures by using specific height and area reduction criteria. The algorithm-based qualitative (ABQ) method outlines a scheme to systematically rule out non-fracture deformities and diagnoses osteoporotic vertebral fractures based on endplate depression.

QM measures anterior, middle, and posterior vertebral body heights and calculates ratios between these heights [24, 25]. Vertebral fractures are defined by thresholds, e.g. three standard deviations difference in height ratios from normal population means or generally 15 % reduction in height ratios [26]. The advantages of the QM scoring systems are the better reproducibility and objectivity compared to the SQ and ABQ methods. The drawback of QM is the relatively high expenditure of time which is critical in clinical routine.

The most widely used scoring method is the SQ grading system introduced by Genant et al. [27]. Osteoporotic vertebral fractures are graded on visual inspection and without direct vertebral measurement: grade 0: normal, grade 1: mildly deformed (approximately 20–25 % reduction in anterior, middle, and/or posterior height and a reduction of area 10–20 %), grade 2: moderately deformed (approximately 25–40 % reduction in any height and a reduction in area 20–40 %), and grade 3: severely deformed (approximately 40 % reduction in any height and area). Furthermore, the spinal fracture index (SFI) can be calculated by summing the individual vertebral body grades. The advantage of SQ methods is the relatively little time needed for performing the grading. The disadvantage of the SQ methods is the lower reproducibility and worse objectivity compared to the QM grading systems. In contrast to the ABQ method, the Genant scoring system does not account for several other important characteristics of vertebral fracture including endplate deformity or buckling of cortices.

The ABQ method uses a scheme to systematically rule out non-fracture deformities [28]. Osteoporotic vertebral fractures are diagnosed based on the assumption that these fractures always involve the fracture of the endplate. Thus, the evidence of endplate fracture and not the variation in vertebral shape is the primary indicator of osteoporotic fracture. The advantage of the ABQ method is the better differentiation of true fractures from non-fracture deformities, in contrast to the QM and SQ scoring systems where no allowance is made for variation in vertebral dimensions at different vertebral levels or short vertebral heights associated with Scheuermann's disease, scoliosis etc. A skilled reader is needed for the ABQ method to differentiate accurately between vertebral fractures and non-fracture deformities, which is a disadvantage compared to the QM and SQ scoring systems.

Radiography of the thoracic and lumbar spine is the standard imaging modality used for the initial assessment of osteoporotic vertebral fractures (Fig. 2). The dose from a lateral radiograph of the thoracic and lumbar spine is about 0.6 mSv [29].

Vertebral Fracture Assessment (VFA) by using newer generations of DXA scanners allows for imaging of the thoracic and lumbar spine to assess prevalent osteoporotic vertebral fractures [30, 31]. It can be combined with the DXA-based BMD measurements which is advantageous. VFA has lower radiation exposures than radiographs with reported doses in the range from 0.002 to 0.05 mSv [29]. Moderate and severe osteoporotic fractures can be accurately identified by VFA, but caution is necessary when vertebrae are evaluated in the presence of degenerative changes of the spine [32]. In these cases, radiographs with their superior spatial resolution allow the detection of more subtle abnormalities. Furthermore, Buehring et al. reported that VFA is dependent on instrument and reader [33].

Multi-detector computed tomography (MDCT) and magnetic resonance imaging (MRI) allow for a 3D visualization of the spine and are used as advanced diagnostic tools for osteoporotic fractures [34]. Both imaging techniques are substantially more expensive than radiography. Compared to MDCT, MRI has the advantage that it lacks ionizing radiation. Bone marrow edema can be detected by using MRI. Its presence can differentiate a recent from an old osteoporotic vertebral fracture (Fig. 3). In contrast to MRI, the integrity and shape of the vertebrae, particularly the

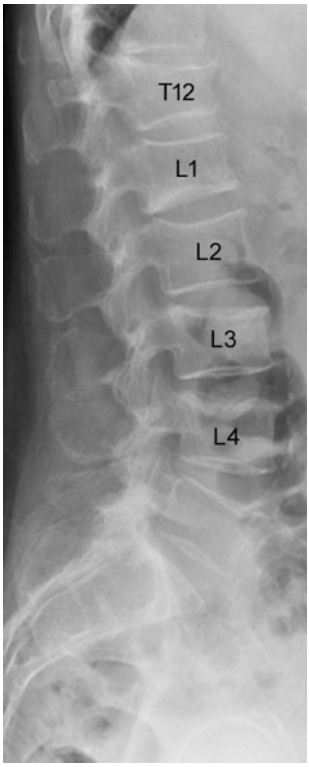


Fig. 2 Lateral radiograph of the lower thoracic and lumbar spine of a 63-year old woman with osteoporotic vertebral fractures from T12 to L4

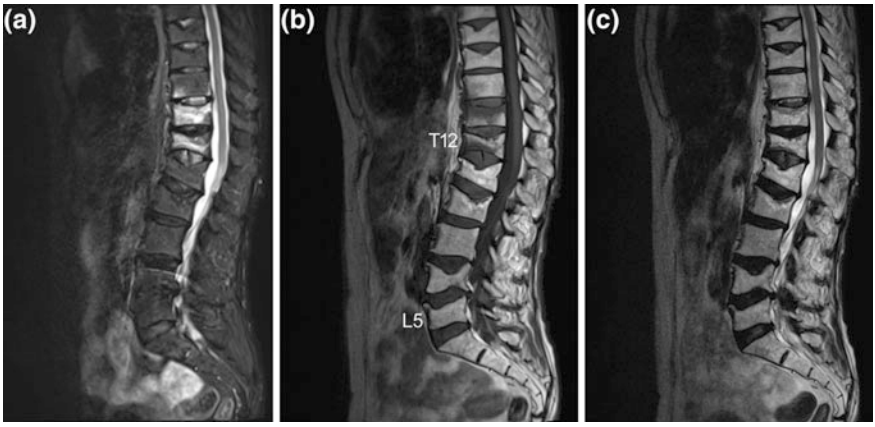


Fig. 3 Sagittal **a** short tau inversion recovery (STIR), **b** T1-weighted, and **c** T2-weighted MRI sequences of the thoracic and lumbar spine. Note the presence of bone marrow edema (hyperintense STIR and hypointense T1 signal, respectively) in T11 and T12 as well as close to the lower endplate in T10 as indication for recent osteoporotic vertebral fractures. In contrast, note the absence of bone marrow edema in the old osteoporotic vertebral fractures of T8, T9, L1, L2, L4, and L5

cortical margins, can be directly visualized by using MDCT. It is less expensive than MRI and may provide important information for differentiating between osteoporotic and malignant fractures and assessing bone matrix changes [34]. Low-dose MDCT protocols for the visualization of the spine were reported with a dose of 2.2 mSv for men and 3.3 mSv for women [35]. Therefore, radiographs should be primarily used for the initial assessment of osteoporotic vertebral fractures. However, clinicians have to be aware that MDCT can more accurately assess vertebral fractures than standard radiographs and has to be used in unclear cases. Bauer et al. compared the performance of lateral radiographs and sagittal reformations of axial MDCT images in detecting osteoporotic vertebral fractures [35]. They examined 65 vertebrae which were harvested from 21 human cadaver spines with a 64-row MDCT scanner. Ninety-five percent of the fractures could be identified by using sagittal reformations of 1 mm slice thickness, but 18 % of the fractures were missed on the radiographs. Thus, the authors concluded that sagittal MDCT reformations could more accurately assess vertebral fractures than standard radiographs. Furthermore, sagittal reformations of axial MDCT in-vivo images of the spine significantly improved the detection of osteoporotic vertebral fractures and other spine abnormalities, compared to axial images [36].

However, osteoporotic vertebral fractures often occur in absence of a specific trauma and are asymptomatic. Thus, they frequently do not come to clinical attention and dedicated imaging for the diagnosis of osteoporotic vertebral fractures is not performed. Consequently, initiation of appropriate therapy is delayed. Therefore, it is highly important to report prevalent osteoporotic vertebral fractures in routine chest radiographs and routine thoracic/abdominal MDCT images, which are one of the most frequent performed radiologic examinations (Fig. 4). However, osteoporotic vertebral fractures are underdiagnosed in these non-dedicated images [37–40]. Therefore, the International Osteoporosis Foundation (IOF) and the European Society of Musculoskeletal Radiology (ESSR) have started a teaching initiative to raise the awareness of radiologists to report prevalent osteoporotic fractures at the spine, which is well suited to improve the accurate diagnosis and reporting of prevalent vertebral fractures on lateral chest radiographs [41]. Furthermore, the development of semi-automatic segmentation techniques for lumbar radiographs and MDCT images as computer-assisted detection tools may support radiologists to correctly diagnose and report prevalent osteoporotic vertebral fractures [42, 43]. Roberts et al. [42] presented a semi-automatic determination of detailed vertebral shape from lumbar radiographs using active appearance models. The vertebral body outlines were manually annotated by radiologists in 670 lumbar radiographs to obtain a training set. This was used to build statistical models of vertebral shape and appearance using triplets of vertebrae. In order to segment the vertebrae, the models were refitted using a sequence of active appearance models of vertebral triplets. The accuracy achieved on normal vertebrae was good. However, the accuracy performance deteriorated with increasing fracture grade, but even in fractured vertebrae, point-to-line accuracy was below 2 mm in 79 % of the cases. The authors concluded that the located detailed vertebral shapes may enable the development of more powerful quantitative classifiers of osteoporotic vertebral

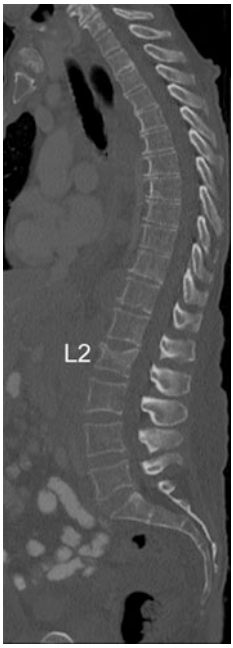


Fig. 4 Sagittal reformation with a slice thickness of 3 mm of a routine thoracic and abdominal MDCT. Note the osteoporotic vertebral fracture of L2

fracture. These active appearance models can also be used in DXA-based VFA images to detect osteoporotic vertebral fractures [44]. Baum et al. [43] developed a prototype algorithm for automatic spine segmentation in routine thoracic and abdominal MDCT images and used this algorithm to automatically detect osteoporotic vertebral fractures. The algorithm automatically localized and identified the vertebrae. Then, each vertebra was automatically segmented by using corresponding vertebrae surface shape models that were adapted to the original images (Fig. 5). Finally, anterior, middle, and posterior height of each segmented vertebra was automatically determined, and anterior-posterior-ratio and middle-posterior-ratio were computed to diagnose osteoporotic vertebral fractures. The prototype algorithm demonstrated a good performance for the automatic detection of prevalent and incident osteoporotic vertebral fractures cross-sectionally and longitudinally, respectively.

4 BMD Measurements

As outlined in the Background section, measurements of BMD are performed to diagnose osteoporosis. Subjects at high risk for osteoporotic fractures can be identified, so that pharmacological treatment can be initiated [1]. Furthermore, BMD examinations are used to monitor treatment response [34]. Fractured

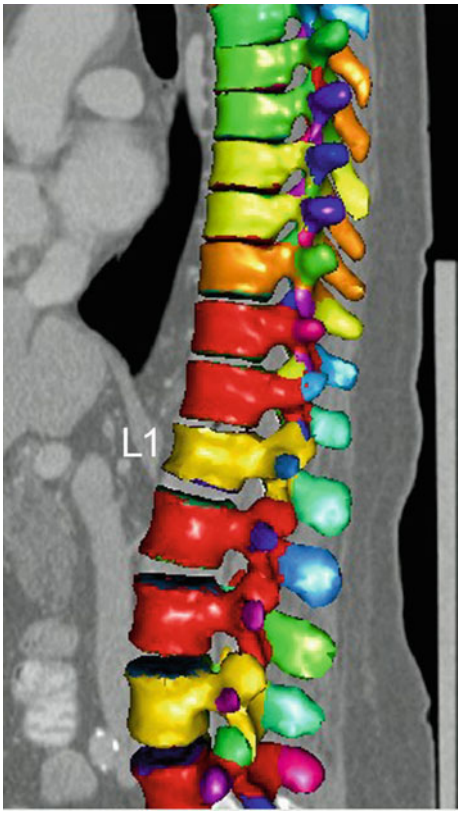


Fig. 5 Automatic detection, identification, and segmentation of the vertebrae in routine thoracic and abdominal MDCT images. Note the automatically detected osteoporotic vertebral fracture of L1

vertebrae are not analyzed, since inclusion of the endplate will cause overestimation of BMD.

DXA is the most common technique to measure BMD [45]. It is usually performed at the forearm, spine, and hip (Fig. 6). BMD is measured as areal value in g/cm^2 calcium hydroxyapatite. Compared with alternative bone densitometry techniques, DXA-based BMD results of the spine and hip are the only ones which can be interpreted using the WHO T-score definition of osteoporosis. The T-score is calculated by comparing the measured BMD value of the subject with the mean BMD value of a reference population consisting of healthy young adult women:

$$T\text{-}score = \frac{measured\ BMD_{subject} - mean\ BMD_{reference\ population}}{standard\ deviation\ BMD_{reference\ population}}$$

Subjects with T-scores between −2.5 and −1.0 are classified as osteopenic and those with T-scores <−2.5 as osteoporotic [10]. The Z-score expresses the comparison of the measured BMD value of the subject with the mean BMD value of an age-matched reference population:

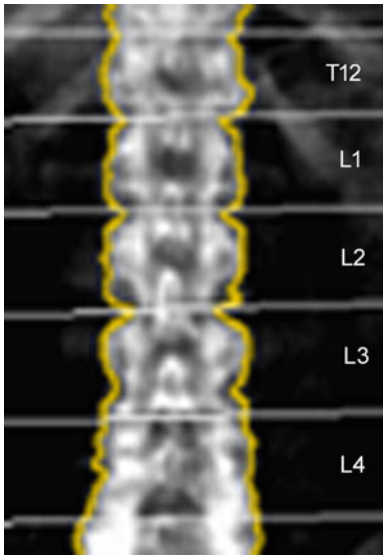


Fig. 6 Representative DXA image from T12 to L4. Measurements of areal BMD (mg/cm²) are obtained for each vertebra

$$Z\text{-}score = \frac{measured\ BMD_{subject} - mean\ BMD_{age-matched\ population}}{standard\ deviation\ BMD_{age-matched\ population}}$$

The radiation dose for DXA measurements is relatively low and amount to 0.013 μSv at the spine [29]. BMD can be accurately determined by DXA which is particularly important for the longitudinal assessment of treatment response. Reproducibility errors expressed as coefficient of variation (CV) ranged between 1.0 and 1.5 % for the spine [45]. However, aortic sclerosis, degenerative disc disease, and scoliosis represent significant error sources for DXA-based BMD measurements.

Quantitative computed tomography (QCT) at the spine avoids these error sources. QCT-based BMD measurements at the spine are performed with clinical whole-body multi-detector computed tomography (MDCT) scanners and are determined as volumetric values in mg/cm^3 calcium hydroxyapatite [46]. Thus, BMD values obtained by QCT are not size dependent in contrast to DXA-based areal BMD. A further advantage of QCT is the separate measurement of cortical and trabecular BMD (Fig. 7). Since the trabecular compartment is the metabolically more active one, treatment response can be assessed more accurately by using QCT compared to DXA. QCT-based BMD is usually measured in the lumbar vertebrae 1–3 (L1–L3). Subjects with a trabecular BMD averaged from L1 to L3 between 80 and 120 mg/cm^3 are classified as osteopenic and those below 80 mg/cm^3 as osteoporotic [47]. A calibration phantom in the table mat is required to be scanned with the subject to convert the voxels' attenuation values in Hounsfield Unit into BMD values in mg/cm^3 calcium hydroxyapatite. BMD measurements at the spine can be performed as 2D single-slice QCT with a slice thickness of 8–10 mm. The scanning

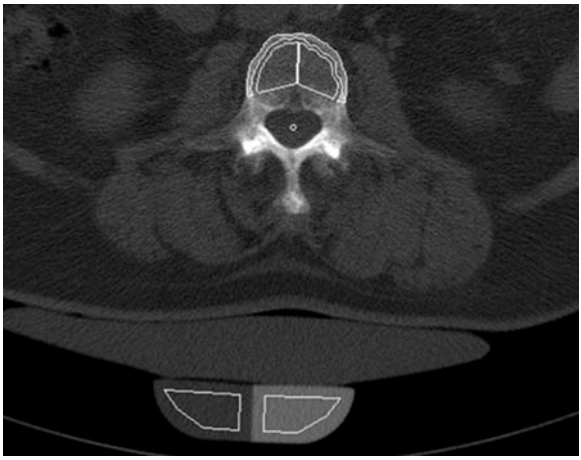


Fig. 7 2D single-slice QCT of L3 with a slice thickness of 10 mm. Volumetric BMD measurements were performed in the cortical and trabecular compartment of L3 (*white contours*). Note the calibration phantom in the table mat to convert the voxels' attenuation values in Hounsfield Unit into BMD values in mg/cm³ calcium hydroxyapatite

plane is selected through the middle of each vertebra parallel to the endplates. The radiation dose for scout image and 3 slices of 10 mm thickness range between 0.2 and 1.0 mSv [29, 46]. Alternatively, 3D QCT measurements can be obtained allowing more sophisticated analysis of cortical and trabecular bone and the imaging of trabecular bone microstructure. However, 3D QCT scans are associated with higher radiation doses (about 1.5 mSv) [29, 46]. Reproducibility errors for 2D single-slice QCT at spine expressed as coefficient of variation (CV) range between 1.4 and 4.0 %, those for 3D QCT between 1.3 and 1.7 % [46]. Compared to DXA, the drawbacks of QCT are the higher radiation dose and the fact that the WHO definition of osteoporosis (T-score <−2.5 using DXA) is not applicable. Therefore, QCT is not as commonly used as DXA and is predominantly performed in selected patient populations and clinical trials [48].

Since routine thoracic and abdominal MDCT is one of the most frequently used radiologic examination, it would be beneficial to use the obtained MDCT images to conduct additionally BMD measurements of the spine [49]. Thus, further radiation exposure is avoided. Particularly, patients with cancer would benefit, since they routinely undergo MDCT and are at increased risk of osteoporosis due to the cancer related treatment [50]. Most routine thoracic and abdominal MDCT examinations are performed with intravenous contrast medium. Therefore, the obtained BMD values have to be converted to standard QCT equivalent BMD values which are better useable for fracture risk prediction. Bauer et al. [51] and Link et al. [52] demonstrated the feasibility to determine BMD values at the lumbar spine in axial images of non-dedicated routine abdominal contrast-enhanced MDCT. Baum et al. described lumbar BMD measurements without dedicated software and with

low time effort in the sagittal reformations of non-dedicated routine abdominal contrast-enhanced MDCT [53]. This offers the possibility to determine lumbar BMD values in the same reformations which are known for a substantial better detection of osteoporotic vertebral fractures [35, 36]. Thus, radiologists can assess vertebral fracture status and BMD in the sagittal reformations in an acceptable time which is critical in clinical routine. Baum et al. determined in ten patients standard QCT-based BMD of L1–L3 and apparent BMD of L1–L3 in the sagittal reformations of routine abdominal contrast-enhanced MDCT images [53]. Apparent BMD values of contrast-enhanced MDCT were on average 56 mg/cm^3 higher than those of standard QCT. A correlation coefficient of r = 0.94 was calculated for the BMD values of MDCT and standard QCT with the conversion equation $BMD_{QCT} = 0.69 \times BMD_{MDCT}$-11 mg/ml. Using this conversion equation, lumbar BMD measurements in the sagittal reformations of routine abdominal contrast-enhanced MDCT images could adequately differentiate patients with versus without osteoporotic vertebral fractures. Furthermore, baseline converted lumbar BMD values predicted incident osteoporotic vertebral fractures during a follow-up of 20 ± 12 months [54]. The BMD measurements in the sagittal reformations were performed by placing manually circular regions of interest (ROIs) in the ventral halves of the trabecular compartment of the vertebral bodies of L1–L3, in each case equidistant to both endplates (Fig. 8). The attenuation values measured in the ROIs in Hounsfield Unit were converted into mg/cm^3 calcium hydroxyapatite using a

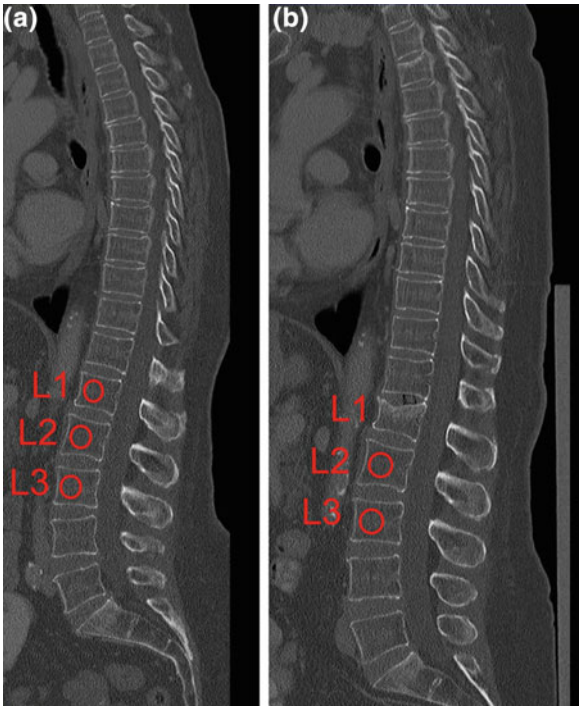


Fig. 8 BMD measurements in sagittal reformations of routine contrast-enhanced MDCT in a patient at baseline (**a**) and follow-up (**b**). Circular ROIs (*red*) were manually placed in L1–L3. Note the incident osteoporotic vertebral fracture of L1 at 8-month follow-up

reference phantom integrated into the table mat. Short- and long-term reproducibility errors for BMD measurements in the sagittal reformations amounted 2.09 and 7.70 %, respectively [53]. Baum et al. [53, 54] used only MDCT examinations with a scan delay of 70 s after intravenous contrast medium injection. Acu et al. [55] pointed out that scan protocols with different scan delay times after intravenous contrast medium injection significantly change the apparent BMD values. They reported increasing apparent BMD values with longer scan delay times. Thus, the scan delay time after intravenous contrast medium injection has to be taken into account for the derivation of standard QCT equivalent BMD values from routine contrast-enhanced MDCT examinations. Lastly, Summers et al. [56] and Pickhardt et al. [57] demonstrated the feasibility to calculate BMD from computed tomographic colonography scans.

5 Measurements of Bone Microstructure

T-scores and BMD values of patients with versus without osteoporotic fractures overlap as outlined in the Background section [11, 12]. Bone strength reflects the integration of BMD and bone quality including bone microstructure [1]. Therefore, substantial research efforts have been undertaken to assess bone microstructure by using high-resolution imaging techniques to improve fracture risk prediction [58, 59]. Trabeculae have a diameter between 50 and 200 μm and the cortical thickness varies between 0.2 to 5 mm. Thus, the spatial resolution of the imaging techniques used for bone microstructure analysis is critical.

High-resolution peripheral quantitative computed tomography (hr-pQCT) allows for an isotropic spatial resolution of 82 μm^3 in-vivo with a relatively low effective dose of approximately 4 μSv for a scan at the distal radius or tibia [60]. However, hr-pQCT systems are limited to peripheral sites and cannot be applied to the spine.

Magnetic resonance imaging (MRI) lacks ionizing radiation. Bone tissue has low MR signal and consequently appears dark in most clinically accessible pulse sequences. Bone marrow has a relatively high MR signal, i.e. has positive contrast, depending on the fat content [fatty (yellow) or hematopoietic (red) bone marrow] and the applied pulse sequence [58, 61]. High-resolution MRI has been performed mostly at the peripheral skeleton such as radius, tibia, and calcaneus due to their easy accessibility. Voxel sizes up to 137 × 137 × 410 μm^3 were reported for high-resolution MRI at the distal radius [62, 63]. Due to higher field strength and sequence development, in-vivo MR imaging at the proximal femur as important clinical fracture site has become feasible [61, 64]. In contrast to peripheral sites, the proximal femur contains not only fatty bone marrow, but also hematopoietic bone marrow. The visualization of the trabeculae in the proximal femur is partly obscured by the dark, hematopoietic bone marrow. An even higher percentage of hematopoietic bone marrow is found in the vertebral bodies resulting in insufficient signal-to-noise ratios to obtain high-resolution MR images of the spine.

High-resolution MRI and hr-pQCT measurements of bone microstructure at peripheral sites revealed important information on the spine [59, 65]. To give an example, Ladinsky et al. [63] demonstrated that trabecular bone microstructure quantified with high-resolution MRI in postmenopausal women contributes to vertebral deformity burden independent of areal vertebral BMD. However, Eckstein et al. [66] reported a substantial heterogeneity of bone strength among clinically relevant skeletal sites and that its loss in osteoporosis may not represent a strictly systemic process. Therefore, the direct assessment of bone microstructure at the spine is advantageous to assess vertebral fracture risk and evaluate treatment response at this anatomical location.

MDCT is the only imaging technique for high-resolution bone imaging at the spine in-vivo [58]. Clinical whole-body MDCT can achieve a maximal in-plane spatial resolution of about 250 × 250 μm^2 with an axial slice thickness of 500 μm [67]. Thus, MDCT systems do not have the sufficient spatial resolution to reveal the true trabecular bone microstructure. However, trabecular bone microstructure parameters and finite element models (FEMs) assessed with MDCT and μCT (micro-CT) or hr-pQCT as standard of reference showed high correlations and predicted biomechanically determined bone strength equally well [68–70]. Bauer et al. harvested 20 cylindrical trabecular bone specimens from formalin-fixed human thoracic spines [68]. μCT images of the bone specimens were obtained with an isotropic voxel size of 20 μm^3 and corresponding MDCT images up to a voxel size of 230 × 230 × 500 μm^3. Trabecular bone microstructure parameters obtained from μCT and MDCT showed R^2 values up to 0.84. Furthermore, MDCT derived trabecular bone microstructure parameters demonstrated high correlations with biomechanically determined bone strength (R^2 values up to 0.81). Similarly, Baum et al. [69] examined formalin-fixed spinal segment units by using hr-pQCT (isotropic voxel size of 41 μm^3) and a clinical whole-body MDCT (spatial resolution of 250 × 250 × 600 μm^3). Corresponding images of a spinal segment unit acquired with MDCT and hr-pQCT as standard of reference are shown in Fig. 9. Correlations between trabecular bone microstructure parameters and biomechanically determined failure load amounted up to r = 0.86 using the hr-pQCT images, and up to r = 0.79 using the MDCT images. Correlation coefficients of failure load versus trabecular bone microstructure parameters obtained with HR-pQCT and MDCT were not significantly different. Furthermore, no differences in the performance of 64- and 320-slice MDCT scanners with respect to the depiction of trabecular bone microstructure were observed [71].

The calculation of bone microstructure parameters and FEMs in MDCT images of the spine requires several steps including image registration and segmentation. Multiple (semi-) automated image registration and segmentation algorithms have been developed to minimize time effort and reproducibility errors, which are particularly important for the assessment of change in longitudinal studies [72–76]. Thereby, regions of interest (ROIs) are drawn in the acquired images to define the outer contour of the vertebrae and consecutively certain areas of the trabecular bone.

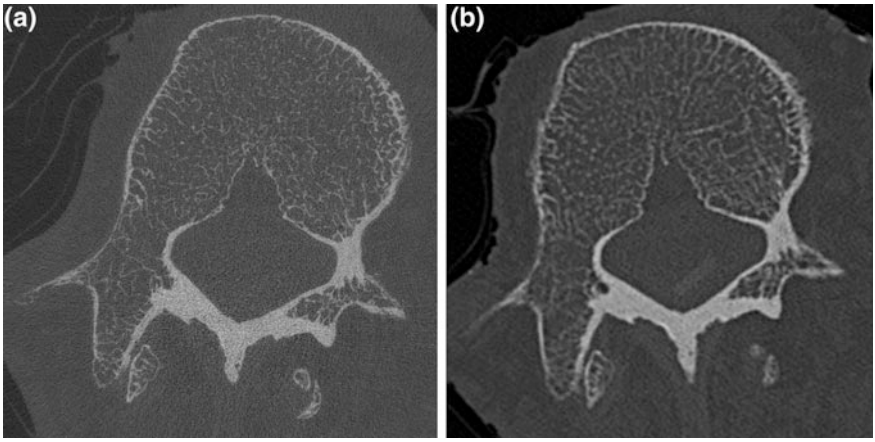


Fig. 9 Corresponding hr-pQCT (**a**) and MDCT (**b**) image of a formalin-fixed spinal segment unit. Spatial resolution was 41 μm³ at hr-pQCT and 250 × 250 × 600 μm³ at MDCT. Note the better depiction of the single trabeculae in the hr-pQCT image

Standard parameters for the assessment of trabecular bone microstructure can be calculated in the binarized MDCT images according to bone histomorphometry using the mean intercept length method [77]: Bone volume divided by total volume (BV/TV; bone volume fraction; [%]), trabecular number (TbN; [mm^{-1}]), trabecular separation (TbSp; [mm]), and trabecular thickness (TbTh; [mm]). In contrast to hr-pQCT and μCT, MDCT derived parameters are labeled as apparent values, since they cannot depict the true trabecular microstructure due to the limited spatial resolution. Furthermore, several advanced measures of trabecular bone microstructure have been introduced, e.g. non-linear topological parameters such as the Minkowski Functionals [78]. The appropriate definition of thresholds for image binarization is critical for the calculation of these trabecular bone microstructure parameters. The absolute values of these parameters vary with different selected thresholds due to partial-volume effects. An optimized, global threshold is usually chosen for MDCT images, so that subjects with dense trabecular bone microstructure do not have only bone voxels and osteoporotic subjects not only marrow voxels. To give an example, Baum et al. [69] applied an optimized, global threshold of 200 mg hydroxyapatite/cm^3 on vertebral bone specimens. To avoid the dependence of the results on the selected threshold, trabecular bone microstructure parameters have been introduced which do not require a threshold, e.g. the scaling index method [78, 79]. It reveals the local dimensionality of each voxel (i.e. more plate-like or rod-like structure). Thus, the transformation of trabecular bone from plate- to rod-like structures due to osteoporosis can be identified (Fig. 10). Elastic and shear moduli obtained from FEMs represent an alternative to the trabecular bone microstructure parameters to predict bone strength. FEMs can be calculated not only in ROIs in the trabecular bone, but also integrally for the whole vertebra including the cortical bone. This is advantageous, since it is well known that the

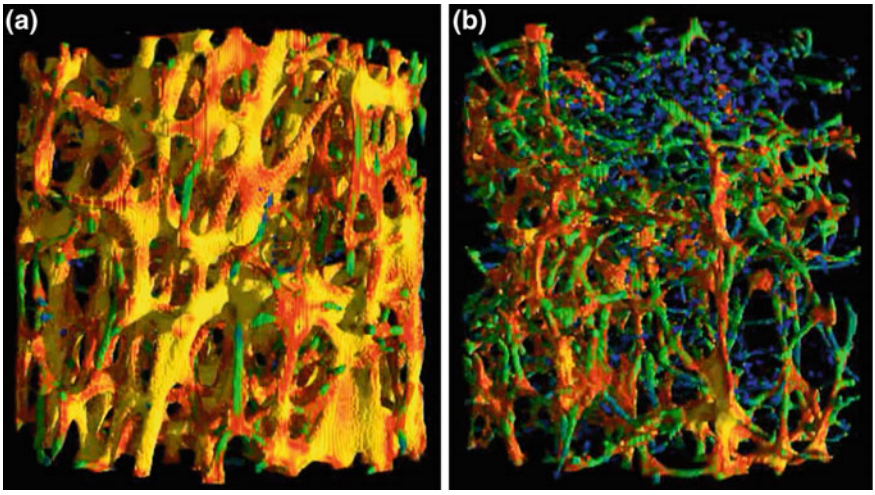


Fig. 10 Trabecular bone specimens of T10 from a normal (**a**) and an osteoporotic (**b**) subject. Color-coded 3D visualizations of the isotropic scaling indices α of micro-CT scans with a spatial resolution of 26 μm^3. α-values representing the local dimensionality of the bone voxels are increasing from *blue* over *green* to *red* color. Rod-like structures have lower local dimensionality than plate-like structures. **a** shows greater local dimensionality compared to **b** resulting from the transformation of trabecular bone from plate- to rod-like structures due to osteoporosis

cortical bone contributes substantially to the mechanical properties of the bone [80]. Information about bone geometry and bone density distribution is used for MDCT-based FEMs at the spine [81]. A representative FEM of a vertebra is displayed in Fig. 11.

MDCT-based trabecular bone microstructure parameters and FEMs at the spine were successfully validated in-vitro by using biomechanically determined failure load as gold standard [68–70, 78, 79, 82–84]. Trabecular bone microstructure parameters and FEMs provided better diagnostic performance for differentiating subjects with versus without osteoporotic vertebral fracture than BMD measurements. Ito et al. [67] computed MDCT-based trabecular bone microstructure parameters in the vertebrae of 82 postmenopausal women in-vivo including 39 women with and 43 without osteoporotic vertebral fracture. The microstructure parameters revealed higher relative risk for prevalent osteoporotic vertebral fracture than vertebral BMD as determined by DXA. The authors reported reproducibility errors for the trabecular microstructure parameters in cadaver specimens in the range from 0.67 to 12.30 %. Vertebral FEMs were not only able to detect prevalent osteoporotic vertebral fractures better than BMD measurements, but also predicted incident fractures in-vivo [85, 86]. Wang et al. [86] performed baseline MDCT-based FEMs to determine L1 vertebral compressive strength and a load-to-strength ratio in 306 men aged 65 years and older. An incident osteoporotic vertebral fracture was diagnosed in 63 subjects during follow-up of averaged 6.5 years. The area-under-the-curve for areal BMD (AUC = 0.76) was significantly lower than for

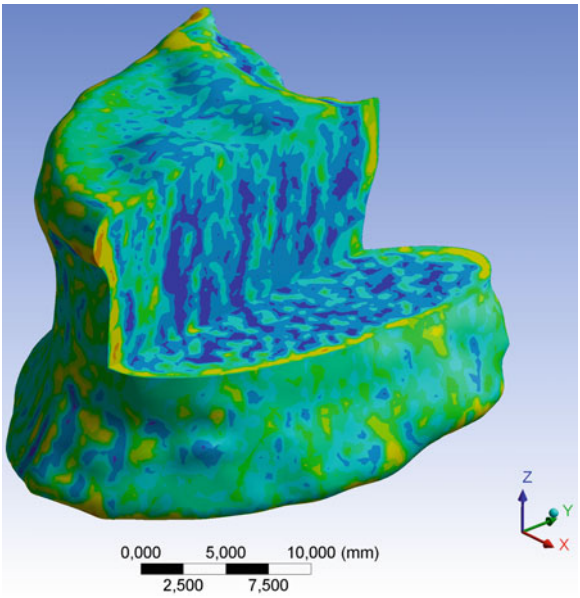


Fig. 11 MDCT-based FEM of a vertebra (T10) in-vitro. The BMD distribution is color-coded and used for the assignment of the material properties for each element of the FEM

strength (AUC = 0.83), volumetric BMD (AUC = 0.82), and the load-to-strength ratio (AUC = 0.82). Thus, FEM-based vertebral compressive strength and volumetric BMD consistently improved vertebral fracture risk assessment compared to areal BMD as assessed by DXA.

Furthermore, trabecular bone microstructure parameters and FEM revealed drug effects which were partly not captured by BMD measurements. Graeff et al. [87] performed high-resolution MDCT imaging of T12 in 65 postmenopausal women with established osteoporosis after 0, 6, and 12 months of teriparatide treatment. Interestingly, changes in trabecular bone microstructure parameters exceeded and were partially independent of changes in BMD. Thus, longitudinal analysis of trabecular bone microstructure at the spine offers information beyond BMD. Similar findings were reported with regard to the effects of teriparatide, alendronate, and risedronate on vertebral bone strength as assessed by MDCT-based FEMs [88–90].

These findings underline the importance of high-resolution bone imaging for fracture risk assessment and therapy monitoring. However, in-vivo MDCT imaging for trabecular bone microstructure analysis and FEMs is associated with an effective dose of estimated 3 mSv according to Graeff et al. [87]. This dose is in the upper range of the medically indicated radiation exposure. Therefore, these measurements are currently limited to research trials and cannot be used in clinical routine. In the future, it remains critical for MDCT imaging of bone microstructure at the spine to considerably reduce the radiation exposure. Newly developed CT reconstruction algorithms (e.g. iterative reconstruction) and flat-panel CT devices may have the potential to reliably assess trabecular bone microstructure with less radiation dose [91].

6 Measurements of Bone Marrow Fat Content

The bone marrow is the non-mineralized component of bone. The interaction of the mineralized and non-mineralized component plays an important role in the pathophysiology of age-related bone loss [92]. The bone marrow is more metabolically active and responsive than the mineralized component of bone. Therefore, it has been hypothesized that the quantification of bone marrow fat content could be used as a new diagnostic and therapeutic approach for osteoporosis. Bredella et al. [93] demonstrated that vertebral bone marrow fat is positively associated with visceral fat. Furthermore, it has been reported that visceral adiposity and the metabolic syndrome have potential detrimental effects on bone health [94, 95]. On the contrary, obese women are at decreased risk for developing osteoporosis [96].

In the lights of these conflicting results, considerable research effort has been undertaken recently to gain more insights into the bone marrow metabolism and its relationship with BMD, bone strength, and other body fat depots [97]. Proton single-voxel magnetic resonance spectroscopy (^{1}H-MRS) is the most widely used method to non-invasively quantify bone marrow fat in-vivo [98]. Other non-spectroscopic MR methods include T_1-weighted imaging and chemical shift-based water-fat separation techniques, e.g. the Dixon methods and the Iterative Decomposition of water and fat with Echo Asymmetry and Least-squares estimation (IDEAL) method [99–101]. All methods lack ionizing radiation. ^{1}H-MRS is considered as gold standard [102, 103]. It requires to prescribe a volume of interest in the exact desired anatomical location which can be technically demanding. Figure 12 shows a representative ^{1}H-MRS-based spectrum of L4: the spectrum shows the methyl ($-(CH_2)_n-CH_3$) peak at 0.9 ppm (peak 1), the superposition of the methylene ($-(CH_2)_n-$) peak at 1.30 ppm and the β-carboxyl ($-CO-CH_2-CH_2-$) peak at 1.59 ppm (peak 2), the superposition of the α-olefinic ($-CH_2-CH=CH-CH_2-$) peak at 2.00 ppm and the α-carboxyl ($-CO-CH_2-CH_2-$) peak at 2.25 ppm (peak 3), the water peak at 4.7 ppm (peak 4), and the olefinic ($-CH=CH-$) peak at 5.3 ppm (peak 5). ^{1}H-MRS-based bone marrow fat fraction (also named fat content) has been previously defined by the methylene ($-(CH_2)_n-$) fat peak at 1.30 ppm and the water peak at 4.7 ppm as the relative fat signal intensity amplitude in terms of a percentage of total signal intensity amplitude (fat and water) [102, 103]. Li et al. [104] performed ^{1}H-MRS in six subjects twice at four vertebral body levels (L1–L4) on the same day with repositioning between the two scans. The averaged coefficient of variation (CV) of vertebral bone marrow fat content was 1.7 %, suggesting good in-vivo reproducibility. T_1-weighted MR imaging is a conventional practice which is not technically demanding in terms of acquisition. A threshold is usually applied on the T_1-weighted images to define voxels as bone marrow voxels (i.e. bone marrow adipose tissue volume) in the segmented vertebra [105]. The bone marrow fat fraction is defined as the number of bone marrow voxels divided by the total number of voxels in the segmented vertebra. The applied threshold is often set at the same level as subcutaneous adipose tissue on the grey scale [101, 105, 106]. The intra- and interobserver reproducibility for the assessment of bone marrow adipose tissue in

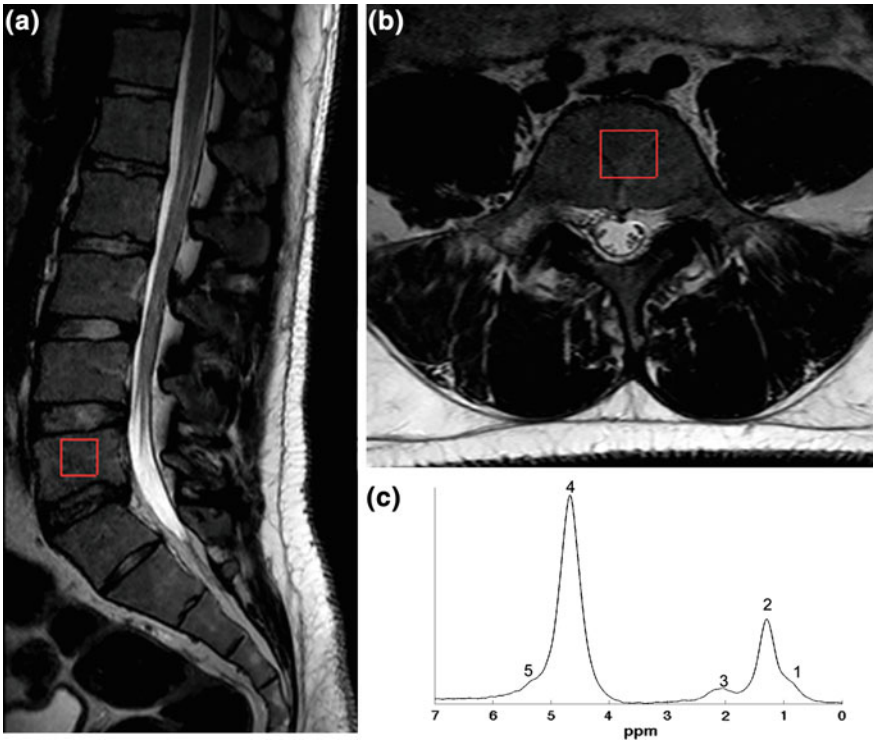


Fig. 12 ^{1}H-MRS of L4 at 3T: Sagittal (**a**) and axial (**b**) T_2-weighted images with the red box in L4 indicating the position for the employed single-voxel ^{1}H-MRS. The ^{1}H-MRS-based spectrum of L4 (**c**) shows the methyl ($-(CH_2)_n-CH_3$) peak at 0.9 ppm (peak 1), the superposition of the methylene ($-(CH_2)_n-$) peak at 1.30 ppm and the β-carboxyl ($-CO-CH_2-CH_2-$) peak at 1.59 ppm (peak 2), the superposition of the α-olefinic ($-CH_2-CH=CH-CH_2-$) peak at 2.00 ppm and the α-carboxyl ($-CO-CH_2-CH_2-$) peak at 2.25 ppm (peak 3), the water peak at 4.7 ppm (peak 4), and the olefinic ($-CH=CH-$) peak at 5.3 ppm (peak 5)

T_1-weighted MR images expressed as coefficient of variation (CV) amounted 1.0 and 2.6 %, respectively [101]. The main error source for the calculation of bone marrow adipose tissue volume in T_1-weighted MR images results from partial volume effects and the threshold selection, making the technique semi-quantitative [101]. Chemical shift-based water-fat separation techniques extract the bone marrow fat fraction based on the different precession frequencies of water and lipid hydrogen protons. Confounders such as T_2* decay, T_1 effects, the multi-spectral nature of fat, and noise bias have to be addressed for reliable fat quantification [107–109]. After the correction of these confounding factors, the proton density fat fraction (PDFF) can be quantified [110]. Regions of interest (ROIs) have to be placed in vertebral body, e.g. by using the first in-phase series in mid-sagittal view of the vertebral bodies [107]. Then, the ROIs are copied onto the reconstructed PDFF map to determine the vertebral bone marrow fat fraction [107].

Shen et al. [101] compared measurements of vertebral bone marrow fat fraction of L3 and the femoral neck in 27 postmenopausal women as assessed by T_1-weighted MR imaging, the Dixon method, and ^{1}H-MRS. Correlations between the obtained bone marrow fat fractions of the three MR methods ranged from r = 0.78 to 0.88 in L3. Karampinos et al. [111] measured bone marrow fat fraction at the proximal femur in 7 healthy volunteers by using chemical shift-based water-fat imaging and ^{1}H-MRS. They reported a significant underestimation of the fat fraction as determined by a ^{1}H-MRS model which did not account for short T_2* species compared to the chemical shift-based water-fat imaging-based fat fraction. However, a good equivalency was observed between the fat fraction by using a ^{1}H-MRS model accounting for short T_2* species and the chemical shift-based water-fat imaging-based fat fraction (R^2 = 0.87). Thus, the comparison of bone marrow fat fraction derived from different MR methods is limited and potential confounders have to be taken into account.

Griffith et al. [102] performed DXA measurements of L1–L4 and ^{1}H-MRS of L3 in 82 men with a mean age of 73 years. According to their DXA results, 42 subjects were classified as healthy (mean T-score of 0.8 ± 1.1), 23 subjects as osteopenic (mean T-score of 1.6 ± 0.4), and 17 subjects as osteoporotic (mean T-score of 3.2 ± 0.5). ^{1}H-MRS-based vertebral bone marrow fat content of L3 was significantly increased in osteoporotic subjects (58.23 ± 7.8 %) and osteopenic subjects (55.68 ± 10.2 %) compared to subjects with normal BMD (50.45 ± 8.7 %). Similar results were observed in a study population of 103 postmenopausal women older than 65 years of age [103]. Based on DXA measurements of L1–L4, 18 women had a normal BMD, 30 women were osteopenic, and 55 women had osteoporosis. Vertebral bone marrow fat content of L3 was significantly increased in the osteoporotic group (67.8 ± 8.5 %) when compared with that of the normal BMD group (59.2 ± 10.0 %). Similar results were observed for T_1-weighted MRI and chemical shift-based water-fat imaging. Shen et al. [105] obtained T_1-weighted MRI in 210 healthy African-American and Caucasian men and women aged 38–52 years. Hip and lumbar spine BMD were measured by DXA. Pelvic, hip, and lumbar spine bone marrow adipose tissue as assessed by T_1-weighted MRI showed negative correlations with hip and lumbar spine BMD (r = −0.399 to −0.550). These negative relationships remained significant after adjusting for demographics and body composition. Kühn et al. studied 51 patients (28 female; mean age 69.7 ± 9.0 years) who underwent DXA from L1–L4 to measure BMD and chemical shift-based water-fat imaging to determine vertebral bone marrow fat fraction [107]. The 173 investigated vertebral were divided into three groups (healthy, osteopenic, and osteoporotic) based on their T-score. The area-under-the-curve to differentiate between normal and osteoporotic vertebrae was statistically significant with AUC = 0.656. The authors therefore concluded that osteoporosis is associated with increased vertebral bone marrow fat.

Furthermore, ^{1}H-MRS allows the calculation of the unsaturation level, previously defined using the formula [112, 113]:

$$unsaturation\ level = \frac{I_{(-CH=CH-)}}{I_{(-CH=CH-)} + I_{(-CH_2CH=CH-CH_2-)} + I_{(-(CH_2)n-)}}$$

where $I_{(–CH=CH–)}$, $I_{(–CH2–CH=CH–CH2–)}$, and $I_{(–(CH2)n–)}$ are the signal amplitudes of the olefinic (–CH=CH–) peak at 5.3 ppm, the α-olefinic ($–CH_2–CH=CH–CH_2–$) peak at 2.00 ppm, and the methylene ($–(CH_2)_n–$) peak at 1.30 ppm, respectively. Patsch et al. obtained ^{1}H-MRS-based unsaturation levels of L1–L3 in 69 postmenopausal women [113]. Thirty-six subjects (47.8 %) had spinal and/or peripheral osteoporotic fractures. After adjustment for age, race, and QCT-based lumbar BMD, the authors observed that the prevalence of osteoporotic fractures was associated with −1.7 % lower unsaturation levels. Interestingly, DXA-based BMD did not differ between fracture and non-fracture patients. These results suggest that altered bone marrow fat composition is associated with osteoporotic fractures. Thus, ^{1}H-MRS of vertebral bone marrow may serve as a tool for BMD independent fracture risk assessment. However, future studies are needed to investigate this association in further detail.

7 Summary

Osteoporosis is classified as a public health problem due to its increased risk for fragility fractures. Computer-assisted diagnostic tools for spine radiographs, DXA and MDCT images have been developed to support radiologists to correctly diagnose and report osteoporotic vertebral fractures. The assessment of fracture risk at the spine is based on the assessment of clinical risk factors and the measurements of BMD by using DXA or QCT. Standard QCT equivalent BMD values can be derived from routine contrast-enhanced MDCT examinations without further radiation exposure. MDCT-based trabecular bone microstructure parameters and FEMs have shown to improve the prediction of bone strength beyond DXA-based BMD and revealed pharmacotherapy effects, which were partly not captured by BMD. Furthermore, ^{1}H-MRS and chemical shift-based water-fat imaging techniques allow new insights into the vertebral bone marrow metabolism and its relationship with BMD, bone strength, and other body fat depots.

Conflict of Interest The authors state no conflict of interest.

References

1. NIH (2001) NIH consensus development panel on osteoporosis prevention, diagnosis, and therapy, March 7–29, 2000: highlights of the conference. South Med J 94(6):569–573
2. Hallberg I, Bachrach-Lindstrom M, Hammerby S, Toss G, Ek AC (2009) Health-related quality of life after vertebral or hip fracture: a seven-year follow-up study. BMC Musculoskelet Disord 10:135

3. Papaioannou A, Kennedy CC, Ioannidis G, Sawka A, Hopman WM, Pickard L, Brown JP, Josse RG, Kaiser S, Anastassiades T, Goltzman D, Papadimitropoulos M, Tenenhouse A, Prior JC, Olszynski WP, Adachi JD (2009) The impact of incident fractures on health-related quality of life: 5 years of data from the Canadian Multicentre Osteoporosis Study. Osteoporos Int 20(5):703–714
4. Ioannidis G, Papaioannou A, Hopman WM, Akhtar-Danesh N, Anastassiades T, Pickard L, Kennedy CC, Prior JC, Olszynski WP, Davison KS, Goltzman D, Thabane L, Gafni A, Papadimitropoulos EA, Brown JP, Josse RG, Hanley DA, Adachi JD (2009) Relation between fractures and mortality: results from the Canadian Multicentre Osteoporosis Study. CMAJ 181(5):265–271
5. Jalava T, Sarna S, Pylkkanen L, Mawer B, Kanis JA, Selby P, Davies M, Adams J, Francis RM, Robinson J, McCloskey E (2003) Association between vertebral fracture and increased mortality in osteoporotic patients. J Bone Miner Res 18(7):1254–1260
6. Bliuc D, Nguyen ND, Milch VE, Nguyen TV, Eisman JA, Center JR (2009) Mortality risk associated with low-trauma osteoporotic fracture and subsequent fracture in men and women. JAMA 301(5):513–521
7. Svedbom A, Hernlund E, Ivergard M, Compston J, Cooper C, Stenmark J, McCloskey EV, Jonsson B, Kanis JA (2013) Osteoporosis in the European Union: a compendium of country-specific reports. Arch Osteoporos 8(1–2):137
8. Cauley JA (2013) Public health impact of osteoporosis. J Gerontol A Biol Sci Med Sci 68 (10):1243–1251
9. Burge R, Dawson-Hughes B, Solomon DH, Wong JB, King A, Tosteson A (2007) Incidence and economic burden of osteoporosis-related fractures in the United States, 2005–2025. J Bone Miner Res 22(3):465–475
10. WHO Study Group (1994) Assessment of fracture risk and its application to screening for postmenopausal osteoporosis. Report of a WHO study group, World Health Organization technical report series, pp 843:1–129
11. Schuit SC, van der Klift M, Weel AE, de Laet CE, Burger H, Seeman E, Hofman A, Uitterlinden AG, van Leeuwen JP, Pols HA (2004) Fracture incidence and association with bone mineral density in elderly men and women: the Rotterdam study. Bone 34(1):195–202
12. Siris ES, Chen YT, Abbott TA, Barrett-Connor E, Miller PD, Wehren LE, Berger ML (2004) Bone mineral density thresholds for pharmacological intervention to prevent fractures. Arch Intern Med 164(10):1108–1112
13. Hopkins RB, Goeree R, Pullenayegum E, Adachi JD, Papaioannou A, Xie F, Thabane L (2011) The relative efficacy of nine osteoporosis medications for reducing the rate of fractures in post-menopausal women. BMC Musculoskelet Disord 12:209
14. Eriksen EF, Diez-Perez A, Boonen S (2013) Update on long-term treatment with bisphosphonates for postmenopausal osteoporosis: a systematic review. Bone 58C:126–135
15. Kanis JA, Oden A, Johansson H, Borgstrom F, Strom O, McCloskey E (2009) FRAX and its applications to clinical practice. Bone 44(5):734–743
16. Hillier TA, Cauley JA, Rizzo JH, Pedula KL, Ensrud KE, Bauer DC, Lui LY, Vesco KK, Black DM, Donaldson MG, Leblanc ES, Cummings SR (2011) WHO absolute fracture risk models (FRAX): do clinical risk factors improve fracture prediction in older women without osteoporosis? J Bone Miner Res 26(8):1774–1782
17. Delmas PD, Munoz F, Black DM, Cosman F, Boonen S, Watts NB, Kendler D, Eriksen EF, Mesenbrink PG, Eastell R (2009) Effects of yearly zoledronic acid 5 mg on bone turnover markers and relation of PINP with fracture reduction in postmenopausal women with osteoporosis. J Bone Miner Res 24(9):1544–1551
18. Lee J, Vasikaran S (2012) Current recommendations for laboratory testing and use of bone turnover markers in management of osteoporosis. Ann Lab Med 32(2):105–112
19. Lindsay R, Silverman SL, Cooper C, Hanley DA, Barton I, Broy SB, Licata A, Benhamou L, Geusens P, Flowers K, Stracke H, Seeman E (2001) Risk of new vertebral fracture in the year following a fracture. JAMA 285(3):320–323

20. Melton LJ III, Atkinson EJ, Cooper C, O'Fallon WM, Riggs BL (1999) Vertebral fractures predict subsequent fractures. Osteoporos Int 10(3):214–221
21. Roux C, Fechtenbaum J, Kolta S, Briot K, Girard M (2007) Mild prevalent and incident vertebral fractures are risk factors for new fractures. Osteoporos Int 18(12):1617–1624
22. van der Klift M, de Laet CE, McCloskey EV, Hofman A, Pols HA (2002) The incidence of vertebral fractures in men and women: the Rotterdam Study. J Bone Miner Res 17 (6):1051–1056
23. Oei L, Rivadeneira F, Ly F, Breda SJ, Zillikens MC, Hofman A, Uitterlinden AG, Krestin GP, Oei EH (2013) Review of radiological scoring methods of osteoporotic vertebral fractures for clinical and research settings. Eur Radiol 23(2):476–486
24. Eastell R, Cedel SL, Wahner HW, Riggs BL, Melton LJ III (1991) Classification of vertebral fractures. J Bone Miner Res 6(3):207–215
25. McCloskey EV, Spector TD, Eyres KS, Fern ED, O'Rourke N, Vasikaran S, Kanis JA (1993) The assessment of vertebral deformity: a method for use in population studies and clinical trials. Osteoporos Int 3(3):138–147
26. Sone T, Tomomitsu T, Miyake M, Takeda N, Fukunaga M (1997) Age-related changes in vertebral height ratios and vertebral fracture. Osteoporos Int 7(2):113–118
27. Genant HK, Wu CY, van KC, Nevitt MC (1993) Vertebral fracture assessment using a semiquantitative technique. J Bone Miner Res 8(9):1137–1148
28. Jiang G, Eastell R, Barrington NA, Ferrar L (2004) Comparison of methods for the visual identification of prevalent vertebral fracture in osteoporosis. Osteoporos Int 15(11):887–896
29. Damilakis J, Adams JE, Guglielmi G, Link TM (2010) Radiation exposure in X-ray-based imaging techniques used in osteoporosis. Eur Radiol 20(11):2707–2714
30. Vokes T, Bachman D, Baim S, Binkley N, Broy S, Ferrar L, Lewiecki EM, Richmond B, Schousboe J (2006) Vertebral fracture assessment: the 2005 ISCD official positions. J Clin Densitom 9(1):37–46
31. Lewiecki EM (2010) Bone densitometry and vertebral fracture assessment. Curr Osteoporos Rep 8(3):123–130
32. Schousboe JT, Debold CR (2006) Reliability and accuracy of vertebral fracture assessment with densitometry compared to radiography in clinical practice. Osteoporos Int 17 (2):281–289
33. Buehring B, Krueger D, Checovich M, Gemar D, Vallarta-Ast N, Genant HK, Binkley N (2010) Vertebral fracture assessment: impact of instrument and reader. Osteoporos Int 21 (3):487–494
34. Guglielmi G, Muscarella S, Bazzocchi A (2011) Integrated imaging approach to osteoporosis: state-of-the-art review and update. Radiographics 31(5):1343–1364
35. Bauer JS, Muller D, Ambekar A, Dobritz M, Matsuura M, Eckstein F, Rummeny EJ, Link TM (2006) Detection of osteoporotic vertebral fractures using multidetector CT. Osteoporos Int 17(4):608–615
36. Muller D, Bauer JS, Zeile M, Rummeny EJ, Link TM (2008) Significance of sagittal reformations in routine thoracic and abdominal multislice CT studies for detecting osteoporotic fractures and other spine abnormalities. Eur Radiol 18(8):1696–1702
37. Delmas PD, van de Langerijt L, Watts NB, Eastell R, Genant H, Grauer A, Cahall DL (2005) Underdiagnosis of vertebral fractures is a worldwide problem: the IMPACT study. J Bone Miner Res 20(4):557–563
38. Fechtenbaum J, Cropet C, Kolta S, Verdoncq B, Orcel P, Roux C (2005) Reporting of vertebral fractures on spine X-rays. Osteoporos Int 16(12):1823–1826
39. Gehlbach SH, Bigelow C, Heimisdottir M, May S, Walker M, Kirkwood JR (2000) Recognition of vertebral fracture in a clinical setting. Osteoporos Int 11(7):577–582
40. Williams AL, Al-Busaidi A, Sparrow PJ, Adams JE, Whitehouse RW (2009) Under-reporting of osteoporotic vertebral fractures on computed tomography. Eur J Radiol 69 (1):179–183

41. Gruber M, Dinges J, Muller D, Baum T, Rummeny EJ, Bauer J (2013) Impact of Specific Training in Detecting Osteoporotic Vertebral Fractures on Routine Chest Radiographs. Rofo 185(11):1074–1080
42. Roberts MG, Oh T, Pacheco EM, Mohankumar R, Cootes TF, Adams JE (2012) Semi-automatic determination of detailed vertebral shape from lumbar radiographs using active appearance models. Osteoporos Int 23(2):655–664
43. Baum T, Bauer JS, Netsch T, Klinder T, Dobritz M, Rummeny EJ, Noël PB, Lorenz C (2013) Automatic detection of osteoporotic vertebral fractures in routine thoracic and abdominal MDCT. Eur Radiol (epub ahead of print). doi:10.1007/s00330-013-3089-2
44. Roberts MG, Pacheco EM, Mohankumar R, Cootes TF, Adams JE (2010) Detection of vertebral fractures in DXA VFA images using statistical models of appearance and a semi-automatic segmentation. Osteoporos Int 21(12):2037–2046
45. Blake GM, Fogelman I (2010) An update on dual-energy x-ray absorptiometry. Semin Nucl Med 40(1):62–73
46. Adams JE (2009) Quantitative computed tomography. Eur J Radiol 71(3):415–424
47. Felsenberg D, Gowin W (1999) Bone densitometry by dual energy methods. Radiologe 39 (3):186–193
48. Adams JE (2013) Advances in bone imaging for osteoporosis. Nat Rev Endocrinol 9 (1):28–42
49. Pickhardt PJ, Pooler BD, Lauder T, del Rio AM, Bruce RJ, Binkley N (2013) Opportunistic screening for osteoporosis using abdominal computed tomography scans obtained for other indications. Ann Intern Med 158(8):588–595
50. Guise TA (2006) Bone loss and fracture risk associated with cancer therapy. Oncologist 11 (10):1121–1131
51. Bauer JS, Henning TD, Mueller D, Lu Y, Majumdar S, Link TM (2007) Volumetric quantitative CT of the spine and hip derived from contrast-enhanced MDCT: conversion factors. AJR Am J Roentgenol 188(5):1294–1301
52. Link TM, Koppers BB, Licht T, Bauer J, Lu Y, Rummeny EJ (2004) In vitro and in vivo spiral CT to determine bone mineral density: initial experience in patients at risk for osteoporosis. Radiology 231(3):805–811
53. Baum T, Muller D, Dobritz M, Rummeny EJ, Link TM, Bauer JS (2011) BMD measurements of the spine derived from sagittal reformations of contrast-enhanced MDCT without dedicated software. Eur J Radiol 80(2):e140–e145
54. Baum T, Muller D, Dobritz M, Wolf P, Rummeny EJ, Link TM, Bauer JS (2012) Converted lumbar BMD values derived from sagittal reformations of contrast-enhanced MDCT predict incidental osteoporotic vertebral fractures. Calcif Tissue Int 90(6):481–487
55. Acu K, Scheel M, Issever AS (2014) Time dependency of bone density estimation from computed tomography with intravenous contrast agent administration. Osteoporos Int 25 (2):535–542
56. Summers RM, Baecher N, Yao J, Liu J, Pickhardt PJ, Choi JR, Hill S (2011) Feasibility of simultaneous computed tomographic colonography and fully automated bone mineral densitometry in a single examination. J Comput Assist Tomogr 35(2):212–216
57. Pickhardt PJ, Lee LJ, del Rio AM, Lauder T, Bruce RJ, Summers RM, Pooler BD, Binkley N (2011) Simultaneous screening for osteoporosis at CT colonography: bone mineral density assessment using MDCT attenuation techniques compared with the DXA reference standard. J Bone Miner Res 26(9):2194–2203
58. Baum T, Karampinos DC, Liebl H, Rummeny EJ, Waldt S, Bauer JS (2013) High-resolution bone imaging for osteoporosis diagnostics and therapy monitoring using clinical MDCT and MRI. Curr Med Chem 20(38):4844–4852
59. Link TM (2012) Osteoporosis imaging: state of the art and advanced imaging. Radiology 263 (1):3–17
60. Burghardt AJ, Link TM, Majumdar S (2011) High-resolution computed tomography for clinical imaging of bone microarchitecture. Clin Orthop Relat Res 469(8):2179–2193

61. Krug R, Burghardt AJ, Majumdar S, Link TM (2010) High-resolution imaging techniques for the assessment of osteoporosis. Radiol Clin North Am 48(3):601–621
62. Baum T, Kutscher M, Muller D, Rath C, Eckstein F, Lochmuller EM, Rummeny EJ, Link TM, Bauer JS (2013) Cortical and trabecular bone structure analysis at the distal radius-prediction of biomechanical strength by DXA and MRI. J Bone Miner Metab 31(2):212–221
63. Ladinsky GA, Vasilic B, Popescu AM, Wald M, Zemel BS, Snyder PJ, Loh L, Song HK, Saha PK, Wright AC, Wehrli FW (2008) Trabecular structure quantified with the MRI-based virtual bone biopsy in postmenopausal women contributes to vertebral deformity burden independent of areal vertebral BMD. J Bone Miner Res 23(1):64–74
64. Krug R, Banerjee S, Han ET, Newitt DC, Link TM, Majumdar S (2005) Feasibility of in vivo structural analysis of high-resolution magnetic resonance images of the proximal femur. Osteoporos Int 16(11):1307–1314
65. Patsch JM, Burghardt AJ, Kazakia G, Majumdar S (2011) Noninvasive imaging of bone microarchitecture. Ann N Y Acad Sci 1240:77–87
66. Eckstein F, Lochmuller EM, Lill CA, Kuhn V, Schneider E, Delling G, Muller R (2002) Bone strength at clinically relevant sites displays substantial heterogeneity and is best predicted from site-specific bone densitometry. J Bone Miner Res 17(1):162–171
67. Ito M, Ikeda K, Nishiguchi M, Shindo H, Uetani M, Hosoi T, Orimo H (2005) Multi-detector row CT imaging of vertebral microstructure for evaluation of fracture risk. J Bone Miner Res 20(10):1828–1836
68. Bauer JS, Issever AS, Fischbeck M, Burghardt A, Eckstein F, Rummeny EJ, Majumdar S, Link TM (2004) Multislice-CT for structure analysis of trabecular bone—a comparison with micro-CT and biomechanical strength. Rofo 176(5):709–718
69. Baum T, Grabeldinger M, Rath C, Grande GE, Burgkart R, Patsch JM, Rummeny EJ, Link TM, Bauer JS (2014) Trabecular bone structure analysis of the spine using clinical MDCT: can it predict vertebral bone strength? J Bone Miner Metab 32(1):56–64
70. Bauer JS, Sidorenko I, Mueller D, Baum T, Issever AS, Eckstein F, Rummeny EJ, Link TM, Raeth CW (2014) Prediction of bone strength by muCT and MDCT-based finite-element-models: How much spatial resolution is needed? Eur J Radiol 83(1):e36–e42
71. Issever AS, Link TM, Kentenich M, Rogalla P, Burghardt AJ, Kazakia GJ, Majumdar S, Diederichs G (2010) Assessment of trabecular bone structure using MDCT: comparison of 64- and 320-slice CT using HR-pQCT as the reference standard. Eur Radiol 20(2):458–468
72. Klinder T, Wolz R, Lorenz C, Franz A, Ostermann J (2008) Spine segmentation using articulated shape models. Med Image Comput Comput Assist Interv 11(Pt 1):227–234
73. Klinder T, Ostermann J, Ehm M, Franz A, Kneser R, Lorenz C (2009) Automated model-based vertebra detection, identification, and segmentation in CT images. Med Image Anal 13 (3):471–482
74. Ma J, Lu L, Zhan Y, Zhou X, Salganicoff M, Krishnan A (2010) Hierarchical segmentation and identification of thoracic vertebra using learning-based edge detection and coarse-to-fine deformable model. Med Image Comput Comput Assist Interv 13(Pt 1):19–27
75. Mastmeyer A, Engelke K, Fuchs C, Kalender WA (2006) A hierarchical 3D segmentation method and the definition of vertebral body coordinate systems for QCT of the lumbar spine. Med Image Anal 10(4):560–577
76. Shen H, Litvin A, Alvino C (2008) Localized priors for the precise segmentation of individual vertebras from CT volume data. Med Image Comput Comput Assist Interv 11(Pt 1):367–375
77. Parfitt AM, Drezner MK, Glorieux FH, Kanis JA, Malluche H, Meunier PJ, Ott SM, Recker RR (1987) Bone histomorphometry: standardization of nomenclature, symbols, and units. Report of the ASBMR Histomorphometry Nomenclature Committee. J Bone Miner Res 2 (6):595–610
78. Räth C, Monetti R, Bauer J, Sidorenko I, Müller D, Matsuura M, Lochmüller EM, Zysset P, Eckstein F (2008) Strength through structure: visualization and local assessment of the trabecular bone structure. New J Phys 10(125010)

79. Sidorenko I, Monetti R, Bauer J, Mueller D, Rummeny E, Eckstein F, Matsuura M, Lochmueller EM, Zysset P, Raeth C (2011) Assessing methods for characterising local and global structural and biomechanical properties of the trabecular bone network. Curr Med Chem 18(22):3402–3409
80. Augat P, Schorlemmer S (2006) The role of cortical bone and its microstructure in bone strength. Age Ageing 35(Suppl 2):ii27–ii31
81. Carpenter RD (2013) Finite element analysis of the hip and spine based on quantitative computed tomography. Curr Osteoporos Rep 11(2):156–162
82. Dall'Ara E, Schmidt R, Pahr D, Varga P, Chevalier Y, Patsch J, Kainberger F, Zysset P (2010) A nonlinear finite element model validation study based on a novel experimental technique for inducing anterior wedge-shape fractures in human vertebral bodies in vitro. J Biomech 43(12):2374–2380
83. Dall'Ara E, Pahr D, Varga P, Kainberger F, Zysset P (2012) QCT-based finite element models predict human vertebral strength in vitro significantly better than simulated DEXA. Osteoporos Int 23(2):563–572
84. Imai K, Ohnishi I, Bessho M, Nakamura K (2006) Nonlinear finite element model predicts vertebral bone strength and fracture site. Spine (Phila Pa 1976) 31(16):1789–1794
85. Imai K, Ohnishi I, Yamamoto S, Nakamura K (2008) In vivo assessment of lumbar vertebral strength in elderly women using computed tomography-based nonlinear finite element model. Spine (Phila Pa 1976) 33(1):27–32
86. Wang X, Sanyal A, Cawthon PM, Palermo L, Jekir M, Christensen J, Ensrud KE, Cummings SR, Orwoll E, Black DM, Keaveny TM (2012) Prediction of new clinical vertebral fractures in elderly men using finite element analysis of CT scans. J Bone Miner Res 27(4):808–816
87. Graeff C, Timm W, Nickelsen TN, Farrerons J, Marin F, Barker C, Gluer CC (2007) Monitoring teriparatide-associated changes in vertebral microstructure by high-resolution CT in vivo: results from the EUROFORS study. J Bone Miner Res 22(9):1426–1433
88. Chevalier Y, Quek E, Borah B, Gross G, Stewart J, Lang T, Zysset P (2010) Biomechanical effects of teriparatide in women with osteoporosis treated previously with alendronate and risedronate: results from quantitative computed tomography-based finite element analysis of the vertebral body. Bone 46(1):41–48
89. Imai K, Ohnishi I, Matsumoto T, Yamamoto S, Nakamura K (2009) Assessment of vertebral fracture risk and therapeutic effects of alendronate in postmenopausal women using a quantitative computed tomography-based nonlinear finite element method. Osteoporos Int 20 (5):801–810
90. Keaveny TM, Donley DW, Hoffmann PF, Mitlak BH, Glass EV, San Martin JA (2007) Effects of teriparatide and alendronate on vertebral strength as assessed by finite element modeling of QCT scans in women with osteoporosis. J Bone Miner Res 22(1):149–157
91. Mulder L, van Rietbergen B, Noordhoek NJ, Ito K. Determination of vertebral and femoral trabecular morphology and stiffness using a flat-panel C-arm-based CT approach. Bone 50 (1):200–208
92. Duque G (2008) Bone and fat connection in aging bone. Curr Opin Rheumatol 20 (4):429–434
93. Bredella MA, Torriani M, Ghomi RH, Thomas BJ, Brick DJ, Gerweck AV, Rosen CJ, Klibanski A, Miller KK (2011) Vertebral bone marrow fat is positively associated with visceral fat and inversely associated with IGF-1 in obese women. Obesity (Silver Spring) 19 (1):49–53
94. Gilsanz V, Chalfant J, Mo AO, Lee DC, Dorey FJ, Mittelman SD (2009) Reciprocal relations of subcutaneous and visceral fat to bone structure and strength. J Clin Endocrinol Metab 94 (9):3387–3393
95. von Muhlen D, Safii S, Jassal SK, Svartberg J, Barrett-Connor E (2007) Associations between the metabolic syndrome and bone health in older men and women: the Rancho Bernardo Study. Osteoporos Int 18(10):1337–1344
96. Albala C, Yanez M, Devoto E, Sostin C, Zeballos L, Santos JL (1996) Obesity as a protective factor for postmenopausal osteoporosis. Int J Obes Relat Metab Disord 20(11):1027–1032

97. Fazeli PK, Horowitz MC, MacDougald OA, Scheller EL, Rodeheffer MS, Rosen CJ, Klibanski A (2013) Marrow fat and bone–new perspectives. J Clin Endocrinol Metab 98 (3):935–945
98. Sheu Y, Cauley JA (2011) The role of bone marrow and visceral fat on bone metabolism. Curr Osteoporos Rep 9(2):67–75
99. Ma J (2008) Dixon techniques for water and fat imaging. J Magn Reson Imaging 28 (3):543–558
100. Reeder SB, Pineda AR, Wen Z, Shimakawa A, Yu H, Brittain JH, Gold GE, Beaulieu CH, Pelc NJ (2005) Iterative decomposition of water and fat with echo asymmetry and least-squares estimation (IDEAL): application with fast spin-echo imaging. Magn Reson Med 54 (3):636–644
101. Shen W, Gong X, Weiss J, Jin Y (2013) Comparison among T1-weighted magnetic resonance imaging, modified dixon method, and magnetic resonance spectroscopy in measuring bone marrow fat. J Obes 2013:298675
102. Griffith JF, Yeung DK, Antonio GE, Lee FK, Hong AW, Wong SY, Lau EM, Leung PC (2005) Vertebral bone mineral density, marrow perfusion, and fat content in healthy men and men with osteoporosis: dynamic contrast-enhanced MR imaging and MR spectroscopy. Radiology 236(3):945–951
103. Griffith JF, Yeung DK, Antonio GE, Wong SY, Kwok TC, Woo J, Leung PC (2006) Vertebral marrow fat content and diffusion and perfusion indexes in women with varying bone density: MR evaluation. Radiology 241(3):831–838
104. Li X, Kuo D, Schafer AL, Porzig A, Link TM, Black D, Schwartz AV (2011) Quantification of vertebral bone marrow fat content using 3 Tesla MR spectroscopy: reproducibility, vertebral variation, and applications in osteoporosis. J Magn Reson Imaging 33(4):974–979
105. Shen W, Scherzer R, Gantz M, Chen J, Punyanitya M, Lewis CE, Grunfeld C (2012) Relationship between MRI-measured bone marrow adipose tissue and hip and spine bone mineral density in African-American and Caucasian participants: the CARDIA study. J Clin Endocrinol Metab 97(4):1337–1346
106. Shen W, Chen J, Punyanitya M, Shapses S, Heshka S, Heymsfield SB (2007) MRI-measured bone marrow adipose tissue is inversely related to DXA-measured bone mineral in Caucasian women. Osteoporos Int 18(5):641–647
107. Kuhn JP, Hernando D, Meffert PJ, Reeder S, Hosten N, Laqua R, Steveling A, Ender S, Schroder H, Pillich DT (2013) Proton-density fat fraction and simultaneous R2* estimation as an MRI tool for assessment of osteoporosis. Eur Radiol 23(12):3432–3439
108. Yu H, Shimakawa A, McKenzie CA, Brodsky E, Brittain JH, Reeder SB (2008) Multiecho water-fat separation and simultaneous R2* estimation with multifrequency fat spectrum modeling. Magn Reson Med 60(5):1122–1134
109. Liu CY, McKenzie CA, Yu H, Brittain JH, Reeder SB (2007) Fat quantification with IDEAL gradient echo imaging: correction of bias from T(1) and noise. Magn Reson Med 58 (2):354–364
110. Reeder SB, Hu HH, Sirlin CB (2012) Proton density fat-fraction: a standardized MR-based biomarker of tissue fat concentration. J Magn Reson Imaging 36(5):1011–1014
111. Karampinos DC, Melkus G, Baum T, Bauer JS, Rummeny EJ, Krug R (2013) Bone marrow fat quantification in the presence of trabecular bone: initial comparison between water-fat imaging and single-voxel MRS. Magn Reson Med (Epub ahead of print). doi:10.1002/mrm.24775
112. Baum T, Yap SP, Karampinos DC, Nardo L, Kuo D, Burghardt AJ, Masharani UB, Schwartz AV, Li X, Link TM (2012) Does vertebral bone marrow fat content correlate with abdominal adipose tissue, lumbar spine bone mineral density, and blood biomarkers in women with type 2 diabetes mellitus? J Magn Reson Imaging 35(1):117–124
113. Patsch JM, Li X, Baum T, Yap SP, Karampinos DC, Schwartz AV, Link TM (2013) Bone marrow fat composition as a novel imaging biomarker in postmenopausal women with prevalent fragility fractures. J Bone Miner Res 28(8):1721–1728

Part II
Image Processing

Computer Aided Detection of Bone Metastases in the Thoracolumbar Spine

Jianhua Yao, Joseph E. Burns and Ronald M. Summers

Abstract Computer-aided detection (CAD) techniques and algorithms for radiologic applications are rapidly growing in scope and sophistication. One important application of CAD techniques in medicine is in the detection and assessment of metastatic disease to the bone. Bone metastases affect approximately 400,000 patients per year in the United States. Early detection of bone metastases is important clinically, as the prognosis can change and the treatment regimen can at that point be altered from one of curative therapy to one of palliative treatment. Both lytic and sclerotic metastatic disease can act to biomechanically weaken the bone, and potentially lead to pathologic fractures. This chapter presents a framework for computer-aided detection of lytic and sclerotic metastatic lesions in the thoracolumbar spine using computed tomography (CT). State-of-art techniques are described in detail in each module of the framework. Thorough validation experiments are designed and results are presented. We also discuss the clinical significance and limitation of the CAD system.

1 Background of Bone Metastases

Bone metastases are a frequent occurrence of cancer, affecting approximately 400,000 patients per year in the United States [1]. In advanced breast and prostate cancer alone, bone metastases are seen in up to 70 % of patients, and in patients with carcinoma in the lung, colon, stomach, bladder, uterus, rectum, thyroid or kidney,

J. Yao (✉) · R.M. Summers
Imaging Biomarkers and Computer-Aided Diagnosis Laboratory,
Radiology and Imaging Sciences Department, Clinical Center,
National Institutes of Health, Bethesda, MD 20892-1182, USA
e-mail: jyao@cc.nih.gov

J.E. Burns
Department of Radiological Sciences, School of Medicine,
University of California, Irvine, CA 92868, USA
e-mail: jburns@uci.edu

© Springer International Publishing Switzerland 2015

S. Li and J. Yao (eds.), *Spinal Imaging and Image Analysis*,
Lecture Notes in Computational Vision and Biomechanics 18,
DOI 10.1007/978-3-319-12508-4_4

the occurrence is in the approximate range 15–30 % [2]. Morbidity associated with solid tumor metastasis to the skeleton is frequent and often debilitating, with secondary skeletal complications (skeletal related events or SRE's) occurring in approximately 50 % of patients [1]. Overall, the spine is the most frequent location for skeletal metastases [3]. Vertebral metastases may be associated with SREs including pain, pathologic compression fractures, spinal cord compression, and hematopoietic abnormalities [2–5]. Prior studies have demonstrated diminished functional independence associated with SREs and reduced patient survival after pathological fractures [6–8]. Patient outcomes are improved with early detection and treatment of vertebral metastases before the onset of significant morbidities [9–11].

Traditionally categorized as sclerotic or lytic, a metastatic bone lesion is now thought to exist as a continuum, and may in fact interconvert as part of its natural history [12, 13]. Patients may have lytic, sclerotic or mixed density metastatic lesions. The majority of patients with breast cancer have predominantly lytic lesions, although a subset of breast cancer patients of at least 15–20 % have predominantly sclerotic lesions. In contrast, the lesions in prostate cancer are predominantly sclerotic. Both lytic and sclerotic metastatic disease can act to biomechanically weaken bone and lead to pathologic fractures. Whereas lytic metastases frankly destroy the osseous matrix to weaken bone, in the case of sclerotic metastases, this weakening occurs via bone replacement with irregularly mineralized and disorganized matrix produced as woven bone [4, 14–16]. Fractures in pathologically weakened bone occur most commonly in the spine in the prototypical example of sclerotic metastases due to prostate cancer [4].

Metastases to the spine can involve the bone, epidural space, leptomeninges, and spinal cord. Various imaging modalities have been employed to diagnose different aspects of bone metastases. Radiographs are a ubiquitous modality in most hospitals and routinely used for screening in the evaluation of back or neck pain. However, up to 40 % of bone lesions will not be identified by plain film X-ray studies, thus presenting many false-negative results [17]. Bone scintigraphy, a physiologic imaging method, has also found use as a screening modality for skeletal metastases, but this technique suffers from obscuration of lesions by overprojecting structures, and is sensitive to the level of vascularization, and also has known false negative issues. Single-photon-emission computed tomography (SPECT), a cross-sectional imaging method, improves on the sensitivity of bone scintigraphy, but regions of radionuclide uptake are often nonspecific and result in many false positives. [^{18}F] fluoro-2-deoxy-d-glucose positron emission tomography (FDG-PET) can detect increased glucose metabolism of neoplastic cells in the bone marrow, making it a sensitive method for assessment of cortical and medullary metastases, but it suffers from relative low spatial resolution. Computed tomography (CT) scans typically have excellent spatial resolution with superb osseous delineation, and enable the direct detection of cortical destruction [17]. PET/CT hybrid scanners, where both PET and CT are acquired in the same session so that the two modalities are inherently registered and multi-spectrum features can be

obtained, have entered into widespread clinical use in recent years. PET/CT thus combines the high spatial resolution anatomic information of CT scanning with the physiologic/metabolic information of the tissue obtained from PET imaging to assess regions of pathology. A disadvantage of all X-ray and gamma ray imaging modalities is the associated risk of ionizing radiation exposure to the patient. Magnetic Resonance (MR) imaging provides a high level of soft tissue contrast, optimal for detection of early bone marrow deposits, and is useful for assessing for spinal cord injury and epidural tumors. Additionally, MRI does not involve ionizing radiation exposure to the patient. Disadvantages of MRI include the relatively long scanning time compared to CT which limits its usefulness for scanning unstable patients in emergent situations, and also makes it more susceptible to patient motion artifact… Moreover, MRI cannot be used for patients with medical appliances such as certain types of pacemakers and cerebral aneurysm coils, and it suffers image distortion in patients with metallic orthopedic hardware [18]. Although numerous and varying imaging modalities are currently available, and may be situationally targeted toward specific classes of neoplastic processes, CT remains a widely used high resolution modality for detection and surveillance of many types of cancer, particularly in a community hospital setting. CT is also the most cost effective 3D imaging modality. In the case of prostate cancer, the major role of CT for skeletal imaging is the detection and anatomic localization of bone metastases [19]. Examples of CT images of lytic and sclerotic bone metastases in the spine are shown in Fig. 1.

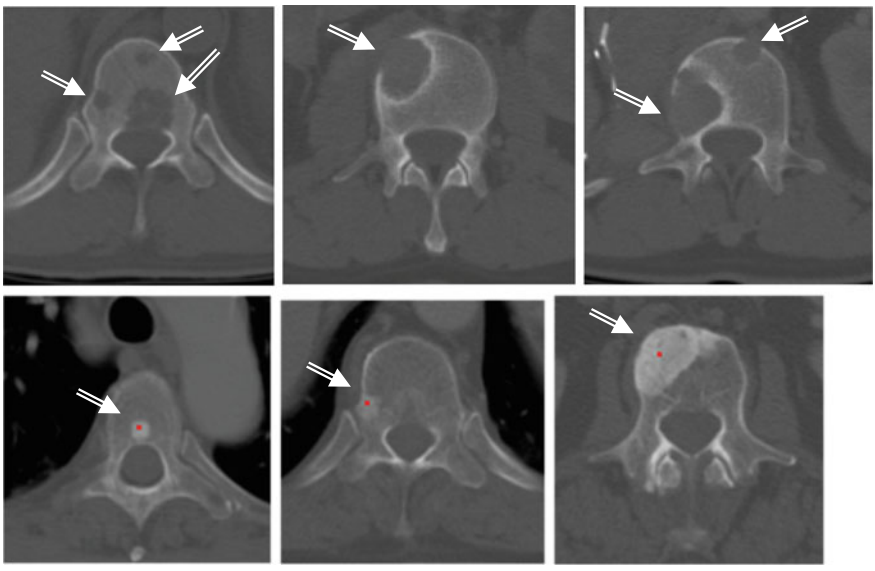


Fig. 1 Examples of lytic and sclerotic metastases in the spine. *First row* lytic metastases. *Second row* sclerotic metastases

2 Background of Computer Aided Detection

Detection of spinal metastases on CT can be a challenge, as the spinal column is a physically extended and complex structure. Lesion presentation may be subtle and unexpected. Adding to this complexity, in the current clinical practice environment, multiple high-throughput CT scanners can produce numerous patient studies in a short time interval, each with thousands of images, restricting time for image assessment. Numerous anatomic structures in each image must be assessed for pathology, typically at multiple window/level settings, effectively increasing the number of images to be reviewed geometrically.

Computer Aided Detection (CAD) researchers have been active in the past two decades and have showed very promising results. CAD is a potential solution to the challenge of rapid assessment of very large datasets. CAD is a software tool that can detect, mark, and quantitatively assess potential pathologies for further scrutiny by a radiologist. The CAD system should be topically focused with the ability to facilitate rapid and accurate assessment of important anatomy or high risk pathology subsets of interest in the CT data, such as the spine. While the final diagnosis is made by the radiologist responsible for interpreting the examination, CAD can help improve sensitivity in identifying lesions potentially overlooked by radiologists in situations of data overload or fatigue.

Figure 2 is a block diagram of a typical CAD system. A typical CAD system has two phases: training and testing. In the training phase, a set of training data is first

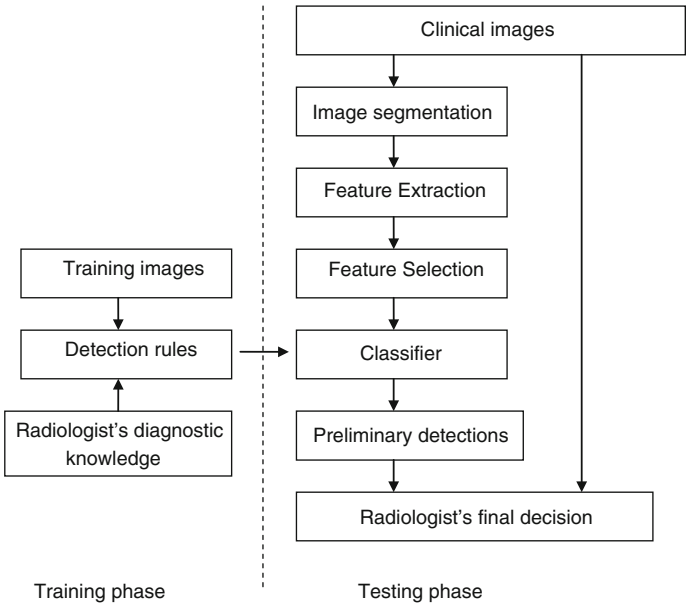


Fig. 2 Block diagram of computer aided detection

collected. Then radiologist's diagnostic knowledge is applied to analyze the training data and to derive detection rules (classifiers). In the testing phase, clinical images are taken as input. An image segmentation step is performed to extract structures of interest and limit the search region. Characteristic features of the structures are computed. Relevant features are then selected and input into a classifier. With the aid of the detection rules developed in the training phase, the classifier distinguishes true lesions from false lesions, or malign lesions from benign lesions. The preliminary detections are then reported to radiologists for final decisions.

Characteristic features for classification include features traditionally used by radiologists and high-order features that are not inherently intuitive. Features include shape features such as circularity, sphericity, compactness, irregularity, elongation, or density features such as contrast, roughness, and texture attributes. Different detection tasks need different sets of relevant features. Feature selection techniques, such as forward stepwise method and genetic algorithm, are applied in the training phase to choose features for further classification [20]. Several classifiers have been proposed for different applications, including linear discriminant analysis, Bayesian methods, artificial neural network, and support vector machine [21].

The quality of a CAD system can be characterized by the sensitivity and specificity of the system detection. Sensitivity refers to the fraction of diseased cases correctly identified as positive in the system (true positive fraction, TPF). Specificity refers to the fraction of disease-free cases correctly identified as negative. "Receiver operating characteristic" (ROC) curves are used to describe the relationship between sensitivity and specificity. The ROC curves show the true-positive fraction (TPF = sensitivity) versus the false-positive fraction (FPF = 1 − specificity). In addition to ROC curves, Free-response ROC (FROC) curves (TPF versus false positive per case) were proposed to more accurately represent the number of false positive detections [22]. The areas under the ROC and FROC curve are measures of the quality of a CAD system. There is often a tradeoff between achieving high specificity and high sensitivity. A successful CAD system should detect as many true lesions as possible while minimizing the false positive detection rate.

CAD systems have been used to detect lesions in the breast, where the increased the true positive rate in breast cancer screening and improved the yield of biopsy recommendations for patients with masses on serial mammograms [23, 24]. CAD has been shown to improve radiologists' performance detecting lung nodules on chest radiography and CT [25, 26] and to increase sensitivity for detecting polyps on CT colonography [27]. In spine imaging, CAD systems had been developed to detect spine abnormality and disease such as lytic lesions [28], sclerotic lesion [29], fractures [30], degenerative disease [31], syndesmophyte [32] and epidural masses [33]. Several commercial systems in mammography, chest CT and CT colonography have already received FDA approval for clinical use.

CAD involves all aspects of medical image processing techniques. For instance, image segmentation and registration are necessary for feature computation, and image visualization and measurement are essential to present the results to clinicians.

3 Spine Metastasis CAD System Overview

Bone metastases can appear lytic, sclerotic, or anywhere in a continuum between these extremes. Therefore, automated detection of these lesions is a complex problem that must be broken down into manageable components. We develop a CAD framework that can handle both types of bone metastases. The CAD system is a supervised machine learning framework. By supplying the CAD system with different sets of labeled training data for lytic and sclerotic bone metastases, we can train the system to detect the two types of bone metastases separately. The framework has two phases: training phase and testing phase. Training phase has four stages: spine segmentation, candidate detection, feature extraction and classifier training. The testing stage also has four stages: spine segmentation, candidate detection, feature extraction and classification. The first three stages of training and testing phases are identical. The training phase is an offline process and can be fine-tuned by researchers, and the testing phase is fully automatic. Each stage will be elaborated upon in the following sections.

4 Spinal Column Segmentation and Partitioning

Spine is a bony structure with higher CT value (pixel intensity) than other tissue types. The CT value is a gauge of X-ray beam attenuation measured by the CT scanner, and historically was normalized into a standardized scaling set referred to as Hounsfield units (HU). We first apply a threshold of 200 HU to mask out the bone pixels. Then a connected component analysis is conducted on the bone mask and the largest connected component in the center of the image is retained as the initial spine segmentation. The bounding box of the initial segmentation is used as the search region for the following segmentation tasks.

The spinal canal links all vertebrae into a column. On a transverse cross section, the spinal canal appears as a low intensity oval region surrounded by high density pedicle and lamina (Fig. 3a). The extraction of the spinal canal is essential in order to accurately localize the spine and form the spinal column. We apply a watershed algorithm to detect the potential spinal canal regions, and then conduct a graph search to locate and extract the spinal canal.

The principle of the watershed algorithm [34] is to transform the gradient of a gray level image into a topographic surface. The algorithm simulates the watershed scenario by puncturing holes at the local minimum of the intensity and filling the region with water. Each region filling with water is called a catchment basin. The spinal canal resembles a catchment basin on a 2D cross sectional image. We adopted the watershed algorithm implementation in ITK [35].

The well-known over-segmentation problem of the watershed algorithm is alleviated by merging adjacent basins. Depth of a basin is defined in Eq. 1.

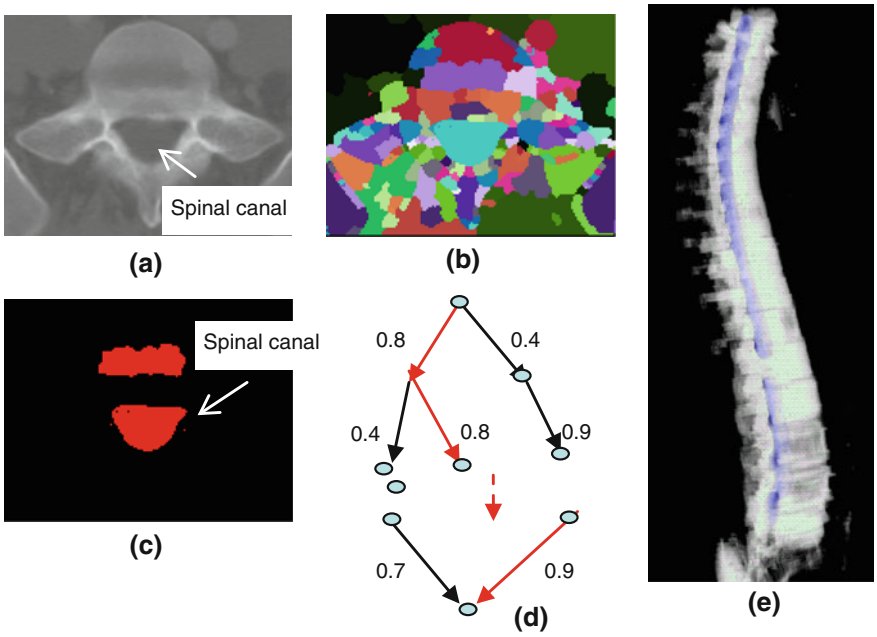


Fig. 3 Spinal canal extraction. **a** CT image; **b** watershed result; **c** spinal canal candidates; **d** directed acyclic graph (*DAG*), number on the edge is the weight between two nodes; and **e** extracted spinal canal (*blue color*)

$$d(b) = average(I(x)), \quad x \in b \tag{1}$$

here $I(x)$ is the image intensity of a pixel x inside the basin b. Given a and b are two neighboring basins, they will be merged if both conditions in Eq. 2 are satisfied

$$\begin{aligned} |d(b) - d(a)| &< d(c_i) - d(a) + \delta_m, \quad \forall c_i \in N(a), c_i \neq b \\ |d(b) - d(a)| &< d(c_i) - d(b) + \delta_m, \quad \forall c_i \in N(b), c_i \neq a \end{aligned} \tag{2}$$

here $N(a)$ denotes neighbors of basin a, δ_m is the merging threshold. After that, all basins that meet the criteria in Eq. 3 and surrounded by bone pixels are recorded as potential candidates for the spinal canal.

$$d(c_i) - d(b) > \delta_d, \quad \forall c_i \in N(b) \tag{3}$$

here δ_d is the depth contrast threshold. Figure 3b, c show the result of the watershed algorithm and the candidates for spinal canals.

As showed in Fig. 3c, multiple canal candidates may exist in one slice due to the partial volume effect or loss density vertebra body region such as lytic bone lesion. We propose a method to extract the correct spinal canal using directed graph search. We first build a directed acyclic graph (DAG) from the canal candidates. The DAG

is illustrated in Fig. 3d. The graph $G(N, A)$ is a structure that consists of a set of nodes N and a set of directional edges E. A node is one canal candidate. A directional edge $\langle n_1, n_2 \rangle$ connects two nodes n_1 and n_2 on adjacent slices, where the weight of $\langle n_1, n_2 \rangle$ is computed as the overlap of n_1 and n_2, as in Eq. 4.

$$\begin{aligned} weight(\langle n_1, n_2 \rangle) &= \frac{|\cap(n_1, n_2)|}{|\cup(n_1, n_2)|} \\ \cap(n_1, n_2) &= \{\langle x_i, y_i \rangle\}, \langle x_i, y_i \rangle \in n_1 \text{ and } \langle x_i, y_i \rangle \in n_2 \\ \cup(n_1, n_2) &= \{\langle x_i, y_i \rangle\}, \langle x_i, y_i \rangle \in n_1 \text{ or } \langle x_i, y_i \rangle \in n_2 \end{aligned} \tag{4}$$

An edge only exists when its weight is greater than 0 (two nodes overlap). DAG has sources on the first slice and sinks on the last slice. A directed graph searching algorithm [36] is applied to find the longest path from source to sink, which is the spinal canal in our case. In Fig. 3d, the longest path is marked with red color. The centerline of the spinal canal is then computed and smoothed using a Bernstein spline [37]. Figure 3e shows the extracted spinal canal (blue color).

Vertebra segmentation is commonly implemented using geometric and statistical models owing to its articulated and complex structure (Fig. 4a) [38–40]. The anatomical models capture the shape, topology and inter-relationship of vertebrae, and therefore convert the image segmentation problem to a model fitting problem. We proposed a four-part vertebral model to segment the vertebral region on a 2D slice. The model includes four main vertebra sub-structures: vertebral body, posterior spinous process, left transverse process and right transverse process (see Fig. 4b). The vertebral body is modeled as a circle with a medial atom in the center and border atoms evenly distributed on the border. The spinous and transverse processes are modeled as slabs with a medial axis and a set of border atoms on each side. The model's multiple-part structure simplifies the problem and makes the segmentation robust. Each model part is essentially a medial model [41]. The medial axis defines the skeleton, and the border atoms define the boundary.

The border of the medial model can be written as an implicit function in the local coordinate of the model

$$v = f(u) \tag{5}$$

A border atom A_i can then be represented in the local coordinate, as $A_i = (u_i, v_i) = (u_i, f(u_i))$. In the coordinate system of the disk model, u is the radian angle around the center and v is the distance to the center. In the coordinate system of the slab model, u is the distance along the medial axis and v is the distance to the medial axis.

The segmentation task is to locate the border atoms so that a maximum model-to-image match can be reached. The matching metric should also preserve the model topology and border smoothness constraint. We design a metric for the model matching:

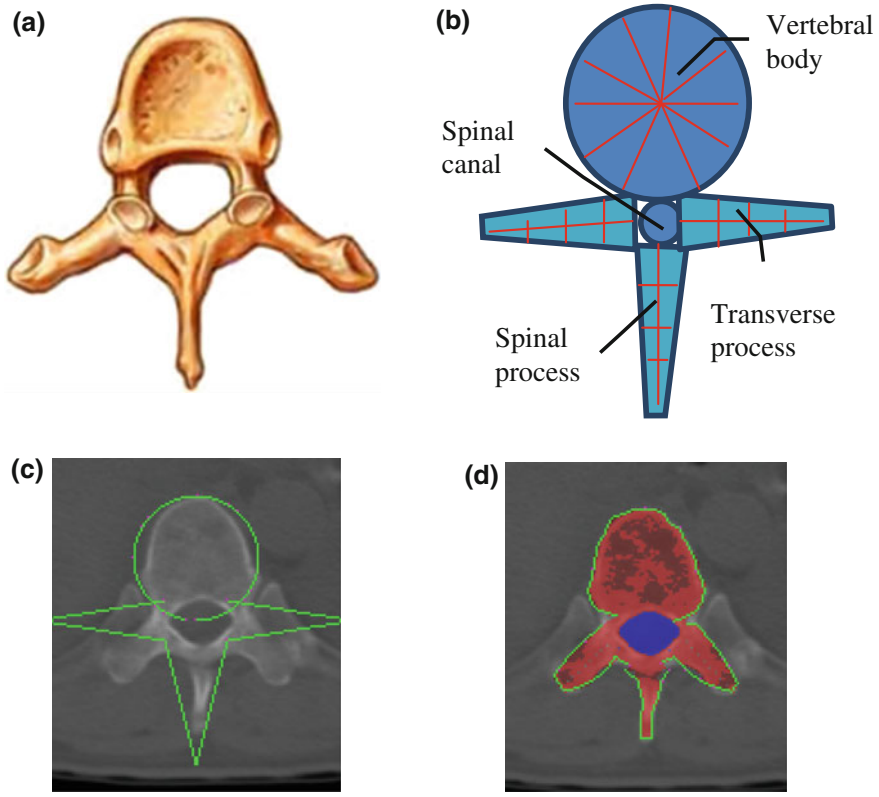


Fig. 4 Spine segmentation. **a** Vertebra anatomical illustration; **b** vertebra template; **c** Initial template superimposed on CT; and **d** segmentation results

$$E = \sum_{i=1}^{n} \left(w_g \vec{g}(u_i, f(u_i)) - w_1 \|\nabla f(u_i)\| - w_2 \left\|\nabla^2 f(u_i)\right\| - w_p p(u_i, f(u_i)) \right) \tag{6}$$

$$\vec{g}(u_i, f(u_i)) \text{ is directional gradient}$$

$$p(u_i, f(u_i)) = \begin{cases} 1 & (u_i, f(u_i)) \text{ is occupied} \\ 0 & \text{otherwise} \end{cases}$$

Here $(u_i, f(u_i))$ are the border atoms. The metric has four components, the directional gradient $\vec{g}(u_i, f(u_i))$ is to match the border atoms with the intensity edge of the image, $\nabla f(u_i)$ and $\nabla^2 f(u_i)$ are smoothness constraints on the border, and $p(u_i, f(u_i))$ is a penalty function to prevent intersecting between model parts. Weights w_g, w_1, w_2 and w_p are set empirically.

The extracted spinal canal defines the initial location and size of the vertebra model. The model matching proceeds sequentially. First the vertebral body is matched, followed by the spinous process, and at the end the transverse processes.

The results of the previous steps are used to determine the initial location and size of the parts in the following steps. In our current model, we define 36 border atoms for the disk model and 20 atoms for the slab models.

5 Spinal Column Partitioning

The spinal column consists of a set of vertebrae separated by inter-vertebral discs (Fig. 5a). Since the spinal column is a curved structure, the standard planar reformations (sagittal and coronal) do not provide clear views of the vertebral separation (Fig. 5b). Curved planar reformation (CPR) (Fig. 5c) [42] is generally considered superior.

After the spinal column is segmented, we need to partition the spinal column into vertebrae at the inter-vertebral disc locations so that we can process the vertebrae separately and also localize the abnormality at the vertebra level. We developed a partitioning approach based on curved reformation along the spinal canal.

The centerline of the spinal canal is used as the central axis for the CPR. We generate the CPR in sagittal and coronal directions. Given that the vertices on the

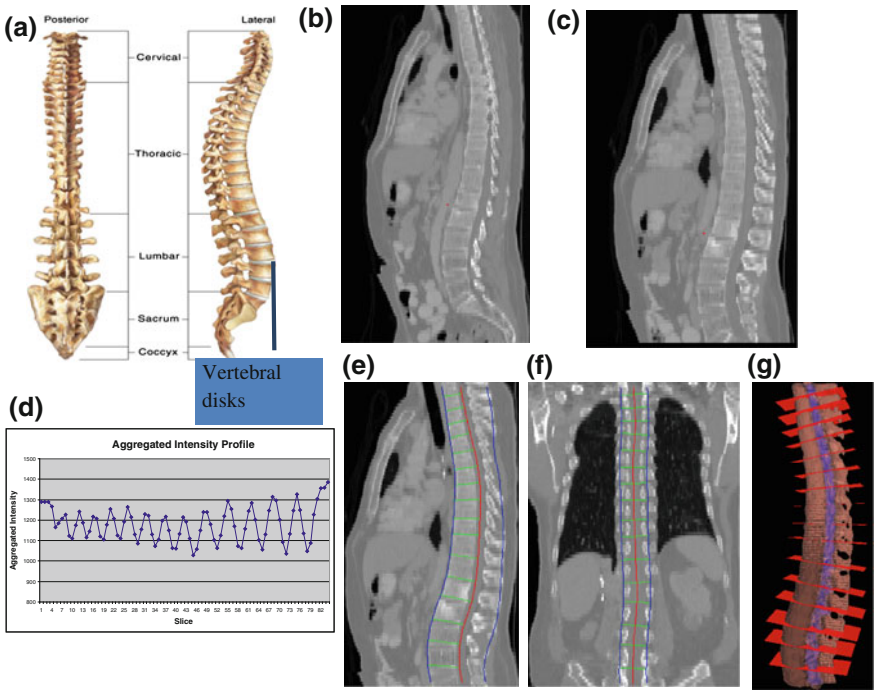


Fig. 5 Spine partitioning. **a** Spinal column; **b** regular sagittal reformation; **c** curved planar reformation in sagittal direction; **d** aggregated intensity profile (*AIP*) along the spinal canal; **e** spinal partitioning in sagittal direction; **f** spine partition in coronal direction; and **g** spine partition in 3D, *red* planes are the partitioning planes

centerline are $(x_j^c, y_j^c, z_j^c), j = 1 \ldots n$, the curved reformation in the sagittal direction is written as

$$I_{Sag}(x_i, y_j) = I_{3D}(x_j^c, x_i, z_j^c) \tag{7}$$

where I_{Sag} is the curved reformatted sagittal image, I_{3D} is the original 3D image. $(x_i,\ y_i)$ is the 2D coordinate in the reformatted image. Similarly, the curved reformation in the coronal direction is written as

$$I_{Cor}(x_i, y_j) = I_{3D}(x_i, y_j^c, z_j^c) \tag{8}$$

where I_{Cor} is the reformatted coronal image. Figure 5b, c show the regular coronal reformation, together with the curved planar reformation in sagittal and coronal directions. The curved reformations clearly better reveal the inter-vertebral disks.

To make use of the CPR for spinal column partitioning, the centerline of the spinal canal is first projected onto the reformatted images. Then the normal is computed at every point on the centerline. The intensity along the normal direction is then aggregated and recorded. Figure 5d shows the aggregated intensity profile (AIP) along the spinal cord at the reformatted coronal view. As observed, the aggregated intensity at the disc location is lower than those at the vertebral body location. However, the difference is still not prominent, especially at cervical spine and highly curved region. We further convolve the aggregated intensity profile with an adaptive disk function, which can be written as,

$$f(x) = \begin{cases} -1 & x \in [-T/2, T/2] \\ 1 & Elsewhere \end{cases} \tag{9}$$

The function is a rectangle function with adaptive width T. In order to determine T, we search the neighborhood in both directions on the AIP for local maximum values.

The intervertebral disks are then located at the lowest response points on the adjusted intensity profile and used to partition the spinal column. Figure 5e, f show the spine partition superimposed on reformatted CPR views and the spinal column with vertebral partitions in a 3D view is shown in Fig. 5g.

6 Metastasis Candidate Detection

After the spine is segmented, the following lesion detection processes are restricted to the segmented spine excluding the spinal canal. We locate bone metastasis candidates in three steps. First a watershed algorithm is applied to extract initial super-pixels as metastasis candidates, followed by a merging routine based on graph cut to alleviate over-segmentation. The resulting 2-D candidates are then merged into 3-D detections. For each 3-D candidate, a set of features is computed

and passed through a detection filter. The candidates which successfully pass through the detection filter are then sent to the next stage for classification.

The watershed algorithm is again applied to detect the metastasis candidates. Watershed algorithm views the gradient of the image intensity as a topographic surface in order to extract relatively homogeneous regions of the image called catchment basins, some of which will be candidates for lesions. The algorithm can be adapted for both lytic and sclerotic lesions. For lytic lesions, low intensity regions surrounded by high intensity regions are detected. Similarly, for sclerotic lesions, high intensity regions surrounded by low intensity regions are detected. Example results of the watershed algorithm are shown in Fig. 6.

We then address the over-segmentation problem in watershed with a post-watershed merging routine using a graph-cuts strategy [43]. Without loss of generality, we use the sclerotic lesion detection to describe the graph-cut strategy. We first initialize each watershed region with a foreground (F) or background (B) label. There are two types of foreground regions: those in the cortical bone region and those in the medullary regions. Any region that has intensity 100 HU higher than its surrounding regions (cortical or medullary) will be initialized as F. The rest of the regions are initialized as B. The regions and their neighbors are fed into a graph-cuts merging routine.

An adjacency graph for watershed regions is constructed by representing adjacent regions as nodes connected by edges [44]. The technique partitions the set of nodes into two disjoint sets F and B in a manner that minimizes an energy function,

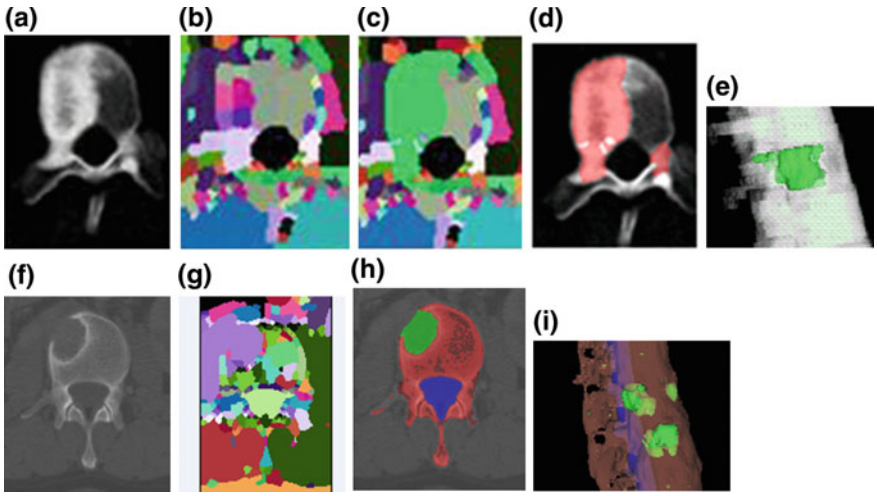


Fig. 6 Candidate detection and segmentation. **a** CT image of a vertebra with sclerotic lesions; **b** watershed result; **c** graph cut result after watershed; **d** candidate sclerotic lesions; **e** 3D segmented sclerotic lesions; **f** CT image of a vertebra with lytic lesions; **g** watershed result; **h** candidate lytic lesions; and **i** 3D segmented lytic lesions

$$E(L) = \sum_{\{p,q\} \in N} V_{Lp,Lq}(p,q) + \sum_{p \in P} D_{Lp}(p) \tag{10}$$

where P is the set of watershed regions, N is the set of pairs of adjacent regions, L is a labeling of all the regions where a given region p can have the label $L_p = F$ or $L_p = B$, V is a smoothness term that penalizes regions with similar densities having different labels, and D is a data term that penalizes a region with low density marked as foreground, or a region with high density marked as background. Thus the technique will merge higher-density regions into the foreground, and lower-density regions into the background. In this case,

$$\begin{aligned} D_B(p) &= K_B sign(I(p) - m_B)(I(p) - m_B)^2 \\ D_F(p) &= K_F sign(m_F - I(p))(I(p) - m_F)^2, \end{aligned} \tag{11}$$

where $K_F = 100$, $K_B = 1$ in our setting, $I(p)$ is the mean intensity of region p, and m_B and m_F are the means of the background and foreground respectively. As for the smoothness term, we chose

$$V_{Lp,Lq}(p,q) = K_s e^{-H^2/2\delta_s^2} \tag{12}$$

where $K_s = (\delta_F + \delta_B)/2$, δ_F and δ_B are the standard deviation of the foreground and background respectively. $H = \sum_r |H_p^r - H_q^r|$ and $H_p^r = \sum_{t \le r} h_p^t$ are the cumulative histogram of region p, and $\delta_s = 10{,}000$.

The smoothness and data penalty functions provide edge weights $w(i, j)$ for a graph G consisting of the adjacency graph of the watershed regions and two additional nodes f and b which both have edges connecting them to every region node:

$$w(f, q) = D_F(q);\ w(p, b) = D_B(p);\ w(p,q) = V_{Lp,Lq}(p, q) \tag{13}$$

A graph cut $[F, B]$ is a partition of the set of nodes such that $f \in F$ and $b \in B$, and the value of the cut is,

$$c(F,B) = \sum_{i \in F, j \in B} w(i,j) \tag{14}$$

A minimal graph cut of G is equivalent to a labeling that minimizes Eq. (10). Such a cut is computed according to a max-flow algorithm referenced in which generates a local minimum within a known factor of the global minimum. The resulting partition $[F, B]$ yields an optimized way of merging watershed regions in which regions corresponding to nodes in F and B are labeled as F and B respectively. Figure 6c demonstrates the effect of this merger. Each merged F region is then regarded as one potential detection.

So far, the candidates are all two-dimensional. Lesions actually extend through the spine in three dimensions. Therefore the next step is to merge the two-dimensional candidates that belong to the same lesion into a single three-dimensional "blob". Two candidates A and B are merged if and only if

1. They lie in adjacent slices z and $z + 1$
2. For all 2-D candidates C on $slice(z + 1)$ and all D on $slice(z)$, it is true that

$$(pr_A(B)/a(A) + pr_B(A)/a(B))/2 > (pr_A(C)/a(A) + pr_C(A)/a(C))/2$$
$$\text{and } (pr_A(B)/a(A) + pr_B(A)/a(B))/2 > (pr_D(B)/a(D) + pr_B(D)/a(B))/2,$$

where $slice(z)$ is the CT slice at height z, $pr_X(Y)$ is the fraction of candidate Y that overlaps with X when projected into the slice of X, and $a(X)$ is the area of candidate X. In other words, the average projectional overlap of A and B is greater than the average projectional overlap of A with any other candidate in the same slice as B, and also greater than the average projectional overlap of B with any candidate in the same slice as A.

After the lesions are detected in 3D, a level set algorithm is applied to obtain the 3D segmentation so that characteristic features can be derived. Level sets are evolving interfaces (contours or surfaces) that can expand, contract, and even split or merge. Level set methods are part of the family of segmentation algorithms that rely on the propagation of an approximate initial boundary under the influence of images forces [45]. The underlying idea behind the level set method is to embed the moving interfaces as the zero level set of a higher dimensional function $\phi(x, t)$, defined as

$$\phi(x, t) = \pm d \tag{15}$$

where $\pm d$ is the signed distance to the interface from point x. That is, x is outside the interface when $\phi(x, t) > 0$, inside the interface when $\phi(x, t) < 0$, and on the interface when $\phi(x, t) = 0$. The evolution of $\phi(x, t)$ can be represented by a partial differential equation:

$$\frac{\partial \phi}{\partial t} + \nabla\phi \cdot x'(t) = 0 \tag{16}$$

Define the scalar speed field F as $F = n \cdot x'(t)$, where $n = \nabla\phi$ is the normal direction, and then the above equation becomes the level set equation:

$$\frac{\partial \phi}{\partial t} + F|\nabla\phi| = 0 \tag{17}$$

usually, the speed function F can be written as an explicit level set scheme:

$$F = F_{prop} + F_{curv} + F_{adv} \tag{18}$$

where F_{prop} is the propagation expansion speed, F_{curv} is the speed on the curvature κ, and F_{adv} is the advection speed. Combining Eqs. 17 and 18, the final equation for level set segmentation can be written as:

$$\frac{\partial \phi}{\partial t} = \alpha \vec{g}(x)\delta|\nabla\phi| + \beta\vec{g}(x)\kappa(x)|\nabla\phi| + \gamma\nabla\vec{g}(x)\nabla\phi \tag{19}$$

here α, β, γ are weighting parameters for each term, δ is the step size, κ is the curvature, and $\vec{g}$ is the speed function. We used the ITK implementation of the level set algorithm in our system [35].

In the fast marching level set, a Gaussian gradient convolution is first applied to the image as the speed function. Then a sigmoid function is applied to remap the speed image. The sigmoid function is designed so that the propagation speed of the front is low when it is close to high image gradients and moves rather fast in the low gradient areas. The sigmoid function can be written as

$$S(I) = (\text{Max} - \text{Min}) \cdot \frac{1}{\left(1 + e^{-\left(\frac{I-b}{a}\right)}\right)} + \text{Min} \tag{20}$$

where I is the intensity of the input pixel, Min and Max are the range for output, a defines the width of the Sigmoid, and b defines the center. a and b control the shape of sigmoid function and the function of the speed image. The determination of a and b is based on the pixel statistics in the region [35]. The speed image for fast marching level set can be written as

$$\vec{g}_f(x) = S(G(I(x))) \tag{21}$$

where $I(x)$ is the image intensity, $G(.)$ is the Gaussian gradient operator, and $S(.)$ is the sigmoid function.

The Laplacian level set defines the speed term based on second derivative features in the image. The speed term is calculated as the Laplacian of the image values. The goal is to attract the evolving level set surface to local zero-crossings in the Laplacian image. In our implementation, the image is first convolved with a few iterations of gradient anisotropic diffusion. Gradient anisotropic diffusion has the attribute to reduce the noise and texture and meanwhile preserve the edge. After the anisotropic diffusion and the Laplacian filter, the speed image for Laplacian level set is

$$\vec{g}_L(x) = \nabla^2(A(I(x))) \tag{22}$$

where I is the image intensity, $A(.)$ is the anisotropic diffusion and ∇^2 is the Laplacian operator.

Segmentation results from the watershed and graph cut algorithms are used as the initialization for the level set algorithm. After the level set algorithm has run, a smooth 3D surface is computed for each detection. The level set results for sclerotic and lytic metastases are shown in Fig. 6e, i respectively.

7 Feature Extraction and Filtering

After the detections are segmented, quantitative features are computed to characterize the detections and distinguish true lesions from false findings. Based on our observation and knowledge about bone metastases, we devised a set of 28 quantitative features in three categories: location, shape and density. Table 1 lists all the features.

The shape features are based on the spatial moment of the detection. The density features are derived from the statistical moments of the intensity histogram of the segmented region. The relative coordinates to the center of the spinal canal is used to derive the location features. Following is the description of each feature:

1. surfaceArea: area of the 3D surface
2. volume: volume enclosed by the surface
3. primaryAxisLength: length of longest axis of the 3D bounding box
4. secondaryAxisLength: length of second longest axis of the 3D bounding box
5. aspectRatio10: ratio between longest and second longest axes
6. aspectRatio20: ratio between longest and third longest axes
7. aspectRatio21: ratio between second longest and third longest axes
8. sphericity: $\Psi = \frac{\pi^{\frac{1}{3}}(6V)^{\frac{2}{3}}}{A}$, here V is the volume, A is the area.
9. shapeComplexity_f1: shape complexity based on radial distance measures [46]
10. shapeComplexity_f2: shape complexity based on radial distance measures
11. shapeComplexity_f21: shape complexity based on radial distance measures
12. meanIntensity: mean intensity inside the detection
13. stdevIntensity: standard deviation of intensity inside the detection
14. skewnessIntensity: skewness of intensity inside the detection
15. kurtosisIntensity: kurtosis of intensity inside the detection
16. interiorIntensity: mean intensity of interior (not including the border)
17. borderIntensity: mean intensity at the border
18. outsideIntensity: mean intensity of region outside the detection

Table 1 Quantitative features for bone metastasis

Shape	Density	Location
surfaceArea	meanIntensity	distToBoundary
volume	stdevIntensity	relCoordx
primaryAxisLength	skewnessIntensity	relCoordy
secondaryAxisLength	kurtosisIntensity	onPedicle
aspectRatio10	interiorIntensity	outerBorderRatio
aspectRatio20	borderIntensity	corticalBorderRatio
aspectRatio21	outsideIntensity	cordBorderRatio
sphericity	outsideIntensityDev	
shapeComplexity_f1	innerOuterContrast	
shapeComplexity_f2	neighborIntensity	
shapeComplexity_f21		

19. outsideIntensityDev: standard deviation of intensity outside the detection
20. innerOuterContrast: difference between mean Intensity and outside Intensity
21. neighborIntensity: intensity of neighboring blobs from watershed algorithm.
22. distToBoundary: minimum distance from center of the detection to the border of vertebra
23. relCoordx: relative x coordinate to the center of spinal canal
24. relCoordy: relative y coordinate to the center of spinal canal
25. onPedicle: whether the detection is inside pedicle region
26. outerBorderRatio: ratio of total borders that are boundary of vertebra
27. corticalBorderRatio: ratio of total borders that are cortical shell of vertebra
28. cordBorderRatio: ratio of total borders that are spinal canal.

Feature filters are then applied to reduce the number of detections and relieve the burden of the classifier in the next step. The filters are designed based on observation of typical bone metastases so that all true detections remain. The set of filters currently in use are,

- Shape filter: aspectRatio10 < 3.5, to eliminate elongated detections
- Size filter: surfaceArea > 0.05 cm^2, to eliminate small detections
- Density filter: innerOuterContrast > 150, to eliminate less prominent detections.

These filters are loosely set and can reduce more than 50 % of false detections without impairing sensitivity.

8 Machine Learning and Classification

Machine learning techniques transfer expert's knowledge into computer algorithms. In a CAD system, a classifier is a mathematical model that determines whether a detection is a true or false finding. The classifier in a CAD system is usually a supervised learning system, i.e., the classifier is trained using annotated data by experts. Well-known classifiers such as neural networks (NN) [47] and support vector machines (SVM) [48] have been widely used in CAD systems [49].

SVM is a relatively new technique for data classification. It uses hyperplanes in a high dimensional feature space to separate data into different classes. SVM is trained with a learning system derived from statistical learning theory, and is generalizable to unknown data. In the training phase, detections are given a class label (lesion, non-lesion) to form feature-class pairs (x, y). Given a training set of S detections (x_1, y_1), (x_2, y_2), …, (x_s, y_s), for p-dimensional feature space $x_i \in \Re^p$ and $y_i \in \{+1, -1\}$, a hyperplane can be optimized to separate the two groups of data (true and false).

$$f(x) = w^T \phi(x) + b = 0 \tag{23}$$

here w and b are separating plane parameters, and $\phi(x)$ is a function to map vector x into a higher dimensional space. $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$ is called the kernel function. We are using radial basis functions as the kernel function, i.e.

$$K(x_i, x_j) = \exp\left(-\left\|x_i - x_j\right\|^2\right) \tag{24}$$

To separate two training classes, SVM is employed to solve the following optimization problem:

$$\begin{aligned} &\min_{w,b,\zeta}\left(\frac{1}{2}w^T w + C\sum_{i=1}^{N}\xi_i\right) \\ &\text{subject to} \quad y_i(w^T\phi(x_i) + b) \geq 1 - \xi_i, \;\; \xi_i \geq 0 \end{aligned} \tag{25}$$

here C is the penalty parameter. The mechanism of SVM is illustrated in Fig. 7, where a hyperplane is fit to separate two groups of dots. SVM allows a soft margin on each side of the hyperplane. For each data point, the distance to the margin of hyperplane is computed. If the point is on the correct side of the plane, the distance is 0. The optimization process is to minimize the total distance of all training points. After the hyperplane is determined, the decision function for the classification rule can be written as

$$h(x) = sign(f(x)) \tag{26}$$

A new detection x, is classified based on which side of the hyperplane it lies, i.e., it is declared a metastasis if $h(x) > 0$, or a non-metastasis if $h(x) < 0$. The feature values in SVM are normalized to the range of $[-1, +1]$. The normalization factor is obtained from the training data and applied to the testing data.

An SVM in higher dimensional space (more features) can lead to more accurate classification. However, SVM in a very high dimensional space may increase the complexity of the model, over-train the data and decrease the generality of the model. One solution is to use an ensemble of classifiers, in which each classifier includes a small number of features.

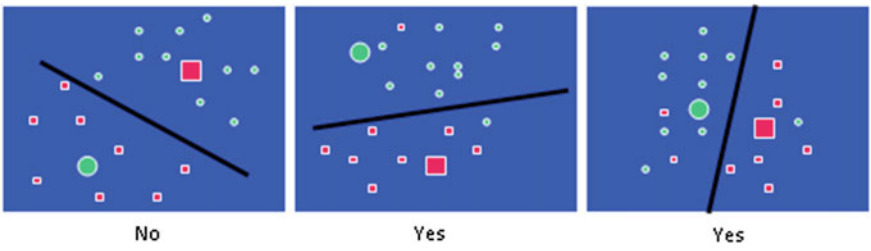


Fig. 7 SVM committee. A committee of three SVMs. Majority vote determines the classification

Ensemble learning combines multiple trained classifiers under the assumption that multiple models are better than one if they are diverse. Popular ensemble approaches include boosting and bagging [50]. Combination strategies for the multiple decisions can be divided into two types: those that adaptively adjust to the training set based on the performance of previous models as in boosting methods and those do not as in bagging [51]. The bootstrap is widely used to estimate the standard error or confidence intervals of an estimate. Bagging is based on the bootstrap technique where the predictions on bootstrapped samples are aggregated to form an ensemble hypothesis. Boosting combines the predictions from re-sampled data based on the previous model's performance such that harder data samples for the system are more likely to be sampled. Bagging has been shown to reduce the variance of classifiers while boosting can reduce both variance and bias. However, it is rarely shown that bagging and boosting combined with feature selection can significantly reduce ensemble training time in practice. In our method, we adopt the boosting approach.

Instead of using one SVM in a very high dimensional feature space, we break the feature space into subsets of low dimension feature spaces (also known as feature vectors). Each feature vector established one SVM, and all SVMs form a committee. We allow overlap of features between different feature subsets. This scheme combines the advantages of using a large number of features and keeping the feature space small for single SVM in the committee. Each member in the committee has one vote for the classification, i.e., if the decision function of the SVM is greater than 0, the vote is 'yes', otherwise the vote is 'no'. The majority vote is used as the decision function of the committee. The committee approach generally produces improved results, provided that the error rate for each member is less than 50 %. Figure 7 demonstrates how the SVM committee works. This is a committee of three SVMs. In the first SVM, there are two misclassified data (big square and big circle), but in the second and third SVM, they are correctly classified. By a majority vote, a correct classification is reached.

In order for the committee to achieve optimal performance, SVM members should be able to compensate each other. If only a few top feature vectors are selected, they usually tend to overlap each other and putting them together will not enhance the differentiating power of the committee. Therefore, a large pool of feature vectors should be available for committee member selection. We developed a progressive feature vector selection method for this purpose [52]. The goal of the feature vector selection is to generate a large pool of feature vectors to be used as candidates for committee members. The task is to select K feature vectors with best performance, and each vector has N features. Here K is a large number, and N is a relatively small number. There are several commonly used feature selection schemes, including exhaustive search, forward stepwise search, and genetic algorithm [20]. Exhaustive search can be very time consuming if hundreds of features are available as candidates. Forward stepwise search is easily trapped in local minimum, and genetic algorithm is sensitive to the initial population.

We proposed a progressive search method to efficiently select a group of K best N-feature vectors [53]. In this method, N-feature vectors are formed progressively in N stages. In each stage, one more feature is added to the vectors selected from the

previous stage. Those new feature vectors are ranked by their performance and only the K top feature vectors are passed to the next stage. The rationale behind this scheme is that feature vectors with the worst performance in $N - 1$st stage are unlikely to be in the top group in the stage N after one more feature is added. Essentially, this method combines the benefit of exhaustive search and forward stepwise search. Only a limited number of vectors are exhaustively examined in each stage, and the performances of feature vectors improve from stage to stage.

9 CAD Performance Evaluation

The quality of a CAD system can be characterized by its sensitivity and specificity in detecting lesions. We used FROC analysis to evaluate the overall performance of our CAD systems. The software used for the FROC analysis of the data was the ROCKIT ROC analysis software subroutine library (ROCKIT; C. E. Metz, B. A. Herman, C. A. Roe, University of Chicago, Ill; http://xray.bsd.uchicago.edu/krl/KRL_ROC/software_index6.htm), which fits a bi-normal distribution to data using a maximum likelihood estimate [54]. The operating point was chosen based on expected sensitivity in the clinical setting.

Analysis for statistical significance of sensitivity value differences between the training and testing sets was performed via a bivariate chi-square test. CT attenuation values and mean volumes of the manually and computer performed lesion segmentations were compared via mixed model ANOVA analysis taking into consideration of multiple lesions in one patient. $P < 0.05$ was considered statistically significant. XLSTAT (www.xlstat.com) was used for data analysis.

10 Data Sets

The CAD system was tested on two cohorts of patients. The first cohort has primarily lytic metastasis and the second cohort has primarily sclerotic metastasis.

10.1 Lytic Metastasis Cohort

The study group consisted of 50 patients (30 men and 20 women, age range 18–82, mean age 54.8), divided into training and test cases (29 and 21 patients, respectively). Patients carried the diagnoses of melanoma, renal cell carcinoma, prostate cancer, lung cancer, lymphoma, breast cancer, pheochromocytoma, or other disorders (19, 10, 4, 4, 2, 2, 2, and 7 patients, respectively. Each patient was scanned with either a 4-detector (Lightspeed QX/I, GE Healthcare, 20 patients), 8-detector (Lightspeed Ultra, GE Healthcare, 29 patients) or 16-detector (Mx8000 IDT, Phillips, 1 patient) CT scanner. Images were obtained at 5 mm slice thickness. Data sets consisted of an

average of 124 images (range 60–145). Patients had a CT of the chest, abdomen, and pelvis (44 patients), abdomen and pelvis (2 patients), chest (1 patient), chest and abdomen (1 patient), abdomen (1 patient), or pelvis (1 patient). The standard reconstruction kernel was used for 49 patients and the "B" kernel for 1 patient. The patients received 110–130 cc Iopamidol (Isovue-300, Bracco Diagnostics, Princeton, NJ) intravenous contrast material given via power injector. Using a two-reader consensus, all thoracolumbar spine lesions were identified and qualitatively characterized as lytic, sclerotic, or mixed—at least 20 % lytic voxels in a predominately sclerotic lesion, vice versa, or lesion without a predominant voxel type. The lytic lesions were further characterized as probable or unlikely metastases; only the probable metastases were included in this study. For example, unlikely metastases including Schmorl's nodes, degenerative disc disease, osteopenia, and hemangiomas, were excluded.

The largest lytic area of each lesion was measured by the largest area enclosed by the contour in the x-y plane and if lesions merged at some point, they were considered one lesion. Due to the 5 mm thickness of slices, the z plane was not used. The lesions were stratified by the maximum area on x-y plane. They were put in three size categories, >0, >0.2 and >0.8 cm^2. Among these, we set 0.8 cm^2 minimum lytic area threshold (equivalent to 1 cm in diameter) for lesions of substantial size and clinically critical. There were 28 probable lytic metastases with lytic areas >0.8 cm^2, which corresponds to a circle with a diameter of >1 cm (12 in the training set, 16 in the test set). Patients had between zero and four probable lytic metastases (average 0.6) with areas 0.9–10.6 cm^2 (average 2.7 cm^2). 33 patients with lytic lesions did not have any lesions characterized as probable metastases. A total of 35 mixed lesions and 37 sclerotic lesions were also present in the study group. Lesions were manually segmented by a trained student, who drew a contour along the voxels on the edge of each lesion on each slice that it appeared. The manual segmentation was used as the reference standard segmentation in our study.

The cohort was divided into training and test sets (29 patients in training and 21 in test). There were in total 90 lytic bone metastases (58 in training and 32 in test set). The data are summarized in Table 2.

10.2 Sclerotic Metastasis Cohort

The sclerotic metastasis cohort consisted of CT examinations from 60 patients (mean age 56.2 years, range 12–77 years; 19 females, 41 males). 50 of them demonstrated one or more sclerotic lesions of the spine. 10 of them were control

Table 2 Data summary for lytic metastasis cohort

	Number of patients	Lesions > 0 cm^2	Lesions > 0.2 cm^2	Lesions > 0.8 cm^2
Training set	29	58	44	12
Test set	21	32	27	16
Total	50	90	81	28

cases. The medical records of the patients whose studies were selected for the study were reviewed for demographics and pathology (27 with prostate cancer and the rest with other types of pathology). CT examination technique varied by the scanners, and across the time interval of the case studies. CT examination of the chest, abdomen and pelvis was performed on 56 patients, CT of the abdomen and pelvis on 3 patients, and CT of the chest and abdomen on 1 patient. 58 of the 60 patients received intravenous contrast as part of their examination, with specific information not available on two of the 58 patients. Contrast was given as Isoview 300 (iopamidol injection 61 %; Bracco Diagnostic Inc., Princeton, NJ) to 44 patients in volumes ranging from 40 ml (n = 1) to 130 ml (n = 41), with one patient receiving 125 ml of Isovue-370 (lopamidol Injection 76 %; Bracco Diagnostic Inc., Princeton, NJ), and one receiving 110 ml of Omnipaque-300 (iohexol, iodine content 46.36 %, GE Healthcare Inc., Princeton, NJ).

CT images were reviewed by a musculoskeletal radiologist with 5-years of experience. Manual segmentation was performed over the three-dimensional extent of each lesion, using in-house software with a bone window setting [55]. These marked lesions formed the reference standard for determination of CAD software performance. Large heterogeneous sclerotic regions that appeared to be either an amalgam of inseparable smaller component sclerotic lesions or a single heterogeneous lesion were location-marked and segmented as a single lesion. A total of 552 sclerotic lesions of the spine were greater than 0.84 cm average diameter.

The cohort was divided into training and test sets (17 patients in training and 43 in test). The data are summarized in Table 3. It also summarizes the size and intensity distribution of the metastases in the data set.

Table 3 Summary of sclerotic metastasis cohort

	Reference standard lesion sets			
	Training		Testing	
Number of cases	17		43	
Number of lesions	180		372	
Volume[1] range	Min	Max	Min	Max
	0.3	52.7	0.3	56.9
Volume[1] mean ± standard deviation (SD)	5.4 ± 9.6		3.3 ± 6.8	
Length[2] range	Min	Max	Min	Max
	0.7	8.1	0.8	7.6
Length[2] mean ± SD	2.4 ± 1.6		2.0 ± 1.2	
Range of the mean HU (CT attenuation) per lesion	Min	Max	Min	Max
	91.8	878	161	821
Mean of the mean HU per lesion ± SD	525 ± 150		391 ± 106	
Range for the SD of the mean HU per lesion	Min	Max	Min	Max
	24	322	28	317
Mean of SD of the mean HU per lesion ± SD of SD	179 ± 65		124 ± 54	

1. Units of cm^3, 2. Units of cm

11 Results

The results are presented separately for lytic and sclerotic metastasis detection since they are independent data sets. For each experiment, we present visual detections, FROC analysis, and etiology of false positives and false negatives.

11.1 Lytic Metastasis Detection Results

Figure 8 shows the examples of detection results for lytic metastasis. Figure 9 shows the FROC analysis on the training and test sets, per-lesion analysis in Fig. 9a and per-patient analysis in Fig. 9b. Given the operating point of 0.5 for the SVM value, in the training set, the per-lesion sensitivities were 72.4, 84.1 and 91.6 % for

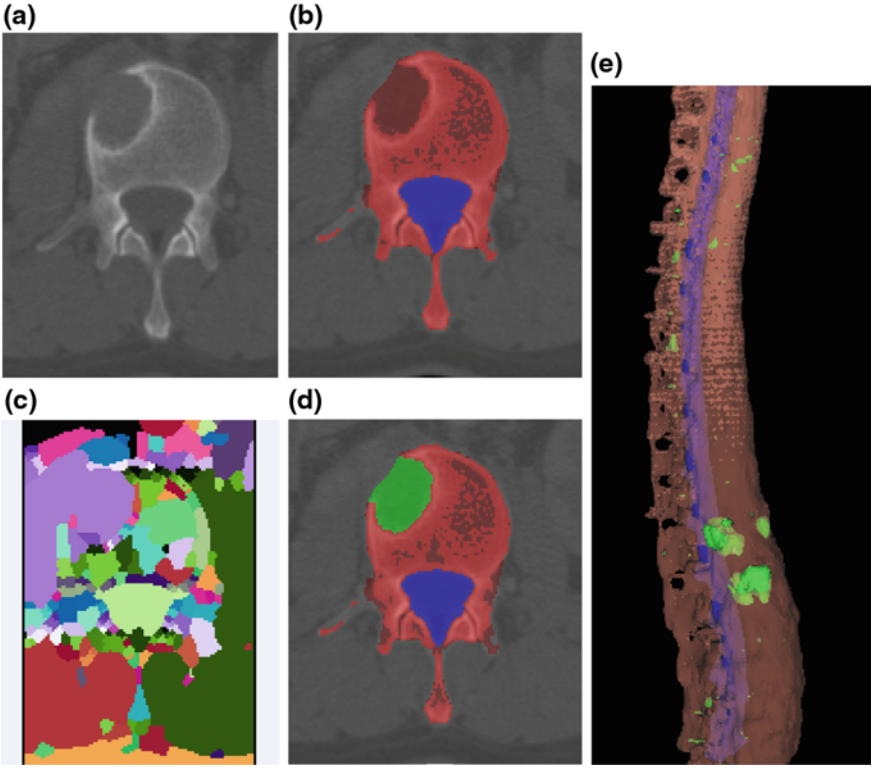


Fig. 8 Lytic metastasis detection. **a** CT image; **b** spine segmentation; **c** watershed segmentation; **d** lesion segmentation; **e** lytic detection in 3D. *Green* lytic lesion detection; *Red* vertebra segmentation; *Blue* spinal canal

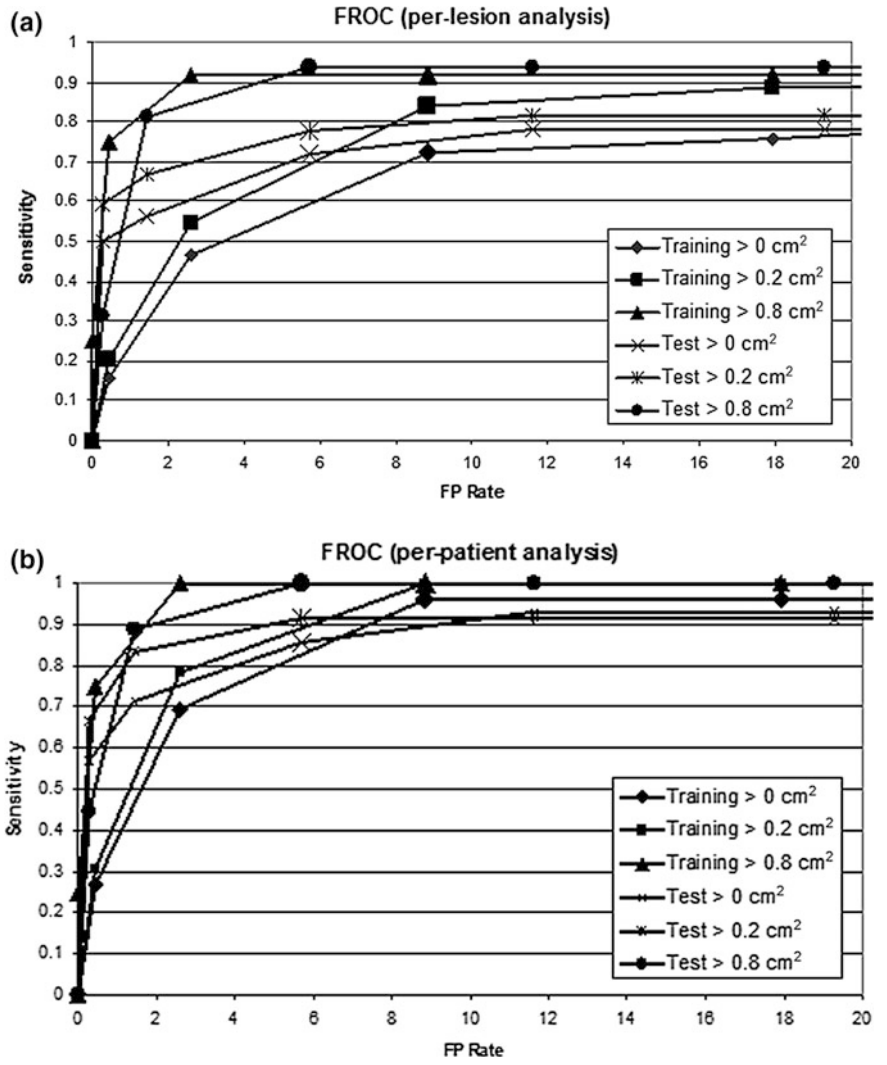


Fig. 9 FROC analysis for lytic metastasis CAD. **a** Per-lesion analysis; **b** per-patient analysis

lesions >0, >0.2 and >0.8 cm^2, and the per-patient sensitivities were 96.2, 100 and 100 %, with 8.8 false positive per patient. In the test set, using the same operating point, the per-lesion sensitivities were 71.9, 77.8 and 93.8 %, and the per-patient sensitivities were 85.7, 91.7 and 100 %, with 5.7 false positive per patient. The difference between the sensitivities in the training and test sets were not statistically significant ($p = 0.56$).

The sources of the 310 false positive detections (FPs) in the training and test sets could be broken down into 7 main categories: (1) "Peripheral vein", on venous

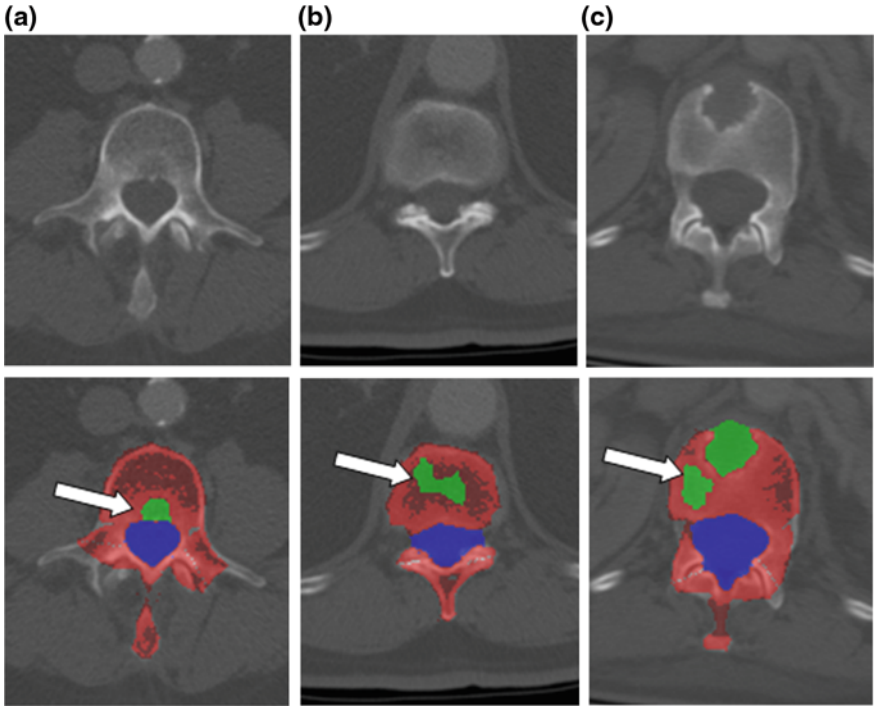


Fig. 10 False positive detections in lytic metastasis CAD. *First row* CT image; *Second row* false positive detections. **a** Basivertebral vein; **b** vertebral disk; and **c** volume averaging

connections between the basivertebral vein and the anterior external venous plexus (106, 34 % of FPs), (2) "Disk", low intensity disks or volume averaging with disks (83, 27 %), (3) Osteopenia (68, 22 %), (4) "Outside", on areas outside the vertebra (37, 12 %), (5) "Basivertebral vein", which enters the posterior vertebral body (6, 2 %), (6) "Normal", a drop in intensity from volume averaging with normal structures such as joints or oblique cuts through the cortex (6, 2 %), and (7) "Spinal canal" (2, 1 %). Two of the FP detections were actually on reference standard lesions that were not segmented on all slices in which they appeared. False positive detections varied greatly amongst patients, numbering 0 to 20 per patient (average 6.2). Some examples of FPs are shown in Fig. 10. We also analyzed the 3 false negative detections (FNs) (two in the training set, one in the test set). Two were in pedicles that were not properly segmented, so they were never detected. The other FN was initially detected, but thrown out by the classifier, most likely due to similarity to a basivertebral vein.

11.2 Sclerotic Metastasis Detection Results

Figure 11 shows one example of sclerotic metastasis CAD results. It shows the output of each stage in the CAD system. A total of 552 sclerotic lesions were electronically location marked and segmented in the ground truth data set of 50 cases, for an average of 11.2 lesions per patient [standard deviation (SD): 10.8], and a range of 0–59 lesions per patient (Table 3). The average number of lesions per patient detected by the CAD system on this data set was 9.9 (SD: 10.1) with a range 0–57 lesions detected per patient. All 50 case CT studies accumulated during the routine clinical review phase of the study appeared positive for lesions upon qualitative visual assessment.

Figure 12 shows the FROC analysis for the sclerotic metastasis CAD system. The operating point value for the training set was chosen to target a sensitivity of 90 % or greater, while maintaining a clinically reasonable false positive rate. This same operating point was then used for the testing set. The CAD system sensitivity for the training set was 0.90 (95 % confidence interval: 0.86, 0.92) at a false positive rate (FPR) of 8.9 (95 % CI: 6.3, 12.0) per patient. The testing set sensitivity was 0.84 (95 % CI: 0.80, 0.87) with an FPR of 11.3 (95 % CI: 9.3, 13.6) per patient.

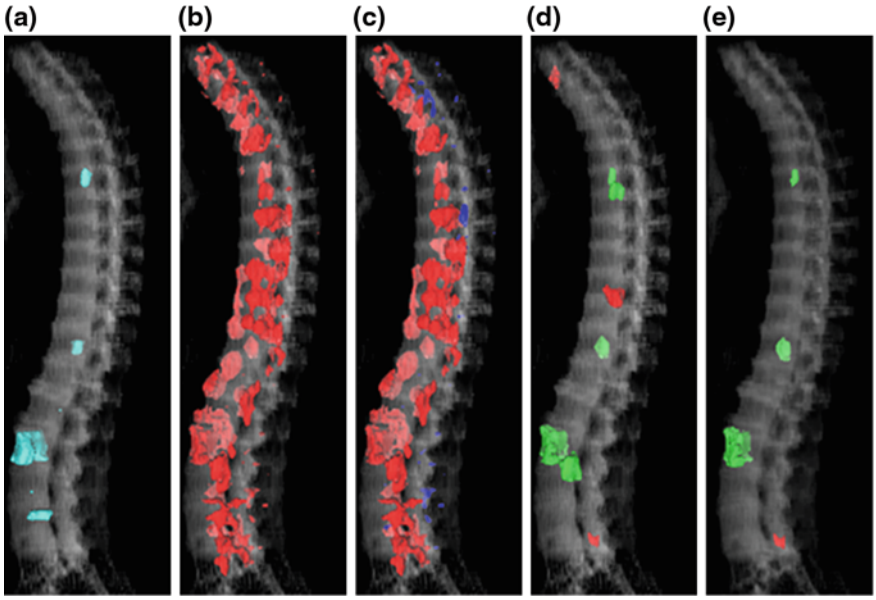


Fig. 11 Sclerotic metastasis CAD. 75 year-old patient with prostate cancer. **a** 3-D segmented spine with ground truth lesions marked in *light blue*. **b** Detections (*red*) from 3D merging algorithm, prior to detection filter screening (*red*). **c** *Red* candidate lesions remaining after screening by detection filter. *Blue* detections eliminated by filter. **d**, **e** Lesions remaining post SVM classification with cutoff 0.48 (**d**) and 0.55 (**e**). *Green* TP lesions. *Red* FP lesions. False positive detections due to degenerative change and partial volume vertebral end plates

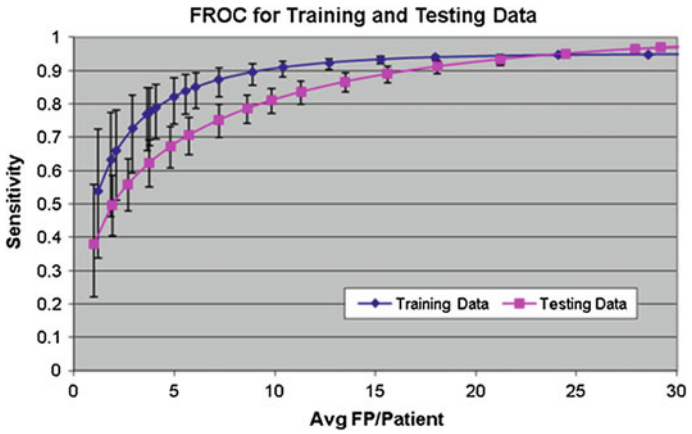


Fig. 12 FROC analysis for sclerotic metastasis CAD

The testing and training sets did not demonstrate a statistically significant difference in sensitivity, with a bivariate chi-square test statistic of 3.7 ($p = 0.055$).

The etiology of false positive and false negative detections was evaluated. If a FP was found to represent a true lesion, not marked by the radiologists creating the reference standard set but detected by the CAD system, it was manually removed from the FP statistic. There were a total of 15 true lesions detected by the CAD system, not marked in the reference standard set. These lesions were not included in the reference standard set because of small size and low attenuation in eight (53 %), because of low attenuation and location at endplate in three (20 %), because of small size in two (13 %), because the lesion was in an L5 with sacralization that was thought to be outside the region of interest at ground truth marking in one (7 %), and because of inadvertent deletion of one electronic lesion segmentation data file during creation of reference standard set in one (7 %). FP detections were most often attributable to degenerative sclerosis (174 [28.1 %] of 620 actual detections) and misclassification of vertebral endplate bone cortex lying parallel to the (axial) imaging plane (173 [27.9 %] of 620) (Table 4). Other causes are noted in Table 4. There were 93 false-negative findings, with 37 (40 %) caused by vertebral body endplate proximity, 32 (34 %) caused by low attenuation, 17 (18 %) caused by small size, and seven (8 %) caused by other reasons, such as the finding was out of the search region. Figure 13 shows a few examples.

Quantitative metrics were calculated for both the ground truth and computer-detected lesions. The difference in mean lesion volume between ground truth and computer-aided detections was not statistically significant, with an approximate volume difference of 0.3 % in the training set ($t = 0.02$, $p = 0.987$), and 0.9 % difference in the testing set ($t = 0.07$, $p = 0.943$). The difference in mean lesion CT attenuation (HU) between manually segmented lesions and computer-aided detections was not statistically significant, with 8.0 % difference for the training set ($t = 2.7$, $p = 0.006$) and 10 % for the testing set ($t = 4.95$, $p < 0.0001$).

Table 4 False positive etiologies for sclerotic metastasis CAD

Degenerative Sclerosis		Volume Averaging at pedicle cortex		Cortex of vertebral endplate		Sclerotic intravertebral channel wall basivertebral veins		Sclerotic wall Schmorl's node		Other (bone island, physeal scar, etc.)	
174 (28.1 %)		121 (19.5 %)		173 (27.9 %)		40 (6.4 %)		20 (3.2 %)		92 (14.8 %)	
Training set	Testing set	Training set	Testing set	Training set	Testing set	Training set	Testing set	Training set	Testing set	Training set	Testing set
41	133	47	74	48	125	4	36	5	15	27	65
*635 total false positive etiologies											

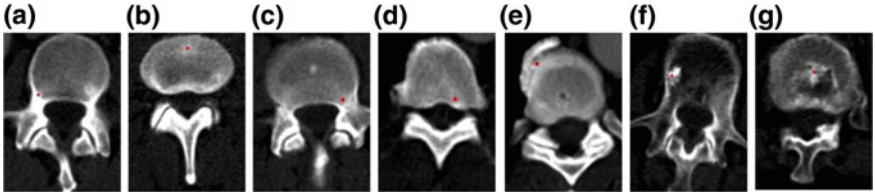


Fig. 13 False positive detections in sclerotic metastasis CAD. FPs (*red dot*) due to: **a** cortex of neuroforamen, **b** cortex of endplate, **c** bone cortex/neurocentral synchondrosis, **d** sclerotic margin basivertebral bundle, **e** degenerative sclerosis of osteophyte, **f** bone island, **g** sclerotic Schmorl's node margin

12 Discussion

While CT is not the study of choice for initial whole body screening the whole patient for bone metastases–skeletal scintigraphy or conventional radiography are the most common choices [56–60]—metastases still must be identified on CT when possible. However, since the cost of CT is decreasing and the radiation exposure is also reducing, CT is becoming more affordable and available for screening purposes. Additionally, the ability to detection metastases on CT continues to grow in importance with the increasingly widespread availability and use of integrated PET/CT imaging in the detection and follow up of metastatic disease. Using bone window settings, CT shows a high level of detail in bone, distinguishing amongst materials of different radio densities [61, 62]. For depicting metastases to the spine, CT is superior to skeletal scintigraphy and conventional radiography [11, 18] and has performed at sensitivities ranging from 93 to 100 % [17, 63, 64]. However, these lesions can be subtle and easily overlooked by a radiologist, especially when bone windows are underutilized [58] and the radiologist has not been specifically directed by the referring physician to look for metastases.

Detecting spinal lesions by computer is challenging, owing to the variation in bone attenuation within and amongst patients as well as the diversity of non-metastatic abnormalities such as degenerative disk disease. The problem must be broken down into manageable components that can be addressed sequentially. Our system detects lytic and sclerotic metastases separately and the results can be combined. The first task in detecting spinal metastases is to locate and segment the spine, excluding other structures. This is most difficult in the thoracic spine, where the ribs are often detected along with the vertebrae. We included location criteria in our filter and classifier to account for this, but a number of the "outside" false positives were detections on costovertebral joints, as they are low intensity regions surrounded by high attenuation cortex. The synovial joints between adjacent vertebrae and the nearby contrast-filled IVC are sometimes segmented along with the lumbar vertebrae, resulting in "outside", non-bone related false positives.

While adjacent high intensity structures pose a challenge for segmentation of the spine from non-spine structures, the intrinsically low intensity intervertebral disks

pose a challenge during both segmentation and lytic lesion detection. Disks and lytic lesions are both low intensity regions, so disks may be mistaken for lesions and disks may cause volume averaging with the vertebrae. This accounts for 27 % of false positive detections in the training set. These false positive detections may be reduced if the intervertebral disks can be automatically identified within the spinal column.

The 5 mm slice thickness of our dataset posed another challenge to the segmentation of the spine and detection of lesions. This thickness is common for routine CTs of the chest, abdomen, and/or pelvis. Our system is designed to find unexpected spinal metastases on examinations ordered for other indications. Thick slices lead to volume averaging, causing parts of the vertebral body to have intensity similar to surrounding soft tissues. Substantial leakage (segmentation of undesired structures) could happen when region-based segmentation is applied. We adopted a multi-pass technique to address this problem. First a high threshold was applied to get the initial segmentation, then morphological operations and rolling balls were applied to close the holes and gaps. Thick slices increase "normal" false positive detections, especially in the vertebral arch, when oblique cuts result in volume averaging of vertebral cortex and adjacent soft tissues. Finally, the slice thickness makes lesions more difficult to detect, as most only appear on one slice. Other low intensity structures that cause false positive detections are the spinal canal, the basivertebral vein, and the vein's connections to the anterior external venous plexus. These false positives have characteristic features, especially in terms of location and shape, which may be used in future CAD systems to recognize and eliminate them from the CAD potential lesion list that is presented to radiologists for consideration.

False negatives can be attributed to two main causes. Two of them were due to failures in segmentation of the pedicle, while the final false negative was due to a failure of characterization. Further work could increase the accuracy of segmentation of the pedicle, while a larger training set may help the classifier distinguish between true lesions and false positives.

Quantitative metrics for lesion volume and CT attenuation were calculated for comparison of CAD system and manually performed lesion characterization. The difference in mean lesion volume and CT attenuation between ground truth and computer-aided detections was not statistically significant. Thus, the CAD system was able to quantitatively characterize the detected lesions with good agreement when compared to the manually segmented data set. Of particular importance, manual tracing of the margins of each lesion for electronic segmentation, with each lesion typically extending over multiple axial image slices required hours of radiologists' time for typical cases. The CAD system was able to automatically analyze each case in less than 2 min on a standard office desktop computer.

There were limitations in CAD system design. First, the bounding region for lesion search was limited to the body and pedicles of the vertebrae. This anatomic simplification was felt justified as this work was intended as a preliminary proof of concept study, and it has been shown previously that the primary loci for (early) metastatic spread to the vertebrae are in the vertebral body and pedicles, and the

majority of the volume of the vertebra are constituted by these regions [65–67]. Bounding box inclusion of the vertebral body and pedicles thus assesses the most likely regions of early metastatic tumor involvement. Future work could include the laminae, transverse and posterior spinous processes for a more complete evaluation and characterization of sclerotic tumor burden. Second, the false positive (FP) rate is relatively high. Two of the three most common causes (volume averaging of cortex at vertebral endplates and pedicles) may be decreased by a more sophisticated segmentation algorithm design, and the third (degenerative sclerosis) by an addition of an algorithm designed to detect degenerative change of the vertebrae. Elimination of these three FP etiologies would eliminate 75 % of false positive detections. Third, images reconstructed with a soft tissue kernel were used to decrease the effect of image noise on software performance. Future design may include increased robustness in the presence of image noise. Finally, sclerotic lesions and lytic lesions were detected separately. An integrated approach to lytic and sclerotic lesion detection and characterization is under investigation.

Since CAD technologies are still under intensive development, most studies of CAD systems to-date have reported its performance in the laboratory setting rather than in the radiology reading room. This CAD system is designed for clinical application as a secondary reader to increase the sensitivity for detection of sclerotic metastatic lesions in the spine. Potential future practical applications include quantification of bone tumor burden and of change in individual lesions and total tumor volumes with generation of metrics, for follow-up examinations in patients undergoing treatment, as well as to assess for localized new or changing density lesions.

In conclusion, we have presented a CAD system that can detect both lytic and sclerotic metastases in the thoracolumbar spine. We have identified some of the common causes of false negative and false positive detections to guide further development of bone CAD systems. The CAD framework is based on supervised machine learning techniques and can be employed to detect other abnormalities in the spine, such as fractures, osteophytes and epidural masses. Additional research will be required to show whether bone CAD systems improve radiologists' diagnostic accuracy and interpretive efficiency.

References

1. Hitron A, Adams V (2009) The pharmacological management of skeletal-related events from metastatic tumors. Orthopedics 32:188
2. Roodman GD (2004) Mechanisms of bone metastasis. N Engl J Med 350:1655–1664
3. Guillevin R, Vallee JN, Lafitte F, Menuel C, Duverneuil NM, Chiras J (2007) Spine metastasis imaging: review of the literature. J Neuroradiol 34:311–321
4. Lee RJ, Saylor PJ, Smith MR (2011) Treatment and prevention of bone complications from prostate cancer. Bone 48:88–95
5. Chirgwin JM, Guise TA (2007) Skeletal metastases: decreasing tumor burden by targeting the bone microenvironment. J Cell Biochem 102:1333–1342

6. Kinnane N (2007) Burden of bone disease. Eur J Oncol Nurs 11(Suppl 2):S28–S31
7. Weinfurt KP, Li Y, Castel LD, Saad F, Timbie JW, Glendenning GA, Schulman KA (2005) The significance of skeletal-related events for the health-related quality of life of patients with metastatic prostate cancer. Ann Oncol 16:579–584
8. Saad F, Lipton A, Cook R, Chen YM, Smith M, Coleman R (2007) Pathologic fractures correlate with reduced survival in patients with malignant bone disease. Cancer 110:1860–1867
9. Bilsky MH, Lis E, Raizer J, Lee H, Boland P (1999) The diagnosis and treatment of metastatic spinal tumor. Oncologist 4:459–469
10. Colman LK, Porter BA, Redmond J, Olson DO, Stimac GK, Dunning DM, Friedl KE (1988) Early diagnosis of spinal metastases by CT and MR studies. J Comput Assist Tomogr 12:423–426
11. Coleman RE (1998) Monitoring of bone metastases. Eur J Cancer 34:252–259
12. Beheshti M, Vali R, Waldenberger P, Fitz F, Nader M, Hammer J, Loidl W, Pirich C, Fogelman I, Langsteger W (2009) The use of F-18 choline PET in the assessment of bone metastases in prostate cancer: correlation with morphological changes on CT. Mol Imaging Biol 12:98–107
13. Mundy GR (2002) Metastasis to bone: causes, consequences and therapeutic opportunities. Nat Rev Cancer 2:584–593
14. Guise TA, Mundy GR (1998) Cancer and bone. Endocr Rev 19:18–54
15. Keller ET, Brown J (2004) Prostate cancer bone metastases promote both osteolytic and osteoblastic activity. J Cell Biochem 91:718–729
16. Saylor PJ, Smith MR (2009) Bone health and prostate cancer. Prostate Cancer Prostatic Dis 13:20–27
17. Muindi J, Coombes RC, Golding S, Powles TJ, Khan O, Husband J (1983) The role of computed tomography in the detection of bone metastases in breast cancer patients. Br J Radiol 56:233–236
18. Sundaram M, McGuire MH (1988) Computed tomography or magnetic resonance for evaluating the solitary tumor or tumor-like lesion of bone? Skeletal Radiol 17:393–401
19. Jeschke S, Schweigreiter E, Janetschek G (2009) Role of imaging in prostate cancer. Imaging Decisions MRI 13:68–87 (Fall/Winter 2009)
20. Li J, Yao J, Summers RM, Petrick N, Manry MT, Hara AK (2006) An efficient feature selection algorithm for computer-aided polyp detection. Int J Artif Intell Tools (IJAIT) 15:893–915
21. Yao J, Dwyer A, Summers R, Mollura D (2011) Computer-aided diagnosis of pulmonary infections using texture analysis and support vector machine classification. Acad Radiol 18:306–314
22. Chakraborty D (2000) The FROC, AFROC and DROC variants of the ROC analysis. In: Beutel J, Kundel H, Van Metter R (eds) Handbook of medical imaging. SPIE Press, Bellingham, pp 771–796
23. Irwig L, Houssami N, van Vliet C (2004) New technologies in screening for breast cancer: a systematic review of their accuracy. Br J Cancer 90:2118–2122
24. Hadjiiski L, Chan HP, Sahiner B, Helvie MA, Roubidoux MA, Blane C, Paramagul C, Petrick N, Bailey J, Klein K, Foster M, Patterson S, Adler D, Nees A, Shen J (2004) Improvement in radiologists' characterization of malignant and benign breast masses on serial mammograms with computer-aided diagnosis: an ROC study. Radiology 233:255–265
25. Awai K, Murao K, Ozawa A, Komi M, Hayakawa H, Hori S, Nishimura Y (2004) Pulmonary nodules at chest CT: effect of computer-aided diagnosis on radiologists' detection performance. Radiology 230:347–352
26. Kakeda S, Moriya J, Sato H, Aoki T, Watanabe H, Nakata H, Oda N, Katsuragawa S, Yamamoto K, Doi K (2004) Improved detection of lung nodules on chest radiographs using a commercial computer-aided diagnosis system. AJR Am J Roentgenol 182:505–510
27. Summers RM, Jerebko AK, Franaszek M, Malley JD, Johnson CD (2002) Colonic polyps: complementary role of computer-aided detection in CT colonography. Radiology 225:391–399

28. O'Connor SD, Yao J, Summers RM (2007) Lytic metastases in thoracolumbar spine: computer aided detection at CT—a preliminary study. Radiology 242:811–816
29. Burns J, Yao J, Wiese T, Munoz H, Jones E, Summers R (2013) Detection of sclerotic metastases in the thoracolumbar spine on computed tomography. Radiology 268:69–78
30. Yao J, Burns JE, Muñoz H, Summers RM (2012) Detection of vertebral body fractures based on cortical shell unwrapping. In: Proceedings of the 15th international conference on medical image computing and computer assisted intervention, Nice, France, pp 509–516
31. Yao J, Muñoz HE, Burns JE, Lu L, Kurdziel K, Choyke P, Summers RM (2013) Computer aided detection of spinal degenerative osteophytes on sodium fluoride PET/CT. In: MICCAI workshop, computational methods and clinical applications for spine imaging, Nagoya, Japan
32. Tan S, Yao J, Ward M (2008) Computer aided evaluation of ankylosing spondylitis using high-resolution CT. IEEE Trans Med Imaging 27:1252–1267
33. Pattanaik S, Liu J, Yao J, Zhang W, Turkbey E, Zhang X, Summers R (2013) Epidural masses detection on computed tomography using spatially-constrained gaussian mixture models. In: MICCAI workshop, computational methods and clinical applications for spine imaging, Nagoya, Japan
34. Vincent L, Soille P (1991) Watersheds in digital spaces: an efficient algorithm based on immersion simulations. IEEE Trans Pattern Anal Machine Intell 13:583–598
35. Ibanez L, Schroeder W (2003) ITK software guide. Kitware Inc., New York
36. Cormen TH, Leiserson CE, Rivest RL (1989) Introduction to algorithms. Mac Graw Hill, New York
37. Wolberg G (1990) Digital image warping. IEEE Computer Society Press Monograph
38. Ghebreab S, Smeulders AWM (2004) Combining strings and necklaces for interactive three-dimensional segmentation of spinal images using an integral deformable spine model. IEEE Trans Biomed Eng 51:1821–1829
39. Heitz G, Rohlfing T, Maurer CRJ (2005) Statistical shape model generation using nonrigid deformation of a template mesh. In: SPIE, San Diego, CA
40. Verdonck B, Nijlunsing R, Gerritsen FA, Cheung J, Wever DJ, Veldhuizen A, Devillers S, Makeram-Ebeid S (1998) Computer assisted quantitative analysis of deformities of the human spine. In: MICCAI, Cambridge, MA, pp 822–831
41. Pizer SM, Thall AL, Chen DT (1999) M-Reps: a new object representation for graphics. University of North Carolina, Chapel Hill TR99-030
42. Vrtovec T, Likar B, Pernus F (2005) Automated curved planar reformation of 3D spine images. Phys Med Biol 50:4527–4540
43. Stawiaski J, Decencière E (2008) Region merging via graph-cuts. Image Anal Stereol 27:39–45
44. Boykov Y, Kolmogorov V (2004) An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. IEEE Pattern Anal Mach Intell 26:1124–1137
45. Sethian JA (1999) Level set methods and fast marching methods. Cambridge University Press, Cambridge
46. Bankman IN, Spisz TS, Ravlopoulos S (2000) Two-dimensional shape and texture quantification. In: Bankman IN (ed) Handbook of medical imaging, processing and analysis. Academic Press, New York, pp 215–230
47. Baxt WG (1995) Application of artificial neural networks to clinical medicine. The Lancet 346:1135–1138
48. Cristianini N, Taylor JS (2000) An introduction to support vector machines. Cambridge University Press, Cambridge
49. Wang S, Summers RM (2012) Machine learning and radiology. Med Image Anal 16:933–951
50. Bauer E, Kohavi R (1998) An empirical comparison of voting classification algorithms: bagging, boosting, and variants. Mach Learn 36:1–38
51. Breiman L (1996) Bagging predictors. Mach Learn 24:123–140
52. Yao J, Campbell S, Hara AK, Summers RM (2004) Progressive feature vector selection scheme for computer aided colonic polyp detection. In: RSNA 2004, Chicago, IL

53. Yao J, Summers RM, Hara AK (2005) Optimizing the support vector machines (SVM) committee configuration in colonic polyp CAD system. In: SPIE medical imaging, San Diego, CA
54. Metz C, Herman B, Roe C (1998) Statistical comparison of two ROC curve estimates obtained from partially-paired datasets. Med Decis Making 18:110–121
55. Jianhua Y, O'Connor SD, Summers RM (2006) Automated spinal column extraction and partitioning. In: 3rd IEEE international symposium on biomedical imaging: nano to macro, 2006, pp 390–393
56. Soderlund V (1996) Radiological diagnosis of skeletal metastases. Eur Radiol 6:587–595
57. Rybak LD, Rosenthal DI (2001) Radiological imaging for the diagnosis of bone metastases. Q J Nucl Med 45:53–64
58. Hamaoka T, Madewell JE, Podoloff DA, Hortobagyi GN, Ueno NT (2004) Bone imaging in metastatic breast cancer. J Clin Oncol 22:2942–2953
59. Rosenthal DI (1997) Radiologic diagnosis of bone metastases. Cancer 80:1595–1607
60. Kagan AR, Bassett LW, Steckel RJ, Gold RH (1986) Radiologic contributions to cancer management. Bone metastases. AJR Am J Roentgenol 147:305–312
61. Coleman RE, Rubens RD (1985) Bone metastases and breast cancer. Cancer Treat Rev 12:251–270
62. Tryciecky EW, Gottschalk A, Ludema K (1997) Oncologic imaging: interactions of nuclear medicine with CT and MRI using the bone scan as a model. Semin Nucl Med 27:142–151
63. Daffner RH, Lupetin AR, Dash N, Deeb ZL, Sefczek RJ, Schapiro RL (1986) MRI in the detection of malignant infiltration of bone marrow. AJR Am J Roentgenol 146:353–358
64. Metser U, Lerman H, Blank A, Lievshitz G, Bokstein F, Even-Sapir E (2004) Malignant involvement of the spine: assessment by 18F-FDG PET/CT. J Nucl Med 45:279–284
65. Algra PR, Heimans JJ, Valk J, Nauta JJ, Lachniet M, Van Kooten B (1992) Do metastases in vertebrae begin in the body or the pedicles? Imaging study in 45 patients. AJR Am J Roentgenol 158:1275–1279
66. Even-Sapir E, Martin RH, Barnes DC, Pringle CR, Iles SE, Mitchell MJ (1993) Role of SPECT in differentiating malignant from benign lesions in the lower thoracic and lumbar vertebrae. Radiology 187:193–198
67. Shah AN, Pietrobon R, Richardson WJ, Myers BS (2003) Patterns of tumor spread and risk of fracture and epidural impingement in metastatic vertebrae. J Spinal Disord Tech 16:83–89

Quantitative Monitoring of Bone Formation in Ankylosing Spondylitis Using Computed Tomography

Sovira Tan

Abstract Ankylosing Spondylitis, an inflammatory disease affecting mainly the spine, can be characterized by abnormal bone structures (syndesmophytes) growing at intervertebral disk spaces. Monitoring the evolution of these syndesmophytes has been a challenge because of their slow growth rate, a problem compounded by the use of radiography and a mainly qualitative rating system. To improve the low sensitivity to change of radiographic reading, we designed a computer algorithm that fully quantitates syndesmophytes in terms of volume using the 3D imaging capabilities of computed tomography. Its reliability was assessed by computing the difference between the results obtained from 2 scans performed on the same day in 9 patients. A longitudinal study performed over 2 years with 33 patients shows that the method holds promise for longitudinal clinical studies of syndesmophyte development and growth. At the end of the first year, 73 % of patients had a volume increase computed by the algorithm compared to only 12 % for the reading of radiographs.

1 Introduction

Ankylosing Spondylitis (AS) is a progressive inflammatory arthritis affecting primarily the spine. It characteristically causes back pain and can lead to structural and functional impairments. As a result, AS patients may suffer from work disability, unemployment, and reduced quality of life. Estimates of prevalence rates range from 0.1 to 1.4 % of the general population. AS is about twice as common in men as in women. It has a known association with an important immunogenetic component of DNA known as HLA-B27. The majority of patients affected by AS are HLA-B27 positive. AS patients may also develop inflammation of tendon–bone

S. Tan (✉)
National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Building 10 CRC, Room 4-1339, Bethesda, MD 20892, USA
e-mail: tanso@mail.nih.gov

© Springer International Publishing Switzerland 2015

S. Li and J. Yao (eds.), *Spinal Imaging and Image Analysis*,
Lecture Notes in Computational Vision and Biomechanics 18,
DOI 10.1007/978-3-319-12508-4_5

junctions and the eye. Commonly affected areas of bony involvement are the spine and sacroiliac joints. Progression of AS is best characterized by abnormal bone (syndesmophytes) formation along the margins of inter-vertebral disk spaces (IDS). Syndesmophytes cause irreversible and progressive structural damage, and over decades, can lead to spinal fusion [1, 2].

Monitoring syndesmophyte evolution is essential for many clinical studies of AS. Recently available treatments, tumor necrosis factor (TNF) inhibitors, have attracted much attention and fostered new hope by substantially reducing signs of inflammation and improving quality of life [3–5]. However it is still an open question whether they slow syndesmophyte growth or not. Most studies seem to show a slight deceleration but without statistical significance [6–10]. The causes of bone formation in AS are still poorly understood. In particular, the involvement of inflammation, which has face value plausibility, constitutes a perplexing and still unanswered question. Evidence of the correlation between inflammation and syndesmophyte growth has been marginal at best despite extensive studies [11–18]. To elucidate the mechanisms of bone formation in AS at a molecular level, correlation between syndesmophyte growth and various biomarkers of bone turnover has been investigated [19, 20]. Predictors of syndesmophyte formation have been sought with only limited success [21, 22]. New promising perspectives on syndesmophyte growth have been opened by genetic studies [23, 24]. In particular, Dickkopf-1 (DKK-1), a regulatory molecule of the Wnt pathway which controls embryonic development, has attracted much attention [25–27].

Unfortunately, all those studies have been hampered by the fact that the current standard for assessing syndesmophyte growth, the visual examination of radiographs, has very poor sensitivity to change. This low sensitivity to change is not only a reflection of the slow growth rate of syndesmophytes. It is also caused by the limitations of radiography, which projects 3D objects onto 2D images with attendant losses of spatial information and ambiguities in density caused by superimposition. Moreover, syndesmophytes on radiographs are usually rated using coarse semi-quantitative reading systems [28, 29]. The modified Stoke Ankylosing Spondylitis Spinal Score (mSASSS) has emerged as the most widely used reading system [30]. The crudeness of the scoring systems further limits sensitivity to change [31]. Figure 1 shows an example of syndesmophyte growth visible on reformatted CT but not radiography.

To overcome the limitations of radiographic methods, we designed a computer algorithm that quantitatively measures syndesmophyte volumes in the 3D space of CT scans [32, 33]. The algorithm is described in the following section. In Sect. 3, we investigate its accuracy and precision. Results of a 2-year longitudinal study are presented in Sect. 4. We review the future challenges of the new method in Sect. 5 before concluding in Sect. 6.

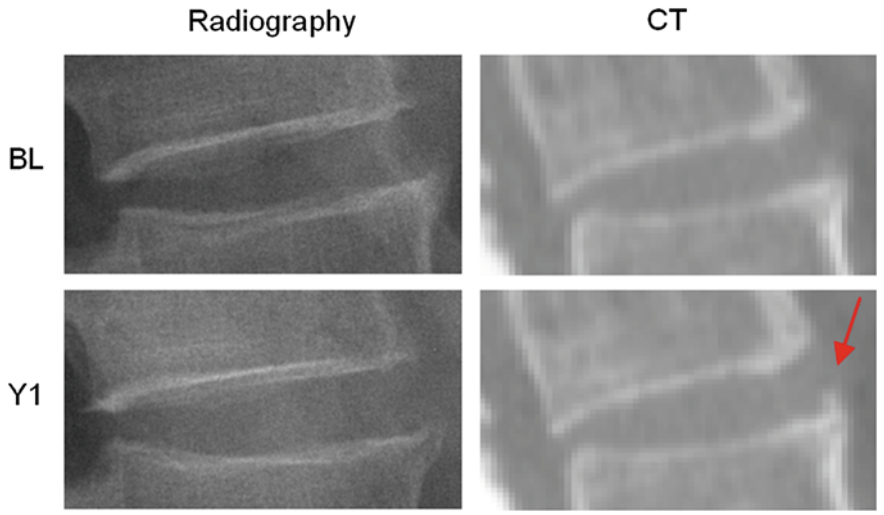


Fig. 1 Example of syndesmophyte growth from baseline (*BL*) to year 1 (*Y1*) visible on CT reformations but not on radiographs

2 The Algorithm

The complete algorithm, summarized in Fig. 2, has of three main parts. First, vertebral bodies are segmented using a 3D multi-stage level set method. Triangular meshes representing the surfaces of the segmentations are made [34]. The 3D surfaces shown in Fig. 2 are triangular meshes obtained from our segmentation results. The vertebral surfaces of corresponding vertebrae are then registered. The purpose of the registration is to extract the syndesmophytes of both vertebrae using the same reference level. Syndesmophytes are cut from the vertebral body using the end plate's ridgeline as the reference level.

2.1 Segmentation of the Vertebral Bodies

Many image processing segmentation techniques have previously been applied to the extraction of vertebral bodies in CT [35–42]. For our algorithm, we chose to use level sets for their flexibility [43]. Flexibility is essential in our application as syndesmophytes can deform the normal vertebral shape in unexpected ways. Level sets are evolving contours or surfaces that can expand, contract, and even split or merge. For the purpose of segmentation they are designed to deform so as to match an object of interest. Many different types of level set exist, depending on the image features chosen to guide the segmentation. For our particular purpose, we selected

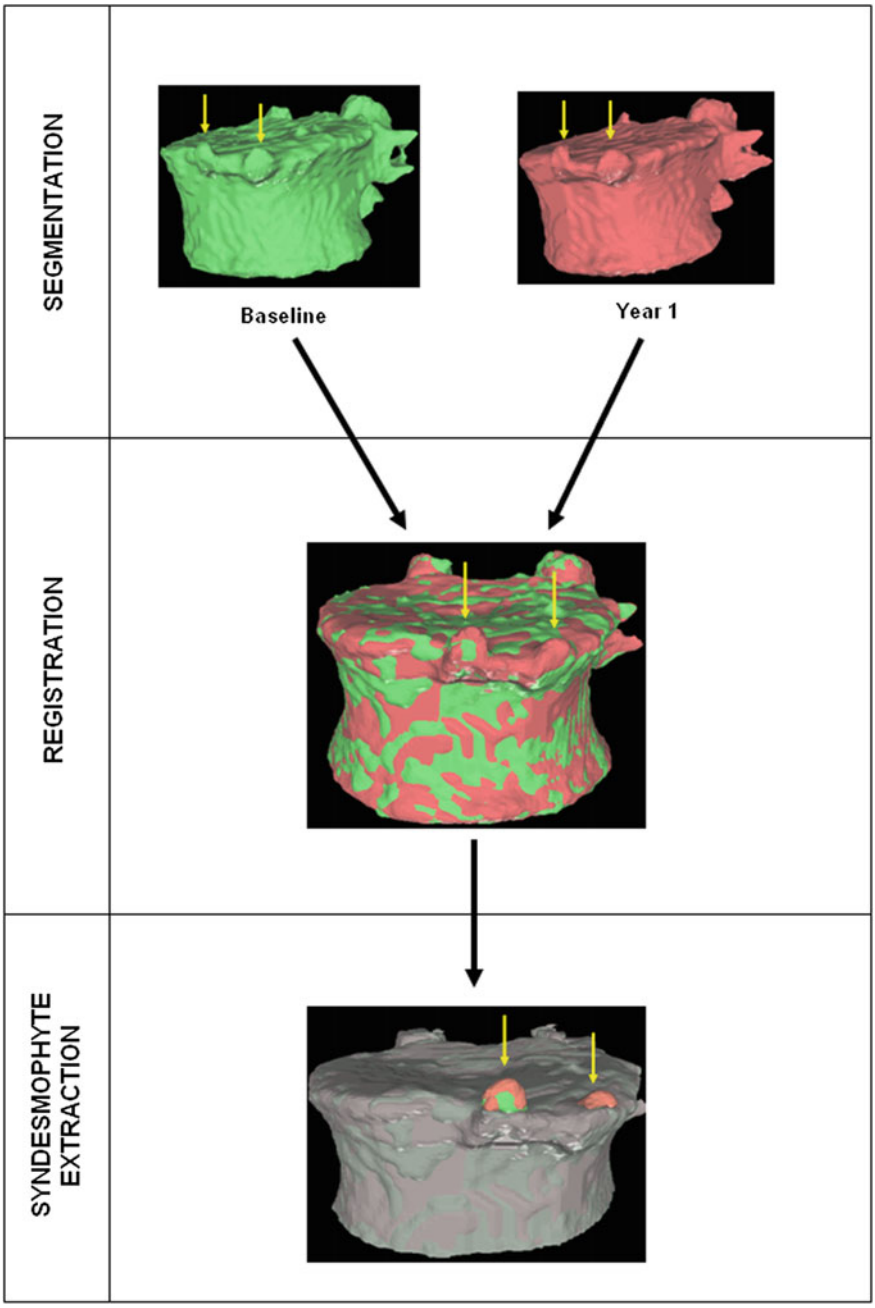


Fig. 2 Overview of the complete algorithm

two level sets based on edge features: the geodesic active contour (GAC) [44] and what we call for convenience the classical level set (CLS) [43]. The GAC evolves according to the equation [44]:

$$\frac{d\psi}{dt} = \alpha g(\vec{x})c|\nabla\psi| + \beta g(\vec{x})\kappa|\nabla\psi| + \gamma\nabla g(\vec{x})\nabla\psi \quad (1)$$

Contours encoded as the zero level set of a distance function $\psi(\vec{x}, t)$: points that verify $\psi(\vec{x}, t) = 0$ form the contour. The three terms on the right-hand side of the equation respectively control the expansion or contraction of the contour (velocity c), the smoothness of the contour using the mean curvature κ and the adherence of the contour to the boundary of the object to be segmented. The last term, often called advection term, is specific to the GAC and is responsible for its robustness to gaps in an object's boundary. The parameters α, β and γ allow the user to weight the importance of each term. The spatial function $g(\vec{x})$, often called speed function, is derived from the images to be segmented and contains information about the objects' boundaries. The design of the speed function is crucial for the success of the segmentation. Depending on the specific needs of the application, information on the object's boundary can be based on image gradient, Laplacian or any other relevant feature. The CLS is equivalent to the GAC without the advection term. The omission of the advection term makes the CLS more flexible.

A vertebral body is composed of trabecular bone surrounded by denser cortical bone. Syndesmophytes are made of cortical bone (Fig. 1). To capture those different components, we adopted a multistage strategy in which successive level sets segment the trabecular and cortical bone. Our algorithm is also multiscale. It was originally uniscale [45] but we found that multiscaling made the segmentation not only faster but also more robust and accurate. Our multiscale, multistage, 3D segmentation algorithm is summarized in the flowchart in Fig. 3. We first linearly subsample our data (step 1). Then the original algorithm is applied to the obtained half-scale volume. The preprocessing (described below) determines the parameters of the sigmoid used to compute the speed function of the first GAC (step 2.1). The first GAC roughly segments the interior of the vertebra (step 2.2). Its seed is the result of a fast marching (FM) stage starting from a seed point roughly placed by the user in the center of the vertebral body and lasting 20 iterations. The second level set, also a GAC, refines this segmentation using a Laplacian convolution of the image as the speed function (step 2.3). The third level set, a CLS, segments the cortical bone (step 2.4). A postprocessing step fills some remaining holes using a dilation followed by an erosion (step 2.5). The resulting segmentation is then supersampled back to full scale (step 3) and refined using a CLS (step 4). A last hole-filling postprocessing is performed (step 5).

The speed function $g(\vec{x})$ should ideally have values close to 1 where there are no boundaries (so that the level set can expand rapidly) and values close to 0 where boundaries are present (so that the level set stops). This can be achieved for instance by writing [46]:

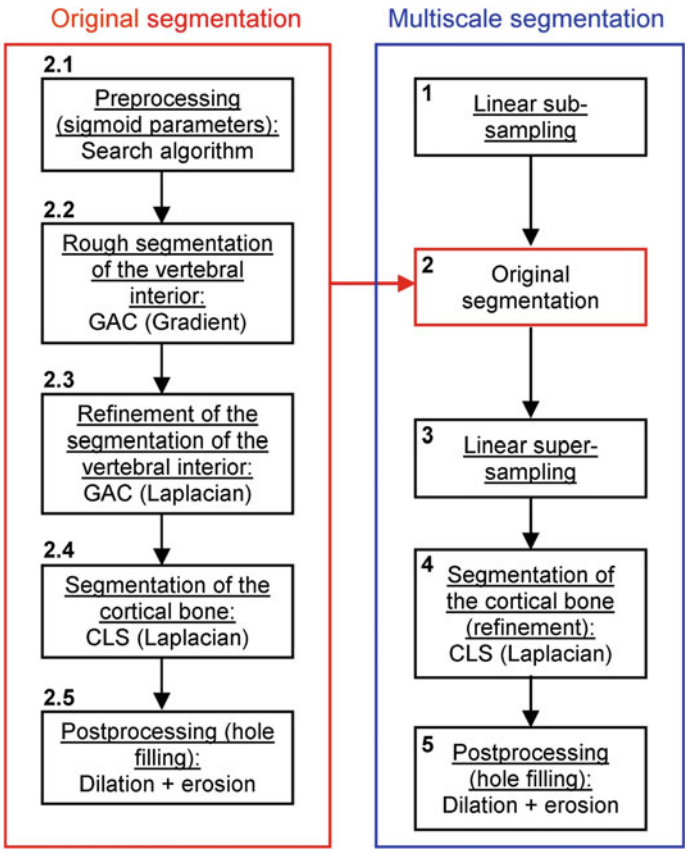


Fig. 3 Flowchart of the algorithm for segmenting vertebral bodies

$$g(\vec{x}) = 1 - \frac{1}{1 + \exp\left(-\frac{I-\xi}{\eta}\right)} \tag{2}$$

where I is the gradient magnitude of the grey level image at voxel $\vec{x}$. The two parameters ξ and η are typically computed using the equations [46]:

$$\eta = \frac{K_1 - K_2}{6} \qquad \xi = \frac{K_1 + K_2}{2} \tag{3}$$

where K_1 is the minimum gradient magnitude value along the object's boundary and K_2 the average gradient magnitude inside the object where the level set is initialized. Those definitions ensure that the level set advances over internal gradients but stops at the minimum gradient along the boundary, as Eq. (2) maps gradients values up to K_2 to approximately 1 and gradient values equal or larger

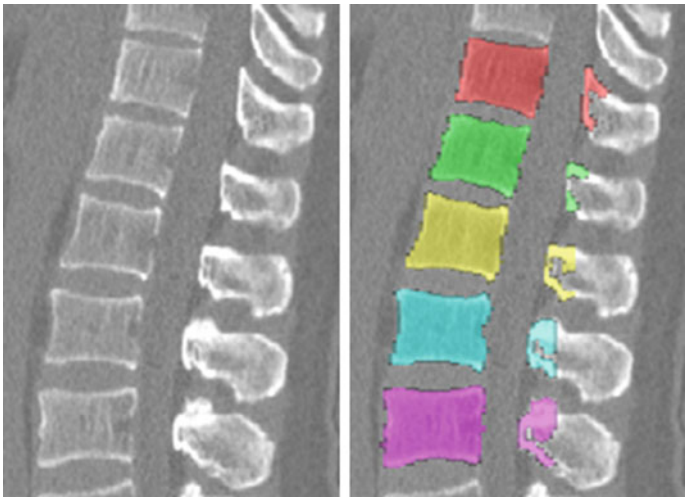


Fig. 4 Example of vertebral body segmentation (original image on the *left*, segmentation results on the *right*)

than K_1 to approximately 0. K_2 can be evaluated as the mean gradient magnitude inside a neighborhood around the seed placed by the user in the center of the vertebral body. K_1 can be determined by a search algorithm. Along lines originating from the center of the vertebral body, the maximal gradient magnitude is considered as belonging to the object's boundary and is recorded. The mean of the 10 % lowest recorded values constitutes our estimate for K_1 [32]. The optimal values for parameters α, β and γ were determined experimentally [32]. Figure 4 shows an example of segmentation obtained by the algorithm.

2.2 Segmentation of the Vertebral Body Ridgelines

The segmentation of vertebral body ridgelines is a preliminary step to both the registration stage (Sect. 2.3) and the syndesmophyte extraction stage (Sect. 2.4). The vertebral body ridgelines provide the landmarks that aid the registration process and the reference level from which syndesmophytes are cut. We extract the ridgelines from the triangular meshes representing the surfaces of the vertebrae using the same level set as Eq. (1), but transposed from the Cartesian domain of rectangular grids to the domain of a surface mesh. While in the usual image grids of CT scans the relevant features are grey level gradients, on a surface mesh, the useful features are curvature measures (the vertebral body surface is more curved at the ridgelines than on the end plates). The curvature measure we used is curvedness (C) [47]:

$$C = \sqrt{\frac{\kappa_1^2 + \kappa_2^2}{2}} \tag{4}$$

where κ_1 and κ_2 are the principal curvatures. Curvedness is a local measure that can be computed at each vertex on the mesh. The larger C is, the more curved the local surface is. The speed function, constructed using Eqs. 2 and 3 but with curvedness replacing grey level gradients, ensures that the level set contour expands in the center of the end plates (low curvedness) and stops at the ridgelines (high curvedness) [32].

The level set evolution equation (Eq. 1) can be implemented on a mesh with two important adjustments relative to level sets in rectangular grids: (1) Gradients and curvatures have to be computed in local coordinate systems defined around each vertex as small enough neighborhoods can reasonably be considered planar. (2) Gradients and curvatures have to be computed using least square estimation methods rather than finite differences [32].

We use the following definitions and notations for level sets on mesh. A function $f(V)$ defined on a mesh associates to each vertex V the quantity $f(V)$. A vertex V is defined by its three coordinates (x, y, z) which can be relative to a global or a local orthonormal frame. V can therefore also be seen as a vector. By immediate neighbor of vertex V, we mean a vertex linked to V by an edge. The 1-ring neighborhood of V is the set of immediate neighbors of V. The 2-ring neighborhood of V consists of its 1-ring neighborhood and all the immediate neighbors of the vertices in the 1-ring neighborhood. The process can be iterated. Thus, the n-ring neighborhood of V is comprised of its (n − 1)-ring neighborhood and all the immediate neighbors of the vertices in the (n − 1)-ring neighborhood.

To implement Eq. 1, the gradients of the distance function ψ and the speed function g have to be evaluated. We do this locally on the mesh in a 1-ring neighborhood around each vertex. The components of $\nabla f(V)$ the gradient of any function f at vertex V can be evaluated by minimizing:

$$E = \sum_{i=1}^{N} \left(\nabla f(V) \cdot \vec{n}_i - \nabla f(V)_i\right)^2 \tag{5}$$

The summation is over the N immediate neighbors of V. The ith neighbor V_i of V defines the unit directional vector $\vec{n}_i$:

$$\vec{n}_i = \frac{V_i - V}{|V_i - V|} \tag{6}$$

The quantity $\nabla f(V)_i$ is the finite difference of function f in the direction of the ith neighbor V_i:

$$\nabla f(V)_i = \frac{f(V_i) - f(V)}{|V_i - V|} \tag{7}$$

The vector components in Eqs. (5)–(7) are relative to local orthonormal frames defined at each vertex V. The gradients $\nabla\psi$ and ∇g in Eq. (1) are computed using Eqs. (5)–(7). $\nabla\psi$ is then used to compute the mean curvature κ.

The gradient $\nabla\psi$ is used to form two functions on the mesh: $\psi_x(V)$ and $\psi_y(V)$, which respectively associate the x and y components of $\nabla\psi$ to each vertex V. We can then evaluate the gradients of $\psi_x(V)$ and $\psi_y(V)$ using Eqs. (5)–(7), which in turn yields ψ_{xx}, ψ_{xy} and ψ_{yy}. Those are the quantities necessary to compute the mean curvature κ of the distance function ψ:

$$\kappa = \frac{\psi_{xx}\psi_y^2 - 2\psi_x\psi_y\psi_{xy} + \psi_{yy}\psi_x^2}{\left(\psi_x^2 + \psi_y^2\right)^{\frac{3}{2}}} \tag{8}$$

The seeding for the mesh level set is also derived from the user placed seed for the vertebral body segmentation. From that seed (roughly in the center of the vertebral body), a vertical line cuts the upper and lower end plates in two points. Those points are used as the seeds for the mesh level sets on the upper and lower end plates. An alternative seeding technique without user input and that relies on the clustering of vertices with low curvedness has recently been proposed [48]. Figure 5 shows an example of contour evolution on the upper end plate of a vertebra. Figure 6 shows several examples of final segmentation results.

2.3 Vertebral Body Registration

Ideally, ridgelines detected on different scans of the same vertebra should be located at identical positions at the junction where the syndesmophytes merge with the end plates. In reality, those positions can be subject to variations, especially for syndesmophytes that do not grow at a right angle in respect to the end plate but laterally and merge with the end plate in a smooth gradual junction. In such cases, the curvature at the junction can be low and the level set might stop at the syndesmophyte's base or slightly leak into the syndesmophyte depending on differing image resolution, sharpness or noise. Figure 7 shows such a discrepancy between baseline and year 1 ridgelines, with a small leak at year 1 (red arrow). Bone above the ridgeline will be labeled as syndesmophyte. If the syndesmophytes at baseline and year 1 were cut from their respective ridgelines, the leak in the year 1 syndesmophyte would cause a deficit in volume compared to baseline. This difference would not be due to real syndesmophyte change. Because real growth may be small, it is important to reduce the error coming from ridgeline discrepancies. We use registration to correct such inconsistencies. Registration aligns the vertebral

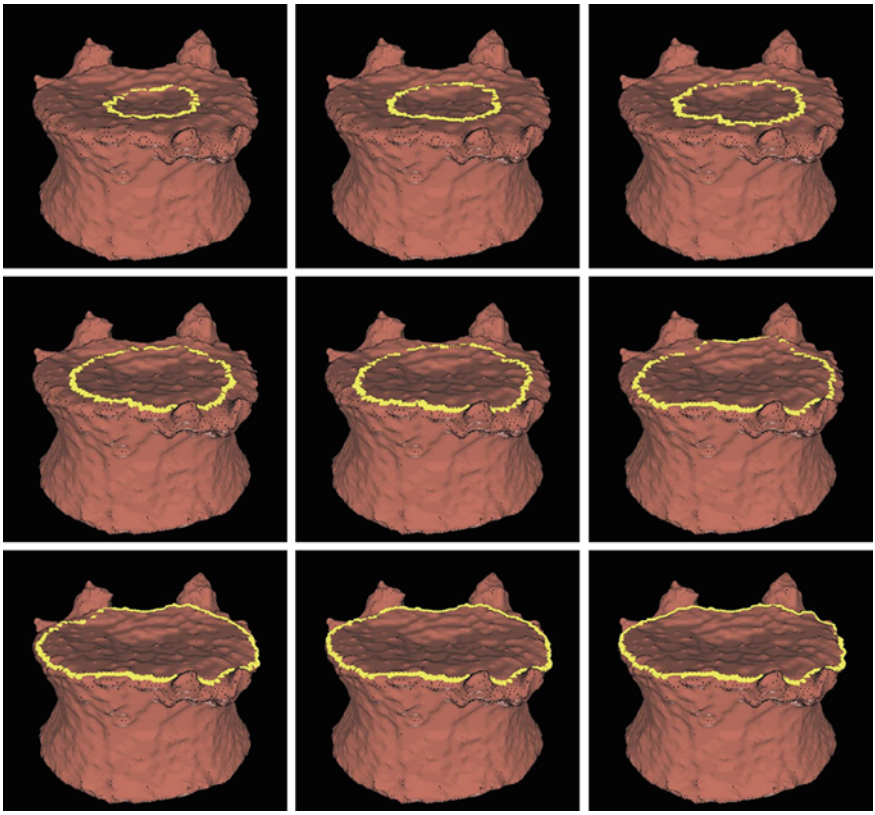


Fig. 5 Example of level set evolution on a mesh

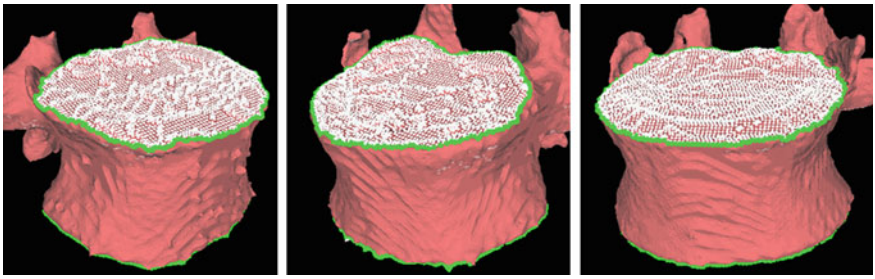


Fig. 6 Examples of end plate (*white*) and ridgeline (*green*) segmentation

bodies of scans (middle of Fig. 2). Once the vertebral bodies are registered, either of the two ridgelines can be used. The important point is to use only one of the ridgelines so that the same syndesmophyte is cut from the exactly the same level on two scans.

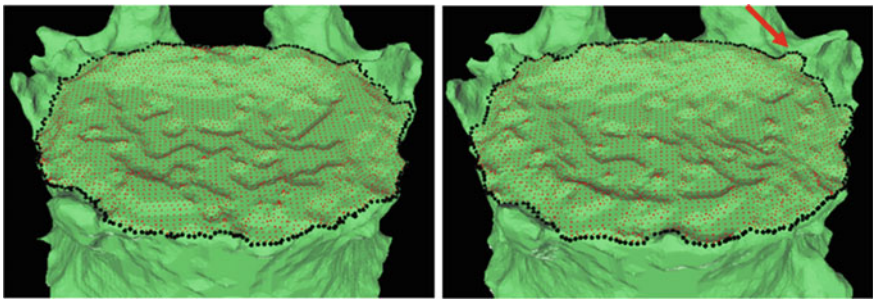


Fig. 7 End plate (*red*) and ridgeline (*black*) segmentation at baseline (*left*) and year 1 (*right*)

We used the iterative closest point (ICP) algorithm to register the surfaces of the vertebrae segmented at baseline and year 1. Given 2 sets of points, the ICP algorithm finds the rigid transformation that minimizes the mean square distance between them [49–51]. We added landmark matching to address the problem of entrapment in local minima. Our ICP algorithm is performed successively on the ridgelines, end plates and the complete surface, the result of each stage serving as the initialization for the following stage [52]. Figure 8 shows some examples of registration results.

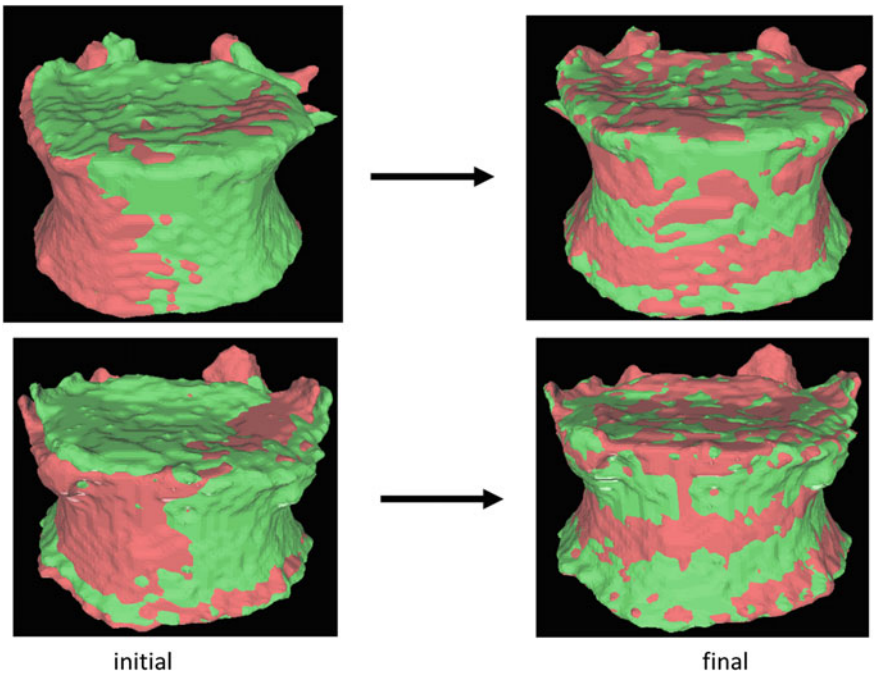


Fig. 8 Two examples of vertebral surface registration

2.4 Syndesmophyte Segmentation

Once corresponding pairs of vertebrae are registered, syndesmophytes can be cut from the vertebral bodies using the ridgeline of the baseline vertebra (using the year 1 or 2 ridgelines is also possible). The algorithm identifies syndesmophytes in each IDS unit. The cutting algorithm marks as syndesmophyte bone voxels lying between the two end plates that bound each IDS. Because of the high precision required by our application, we found it necessary to operate this cutting with subvoxel accuracy. We also address the problem of differing degrees of smoothness in the reconstructions and partial volume effect, and refine the segmentation of syndesmophytes [33].

2.4.1 Syndesmophyte Cutting

Each IDS is bounded by the lower end plate of the superior vertebra, that we note EP1, and the upper end plate of the inferior vertebra, noted EP2. The corresponding ridgelines are respectively noted RL1 and RL2. The cutting algorithm marks as syndesmophyte those previously segmented voxels that are between those 2 end plates. Each candidate voxel is considered in relation to the local ridgelines. If it is below the local level of EP 1/RL1 and above the local level of EP2/RL2 it is marked as syndesmophyte.

However the representation of a continuous space by discrete voxels can introduce inaccuracies in this algorithm. In the first version of our algorithm, a whole voxel was considered either totally above or below the local ridgeline level [32]. However, in reality, most voxels close to the ridgeline level are neither completely above nor completely below that level. Rather, part of the voxel is above while the other part is below. The following algorithm achieves syndesmophyte cutting with subvoxel accuracy. We show how to determine the proportion of a voxel above the local level of EP2/RL2. Determining the proportion of a voxel below the local level of EP1/RL1 is straightforwardly similar.

First we extract the normal to the end plate EP2, $\vec{N}$, using a least square estimate method [32]. Let V be a voxel under consideration. We determine the local ridgeline/end plate level in the following way. The point of RL2 closest to V is found. Neighboring points of EP2/RL2 are averaged to form the point R_V, which, as an average, is an estimate more robust to noise. $\vec{N}$ and R_V define a plane P (orthogonal to $\vec{N}$ and containing R_V), that can be used to cut syndesmophyte from vertebral body. We now determine the position of V relative to this plane. V is a rectangle defined by 8 vertices V_i with $i \in \{1, \ldots, 8\}$. The sign of the scalar product:

$$s(V_i) = sign(\overrightarrow{R_V V_i} \cdot \vec{N}) \tag{9}$$

tells us if V_i is above or below the plane P. If all signs are positive or negative, then voxel V is either completely a syndesmophyte voxel or not. If we have a mix, then

V is a partial syndesmophyte voxel. To determine what proportion of V is syndesmophyte, we subdivide V into smaller rectangles. A voxel V of dimensions p_x, p_y and p_z can be subdivided into M^3 equal subvoxels of dimensions $\frac{p_x}{M}$, $\frac{p_y}{M}$ and $\frac{p_z}{M}$. For this, we simply take as locations of the vertices of the new subvoxels the coordinates $\left(i \cdot \frac{p_x}{M}, j \cdot \frac{p_y}{M}, k \cdot \frac{p_z}{M}\right)$ where M is an integer controlling the number of subdivisions and (i, j, k) are integers. The choice $M = 10$, which means each voxel is divided into $M^3 = 1{,}000$ subvoxels, is a good trade-off between computational speed and gain in precision. The better precision results produced by finer subdivisions (larger M) are limited by diminishing returns. Then, for each subvoxel, it is straightforward to determine if it is above or below P using the same scalar product (Eq. 9). However, since we do not want to pursue the subdivision process further, it is not necessary to test all 8 vertices. We only test one, corresponding to the smallest (i, j, k). For every subvoxel of V, if the test is positive in sign we increment N_S that we define as the number of subvoxels of V found to be syndesmophyte (conversely to determine the proportion of a voxel below the local level of EP1/RL1, we would increment when the test is negative in sign). The corresponding partial syndesmophyte volume is:

$$PSV = \frac{N_S}{M^3} \cdot p_x \cdot p_y \cdot p_z \tag{10}$$

Figure 9 illustrates the difference between whole voxel and subvoxel cutting.

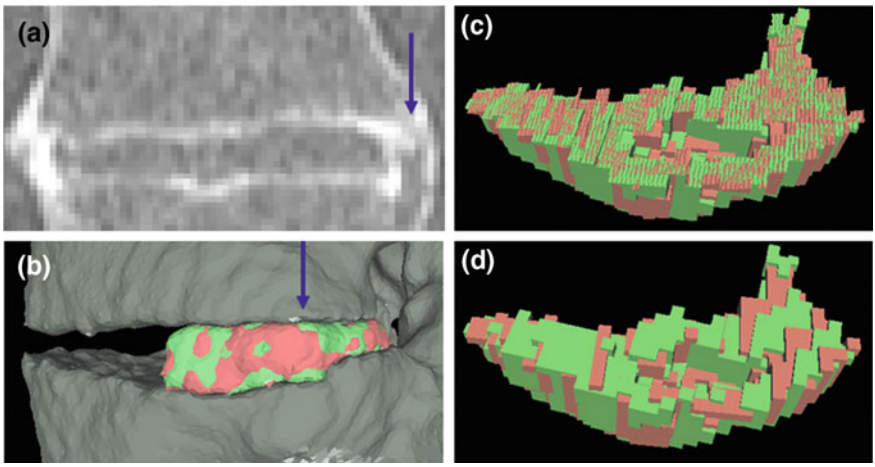


Fig. 9 Comparison between subvoxel and whole voxel cutting. **a** Coronal view of a CT scan of an IDS. **b** Lateral view of the 3D surface reconstruction of the registered *right-hand side* syndesmophytes. View of the registered syndesmophyte upper surfaces after **c** subvoxel and **d** whole voxel cutting from the vertebral body. The view is from the direction of the *blue arrow* in (**a**) and (**b**)

2.4.2 Equalization of Image Smoothness

In a longitudinal study, patients imaged at different times can be scanned using different scanners and/or different scanning parameters. Even when a protocol specifies the scanner and scanning parameters, errors can occur. Images from different scanners and/or with different scanning parameters have different levels of smoothness. The influence of differing degrees of image smoothness on quantitative measurements has been recognized before [53]. In our case, it has an impact on the apparent size of the syndesmophytes. In general, the smoother an image is, the larger the syndesmophyte will appear. To compensate for this effect we devised an algorithm for harmonizing the degree of smoothness of two images. Although we strongly recommend using scanners and scanning parameters in a consistent manner, the ability to compensate for image smoothness differences can allow more flexibility in scanner use when consistency is impractical.

We first devised a measure of image smoothness in a homogeneous region containing only trabecular bone. A region containing both trabecular and cortical bone could produce misleading results. For each voxel in homogeneous region, a mean difference with its neighbors is computed. All those voxel-wise differences are then averaged across the region. This measure can be written:

$$S = \sum_{j=1}^{M} \sum_{i=1}^{N_j} \frac{|GL_j - GL_i|}{MN_j} \tag{11}$$

where GL_j is the grey level of voxel j in the region, GL_i is the grey level of voxel i in the neighborhood of j. M is the total number of voxels in the region. N_j is the total number of neighbors of j that are also in the region. N_j is 26 unless voxel j is at the boundary of the region. To extract a homogeneous region we make use of the segmentations of the vertebral bodies (Sect. 2.1). Eroding those with a structuring element of 5 voxels we obtain homogeneous regions in the trabecular bone. The standard deviation of grey levels in the homogeneous region described can also be used as a smoothness measure. In our experiments, we found that our measure (Eq. 11) performed slightly better in regards to the precision of syndesmophyte volume measurement.

Our procedure for equalizing the smoothness of two images is as follows. We first compute the smoothness measures of the two images. The smoother image has the lower measure, which we call $S_{\min}$. We call the smoothness measure of the other image S. We convolve the least smooth image with Gaussians of increasing standard deviations. We start with a standard deviation of 0.025 mm and increase it by increments of 0.025 mm. After each convolution we compute S. When S becomes smaller than $S_{\min}$ we stop the process. Let us call that measure S_n and the previous one S_{n-1}. We compute the differences $|S_{\min} - S_n|$ and $|S_{\min} - S_{n-1}|$. If the first difference is smaller we use the Gaussian associated with S_n to equalize the smoothness of the two images. Otherwise we use the Gaussian associated with S_{n-1}. Figure 10 shows an example of the procedure. The standard deviation of the

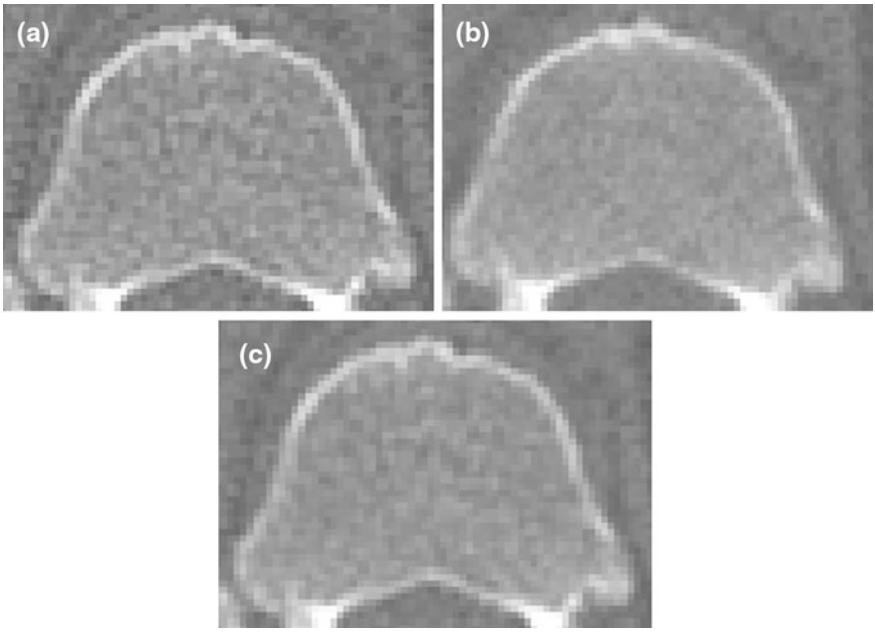


Fig. 10 Smoothness equalization: the least smooth image (**a**) is convolved with a Gaussian (**c**) to match the smoothness level of (**b**)

Gaussian needed to smooth (a) to the level of (b) was 0.125 mm. This stage is a pre-processing step for the following refinement technique.

2.4.3 Density and Laplacian Based Correction

The last step refines the segmentation of the syndesmophytes using the Laplacian filter and gray level density. The output of the Laplacian filter allowed us to pinpoint the boundary between bone and soft tissue. The interface between the two materials can be modeled as a smooth step function. Its Laplacian is positive on one side of the step and negative on the other. The Laplacian divides the interface between 2 materials of different densities with the zero-crossing roughly in the middle. Figure 11 shows an IDS processed with a Laplacian. The color code is green for negative values and red for positive ones. Cortical bone is mainly green. Cortical bone is thin and can be seen as two step functions back to back.

At the boundary between bone and soft tissue, the representation of a continuous space by discrete voxels leads to the creation of voxels containing both materials, a phenomenon usually called partial volume effect. Our algorithm incorporates partial voxels, assigning them a partial volume value depending on their "density", that is, their grey level intensity. The density criterion is obtained in the following manner. From the initial rough syndesmophyte segmentation we estimate the mean voxel

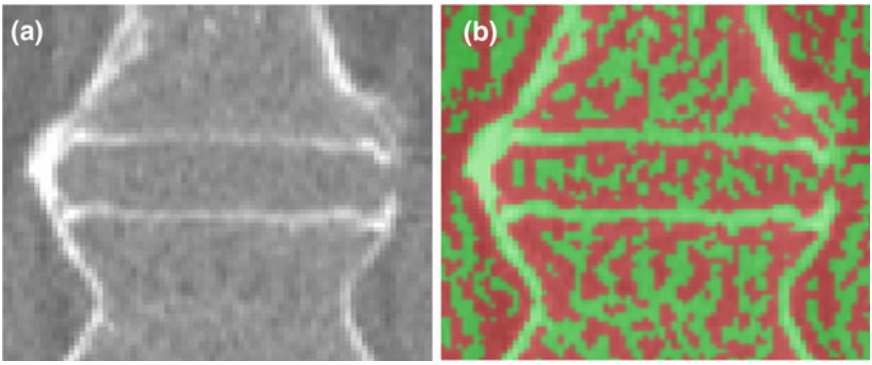


Fig. 11 Effect of the Laplacian filter on an intervertebral disk space: **a** Original image. **b** Laplacian of image (**a**), color-coded with *green* (negative values) and *red* (positive values)

intensity for syndesmophyte, GL_S. Syndesmophytes are surrounded by soft tissue. Considering the neighbors of syndesmophyte voxels we mark the first soft tissue layer, T_1 and second soft tissue layer T_2. From those layers (T_1 and T_2) we extract the mean voxel intensity for soft tissue, GL_T. For a voxel i labeled as syndesmophyte or belonging to T_1 or T_2, our density criterion is based on the measure D_i defined as:

$$D_i = \frac{GL_i - GL_T}{GL_S - GL_T} \tag{12}$$

where GL_i is the grey level of voxel i. The higher the bone content of the voxel, the higher D_i is.

The density and Laplacian criteria are combined in the following manner:

(a) First we consider all syndesmophyte, T_1 and T_2 voxels. If a voxel i (syndesmophyte, T_1 or T_2) verifies the conditions:

$$D_i > D_1 \quad and \quad L_i < 0 \tag{13}$$

where D_1 is a threshold, it is classified as syndesmophyte (L_i is the Laplacian at voxel i). Otherwise it is labeled as soft tissue. This first step mainly corrects leaks. An example is shown in Fig. 12.

(b) The labeling of soft tissue layers T_1 and T_2 and the computing of GL_S and GL_T are updated based on the new more accurate segmentation resulting from step (a). We then process the first soft tissue layer T_1. If a voxel i of T_1 verifies the conditions:

$$D_i > D_2 \quad and \quad L_i < 0 \tag{14}$$

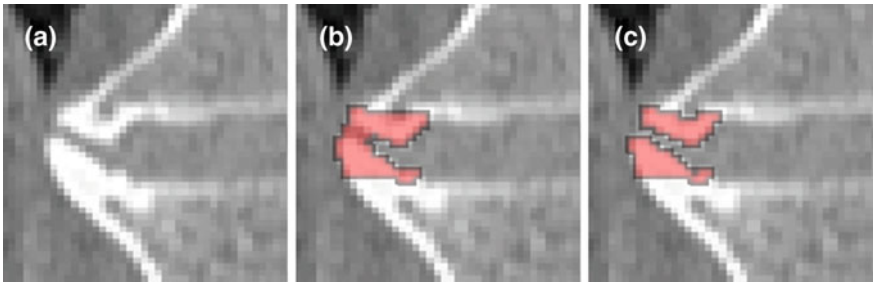


Fig. 12 First stage of the syndesmophyte refinement algorithm: **a** original image, **b** initial segmentation and **c** leak correction

where D_2 is a threshold, it is classified as partial syndesmophyte with proportion of bone corresponding to D_i. This second step adds a layer at the bone/soft tissue boundary where, due to partial volume effect, voxels are likely to contain both types of tissues.

The thresholds D_1 and D_2 control how selective the algorithm is in admitting syndesmophyte voxels. They can be used to add partial bone voxels that were not segmented or exclude soft tissue voxels that were mistakenly labeled as syndesmophyte. Both thresholds can be set between 0 and 1. Lower thresholds are more permissive in syndesmophyte selection. Extensive experimentation led us adopt the set of threshold (0.8, 0.2) for D_1 and D_2 respectively [33].

3 Accuracy and Precision of the Algorithm

3.1 Accuracy and Validity

As an accuracy test, we compared manually and automatically segmented syndesmophytes [33]. Patients were scanned on either a Philips Brilliance 64 or a GE Lightspeed Ultra. For both scanners, voltage and current parameters were 120 kVp and 300 mAs, respectively. Slice thicknesses were 1.5 and 1.25 mm, respectively, for the Philips and GE. Spacing between slices was 0.7 and 0.625 mm for the Philips and GE respectively. Each patient was scanned from the middle of the T10 vertebra to the middle of the L4 vertebra providing 4 IDSs for analysis (T11/T12, T12/L1, L1/L2 and L2/L3). These scanning parameters were used for all the studies including the reliability and longitudinal studies. Using the ITK-SNAP software [54], one operator manually segmented syndesmophytes in two IDSs (L1/L2 and L2/L3) for 6 patients. The agreement between manually and automatically segmented syndesmophytes was evaluated using the overlap similarity index (OSI), also known as the Dice similarity coefficient [55]:

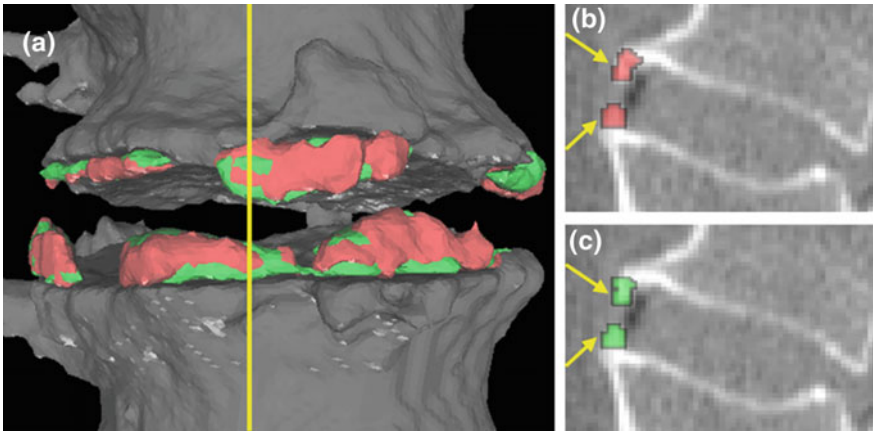


Fig. 13 Comparison between manual (*red*) and automated (*green*) segmentation of syndesmophytes on **a** 3D surface reconstructions, **b**, **c** sagittal slices. The *yellow line* in **a** indicates the position of the sagittal slices. The overlap similarity index in this example is 0.77

$$OSI = \frac{2(V_1 \cap V_2)}{V_1 + V_2} \tag{15}$$

where V_1 and V_2 are the two volumes compared. OSI is always comprised between 0 and 1, with 1 indicating perfect overlap. Out of the 12 IDSs processed, the mean (±std) OSI was 0.76 (±0.06). Considering that syndesmophytes are small objects, an OSI of 0.76 indicates good agreement. Figure 13 shows an example of syndesmophyte segmentations by the manual and automated methods.

In a more extensive validation study involving 38 patients, the syndesmophyte volumes computed by the algorithm were compared with the readings of physicians [56]. Two physicians scored 152 IDSs (4 IDSs per patient) using a 4-point grading system (0 = no syndesmophyte; 1 = small isolated syndesmophytes involving less than a quarter of the vertebral rim and no bridging; 2 = syndesmophyte involving more than a quarter of the vertebral rim or focal bridging; 3 = bridging involving more than a quarter of the vertebral rim). The physicians examined the IDSs in the axial, coronal and sagittal views of the CT reconstructions. Figure 14 shows the association of computed volumes with the physicians' ratings. Volumes computed by the algorithm increased with the readers' scores ($p < 0.0001$ using a stratified Kruskal–Wallis trend test accounting for non-independence of observations within patients [57]).

3.2 Reliability/Precision

The precision of the algorithm was evaluated by comparing the results of 2 scans performed on the same day in 9 patients [56]. The protocol was approved by the institutional review board and all subjects provided written informed consent. After

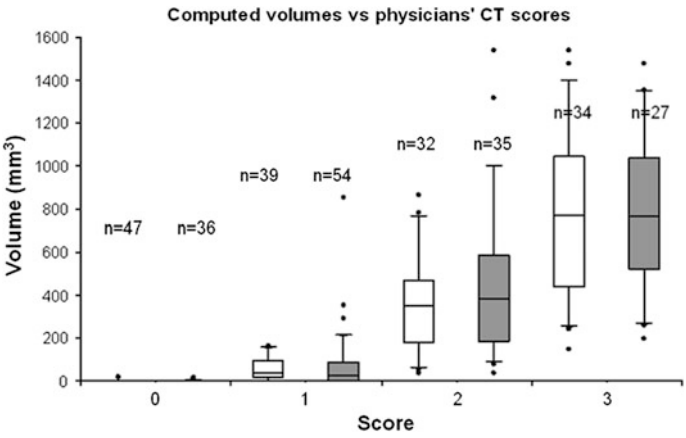


Fig. 14 Boxplots of computed syndesmophyte volume and height by physicians' scores (*white* for one reader, *grey* for the other). N is the number of intervertebral disc spaces

the first scan, patients stood up before lying down again for the second scan. This ensured that they did not lie in exactly the same position and that the variation was in the range expected for patients in a longitudinal study. That enabled us to include the variability originating from CT artifacts such as beam hardening [58]. Syndesmophyte volumes from the 4 IDSs were added to form a total per patient.

Various measures of reliability were computed (Table 1). The mean (±std) difference between the two scans, 18.3 (±19.6) mm^3, only represents 1.31 % of the total mean syndesmophyte volume, 1,396 (±1,564) mm^3. The intraclass correlation coefficient (ICC) was very high. The coefficient of variation (CV) was estimated according to the guidelines of Gluer et al. [59]. Bland-Altman analysis was used to determine the 95 % limits of agreement [60]. Volume measures were heteroskedastic, with larger inter-scan differences for larger syndesmophyte volumes. Bland–Altman analysis was therefore performed on log-transformed values, and the

Table 1 Reliability/precision of computed syndesmophytes volumes

	Syndesmophyte volumes	
	1st scan	2nd scan
Min (mm^3)	55.4	55.5
Max (mm^3)	4,333	4,292
Mean ± std (mm^3)	1,396 ± 1,564	1,404 ± 1,564
	Reliability measures	
Mean ± std of difference (mm^3)	18.3 ± 19.6	
ICC	0.99	
CV (%)	1.31	
95 % limits of agreement (%)	[−0.30, 0.30]	

95 % limits of agreement for volume were in terms of percentage [61]. Using this method it was found that an increase in syndesmophyte volume of more than 3 % represented a change greater than measurement error.

4 Longitudinal Study

For this study, we performed lumbar spine CT scans on 33 patients at baseline, year 1 and year 2 [62]. The same 4 IDSs as in the precision study were processed. Radiographs of these 4 IDSs were also scored by a physician using mSASSS but

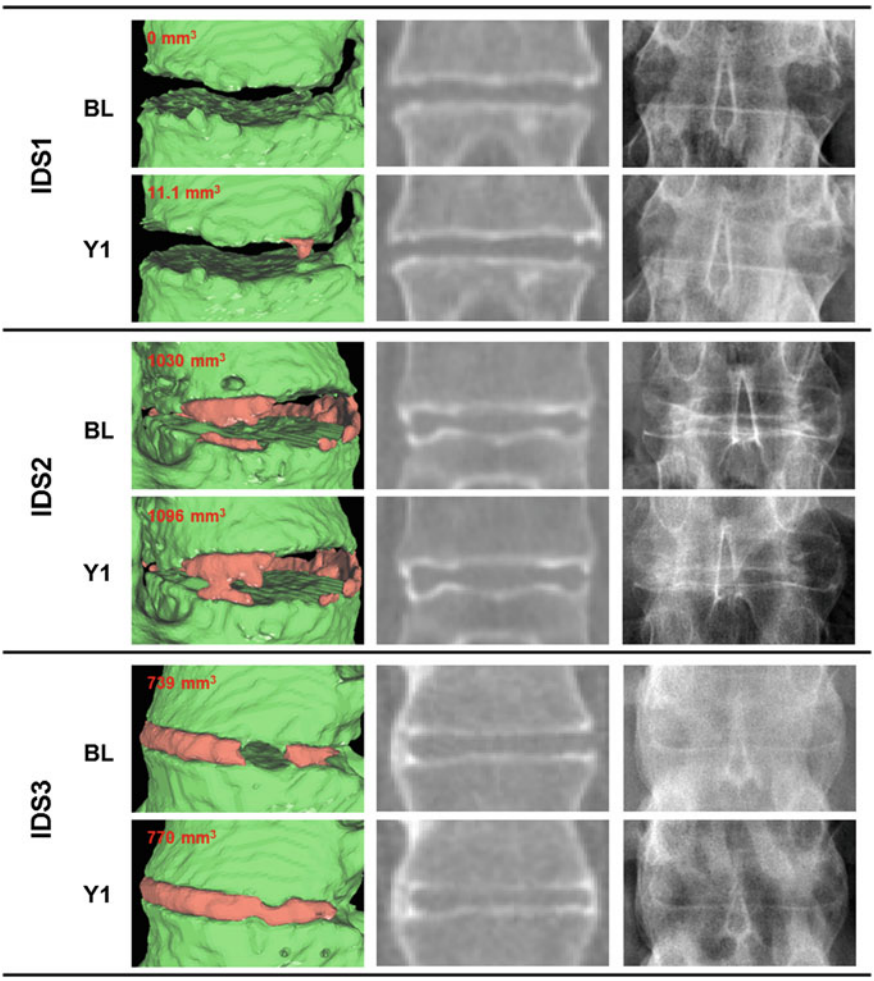


Fig. 15 Examples of syndesmophyte progression from baseline (BL) to year 1 (Y1). From *left* to *right* 3D surface mesh (syndesmophytes in red and vertebral bodies in green), CT slice, radiograph

without the score of 1 which does not represent syndesmophyte growth. The scores were: 0 = no syndesmophyte; 2 = syndesmophyte but not complete bridging; 3 = bridging. Results from the 4 IDSs were added. Figures 15 and 16 show examples of syndesmophyte progression detected by the algorithm but not visible on radiographs from baseline to year 1 and 2 respectively.

The mean (±std) computed syndesmophyte volume change was 87 (±186) mm^3 at year 1 and 201 (±366) mm^3 at year 2, which respectively represents an increase of about 8 and 18 % in respect to the mean baseline volume. At year 1 and 2, respectively 24 (73 %) and 26 (79 %) patients had a volume increase. By contrast,

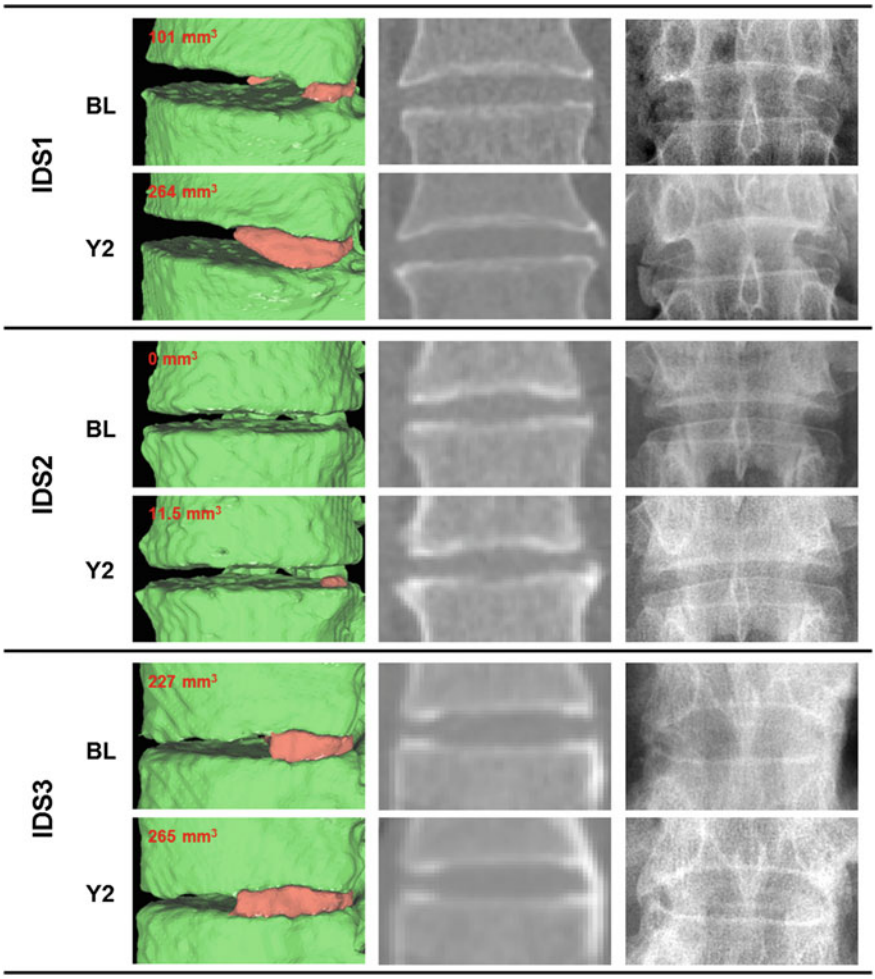


Fig. 16 Examples of syndesmophyte progression from baseline (BL) to year 2 (Y2). From *left* to *right* 3D surface mesh (syndesmophytes in *red* and vertebral bodies in *green*), CT slice, radiograph

Table 2 Change in syndesmophyte volume (CT) and mSASSS (radiography)

		CT	Radiography
	Mean (±std) at baseline	1,095 (±1,278) mm^3	4.2 (±5.6)
Baseline to year 1	Number of patients with change > 0	24 (73 %)	4 (12 %)
	Mean (±std) change	87.0 (±186) mm^3	0.24 (±0.97)
Baseline to year 2	Number of patients with change > 0	26 (79 %)	4 (12 %)
	Mean (±std) change	201 (±366) mm^3	0.30 (±1.4)

only 4 (12 %) had a mSASSS increase at year 1 and 2 (Table 2). From baseline to year 1, 18 patients (55 %) had an increase larger than 3 %, the 95 % limit of agreement derived from Bland-Altman analysis in the reliability study. From baseline to year 2, 23 patients (70 %) had an increase larger than 3 %. Additionally, two patients in whom the algorithm detected no syndesmophytes in all 4 IDSs at baseline developed new syndesmophytes at year 1, and three patients did so at year 2. For these patients, the rate of change cannot be computed because their baseline was 0.

Figure 17 shows the cumulative probability plots for computed volume changes and mSASSS changes. The curves for computed volumes show the progressivity of the disease. The curves for year 1 and 2 are clearly distinguishable and syndesmophyte volume changes are larger for year 2 than for year 1. By contrast, for mSASSS the two curves are nearly identical and both mostly located at zero.

5 Discussion and Future Challenges

The algorithm is still new and has so far been validated on a relatively small numbers of patients. More extensive work is needed to establish the method. The method still requires an operator to place a seed to initiate the segmentation. Automation of this task should be explored. The algorithm also requires high resolution especially in the z direction (slice thickness of 1.5 mm and spacing between slices of 0.7 mm). Additional work is needed to adapt the method to more common lower resolution scans. It is probable that lower resolution will entail lower precision in volume measurements.

Registration makes the choice of the ridgeline (baseline, year 1 or year 2) unimportant. In our work, we chose the baseline ridgeline as the reference. Averaging ridgelines may be advantageous, since an average is generally more robust to errors. Although registration will ensure that the same errors are made for the scans to be compared and will therefore not impact the computed syndesmophyte volume differences, it is always benefic to start with the most accurate ridgeline. Many methods can be proposed to define the average of 2 or more curves. In our case

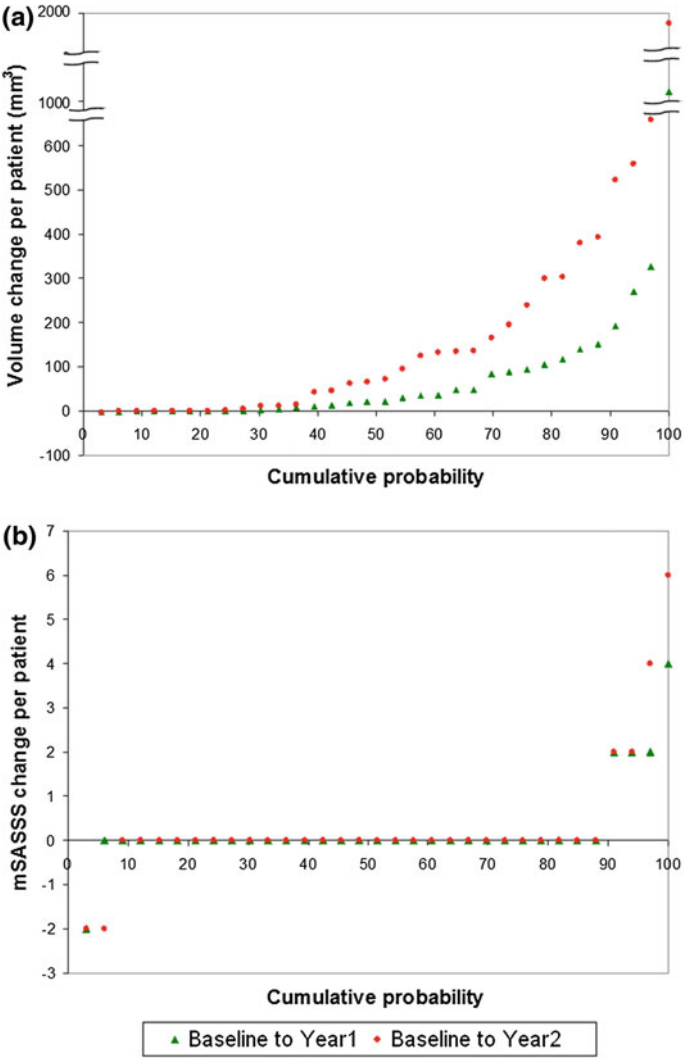


Fig. 17 Cumulative probability plots for changes in **a** syndesmophyte computed volume, **b** modified stoke AS spine score (mSASSS)

where the ridgelines (same vertebral body at different times) should be fairly similar, averaging can for instance be done in the direction normal to the curves. Starting from the baseline curve, for each curve point, a local normal direction can be estimated. On the year 1 or 2 curve, the curve point most aligned with that normal direction can be determined and averaged with the curve point on the baseline curve.

The higher sensitivity to change of our computed volumes reflects both the fully quantitative nature of the method and the improved visualisation of syndesmophytes using CT. Exploiting the 3D imaging capability of CT, we were able to quantitate syndesmophytes along the entire vertebral body rim. It has been suggested that magnetic resonance imaging (MRI) could also be used to image the spine tomographically and with less exposure for patients. Several rating systems for structural chronic damage in AS have been proposed [63–66]. However, few longitudinal studies tracking syndesmophyte growth in MRI have been published [67]. In [62], it was found that computed volumes in CT were much more sensitive to change than MRI readings. Cortical bone is poorly visualised on MRI because its water content is similar to the water content of surrounding tissues. Scoring systems based on MRI are semiquantitative, which also may limit their sensitivity to change. In addition, higher resolution can be achieved in CT and long acquisition time for MRI causes motion artefacts.

A major criticism of the work has centered on the radiation exposure associated with a CT scan. With the protocol used in the study, patients received an average radiation dose of 8.01 mSv compared with 2.59 mSv for lateral radiographs of the cervical and lumbar spine (as would be used in a complete mSASSS assessment) [56, 62]. However, the question of radiation exposure has to be considered in close relation with the information obtained. Although the radiation exposure of CT is substantially higher than the radiation exposure of radiographs, each CT scan provides complete information on syndesmophytes, and, in our study, none needed to be discarded because of poor visualisation. The advantages of low radiation exposure need to be weighed by the usefulness of the information gathered by that exposure. It should be stressed that scanner technology is improving fast and, with the introduction of iterative reconstruction, dose reduction of 50 % or more has been achieved with minimal loss in image quality [68, 69]. The reliability of the algorithm has to be evaluated using such dose saving methods.

Because the algorithm visualizes and quantitates syndesmophytes in their real 3D environment for the first time, it opens the door to new research possibilities. For instance, the distribution of the syndesmophytes around the rim of the vertebral end plates, if not random, could shed some light on the drivers of osteoproliferation, which are still unknown. The testing of drugs that can potentially halt or slow syndesmophyte progression will benefit from the greatly improved sensitivity and reliability of the new method. Similarly, studies that seek to associate syndesmophyte progression with gene expression, biomarkers and lifestyle risk factors (such as smoking or lack of exercise for example) should use a method that can capture syndesmophytes in their totality and quantitatively.

6 Conclusion

To improve the low sensitivity to change associated radiographic reading, we have designed a quantitative measurement of syndesmophytes in CT scans. The method has very good reliability. In a 2-year longitudinal study, the algorithm could detect

syndesmophyte growth in 79 % of the patients compared to only 12 % for radiographic reading. The mean 2-year change represented a 18 % increase in syndesmophyte volume in respect to the baseline volume. This method holds promise for longitudinal clinical studies that need to track syndesmophyte growth.

References

1. Sieper J, Braun J, Rudwaleit M et al (2002) Ankylosing spondylitis: an overview. Ann Rheum Dis 61(Suppl III):iii8–iii18
2. Braun J, Sieper J (2007) Ankylosing spondylitis. Lancet 369:1379–1390
3. Braun J, Brandt J, Listing J et al (2002) Treatment of active ankylosing spondylitis with infliximab: a randomised controlled multicentre trial. Lancet 359:1187–1193
4. Davis JC, van der Heijde D, Braun J et al (2003) Recombinant human tumor necrosis factor receptor (etanercept) for treating ankylosing spondylitis: a randomized, controlled trial. Arthritis Rheum 48(11):3230–3236
5. van der Heijde D, Kivitz A, Schiff MH et al (2006) Efficacy and safety of adalimumab in patients with ankylosing spondylitis: results of a multicenter, randomized, double-blind, placebo-controlled trial. Arthritis Rheum 54(7):2136–2146
6. Baraliakos X, Listing J, Rudwaleit M et al (2005) Radiographic progression in patients with ankylosing spondylitis after 2 years of treatment with the tumour necrosis factor alpha antibody infliximab. Ann Rheum Dis 64:1462–1466
7. Baraliakos X, Listing J, Brandt J et al (2007) Radiographic progression in patients with ankylosing spondylitis after 4 yrs of treatment with the anti-TNF-alpha antibody infliximab. Rheumatology 46:1450–1453
8. van der Heijde D, Landewé R, Baraliakos X et al (2008) Radiographic findings following two years of infliximab therapy in patients with ankylosing spondylitis. Arthritis Rheum 58:3063–3070
9. van der Heijde D, Landewé R, Einstein S et al (2008) Radiographic progression of ankylosing spondylitis after up to two years of treatment with etanercept. Arthritis Rheum 58:1324–1331
10. van der Heijde D, Salonen D, Weissman BN et al (2009) Assessment of radiographic progression in the spines of patients with ankylosing spondylitis treated with adalimumab for up to 2 years. Arthritis Res Ther 11:R127
11. Baraliakos X, Listing J, Rudwaleit M et al (2008) The relationship between inflammation and new bone formation in patients with ankylosing spondylitis. Arthritis Res Ther 10:R104
12. Bennett AN, McGonagle D, O'Connor P et al (2008) Severity of baseline magnetic resonance imaging-evident sacroiliitis and HLA–B27 status in early inflammatory back pain predict radiographically evident ankylosing spondylitis at eight years. Arthritis Rheum 58(11): 3413–3418
13. Maksymowych WP, Chiowchanwisawakit P, Clare T et al (2009) Inflammatory lesions of the spine on magnetic resonance imaging predict the development of new syndesmophytes in ankylosing spondylitis: evidence of a relationship between inflammation and new bone formation. Arthritis Rheum 60:93–102
14. Chiowchanwisawakit P, Lambert RG, Conner-Spady B et al (2011) Focal fat lesions at vertebral corners on magnetic resonance imaging predict the development of new syndesmophytes in ankylosing spondylitis. Arthritis Rheum 63:2215–2225
15. Pedersen SJ, Chiowchanwisawakit P, Lambert RG et al (2011) Resolution of inflammation following treatment of ankylosing spondylitis is associated with new bone formation. J Rheumatol 38:1349–1354

16. van der Heijde D, Machado P, Braun J et al (2012) MRI inflammation at the vertebral unit only marginally predicts new syndesmophyte formation: a multilevel analysis in patients with ankylosing spondylitis. Ann Rheum Dis 71:369–373
17. Lories RJ, Dougados M (2012) Inflammation and ankylosis: still an enigmatic relationship in spondyloarthritis. Ann Rheum Dis 71:317–318
18. Maksymowych WP, Morency N, Conner-Spady B et al (2013) Suppression of inflammation and effects on new bone formation in ankylosing spondylitis: evidence for a window of opportunity in disease modification. Ann Rheum Dis 72:23–28
19. Maksymowych WP, Landewé R, Conner-Spady B et al (2007) Serum matrix metalloproteinase 3 is an independent predictor of structural damage progression in patients with ankylosing spondylitis. Arthritis Rheum 56:1846–1853
20. Pedersen SJ, Sørensen IJ, Lambert RGW et al (2011) Radiographic progression is associated with resolution of systemic inflammation in patients with axial spondylarthritis treated with tumor necrosis factor alpha inhibitors: a study of radiographic progression, inflammation on magnetic resonance imaging, and circulating biomarkers of inflammation, angiogenesis, and cartilage and bone turnover. Arthritis Rheum 63(12):3789–3800
21. van Tubergen A, Ramiro S, van der Heijde D et al (2012) Development of new syndesmophytes and bridges in ankylosing spondylitis and their predictors: a longitudinal study. Ann Rheum Dis 71:518–523
22. Poddubnyy Denis, Haibel Hildrun, Listing Joachim et al (2012) Baseline radiographic damage, elevated acute-phase reactant levels, and cigarette smoking status predict spinal radiographic progression in early axial spondylarthritis. Arthritis Rheum 64(5):1388–1398
23. Ward MM, Hendrey MR, Malley JD et al (2009) Clinical and immunogenetic prognostic factors for radiographic severity in ankylosing spondylitis. Arthritis Rheum 61:859–866
24. Appel H, Ruiz-Heiland G, Listing J et al (2009) Altered skeletal expression of sclerostin and its link to radiographic progression in ankylosing spondylitis. Arthritis Rheum 60:3257–3262
25. Diarra D, Stolina M, Polzer K et al (2007) Dickkopf-1 is a master regulator of joint remodeling. Nat Med 13:156–163
26. Daoussis D, Liossis SN, Solomou EE et al (2010) Evidence that Dkk-1 is dysfunctional in ankylosing spondylitis. Arthritis Rheum 62:150–158
27. Ruiz-Heiland G, Appel H, Poddubnyy D et al (2012) High level of functional dickkopf-1 predicts protection from syndesmophyte formation in patients with ankylosing spondylitis. Ann Rheum Dis 71:572–574
28. MacKay K, Mack C, Brophy S, Calin A (1998) The bath ankylosing spondylitis radiology index (BASRI). Anew, validated approach to disease assessment, Arthritis Rheum. 41:2263–2270
29. Creemers MC, Franssen MJ, van't Hof MA et al (2005) Assessment of outcome in ankylosing spondylitis: an extended radiographic scoring system. Ann Rheum Dis 64:127–129
30. Wanders AJ, Landewé RB, Spoorenberg A et al (2004) What is the most appropriate radiologic scoring method for ankylosing spondylitis? A comparison of the available methods based on the Outcome Measures in Rheumatology Clinical Trials filter. Arthritis Rheum 50:2622–2632
31. Spoorenberg A, de Vlam K, van der Linden S et al (2004) Radiological scoring methods in ankylosing spondylitis. Reliability and change over 1 and 2 years. J Rheumatol 31:125–132
32. Tan S, Yao J, Ward MM et al (2008) Computer aided evaluation of ankylosing spondylitis using high-resolution CT. IEEE Trans Med Imaging 27:1252–1267
33. Tan S, Yao J, Yao L et al (2012) Improved precision of syndesmophyte measurement for the evaluation of ankylosing spondylitis using CT: a phantom and patient study. Phys Med Biol 57:4683–4704
34. Lorensen WE, Cline HE (1987) Marching cubes: a high resolution 3D surface construction algorithm. Comput Graph (Proc SIGGRAPH) 21:163–169
35. McInerney T, Terzopoulos D (1999) Topology adaptive deformable surfaces for medical image volume segmentation". IEEE Trans Med Imaging 18(10):840–850
36. Kang Y, Engelke K, Kalender WA (2003) A new accurate and precise 3-D segmentation method for skeletal structures in volumetric CT data. IEEE Trans Med Imaging 22(5):586–598

37. Ghebreab S, Smeulders AWM (2004) Combining strings and necklaces for interactive three-dimensional segmentation of spinal images using an integral deformable spine model. IEEE Trans Biomed Eng 51(10):1821–1829
38. Mastmeyer A, Engelke K, Fuchs C, Kalender WA (2006) A hierarchical 3D segmentation method and the definition of vertebral body coordinate systems for QCT of the lumbar spine. Med Image Anal 10:560–577
39. Yao J, O'Connor SD, Summers RM (2006) Automated spinal column extraction and partitioning. In: Proceedings of the IEEE international symposium on biomedical imaging, pp 390–393
40. Klinder Tobias, Ostermann Jörn, Ehm Matthias et al (2009) Automated model-based vertebra detection, identification, and segmentation in CT images. Med Image Anal 13:471–482
41. Aslan MS, Ali A, Rara H, Farag AA (2010) An automated vertebra identification and segmentation in CT images. In: Proceedings of the IEEE international conference on image processing, pp 233–236
42. Stern Darko, Likar Bostjan, Pernus Franjo et al (2011) Parametric modelling and segmentation of vertebral bodies in 3D CT and MR spine images. Phys Med Biol 56:7505–7522
43. Sethian JA (1999) Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision and materials science. Cambridge University Press, Cambridge, 1999
44. Caselles V, Kimmel R, Sapiro G (1997) Geodesic active contours. Int J Comput Vision 22:61–79
45. Tan S, Yao J, Ward MM et al (2006) Level set based vertebra segmentation algorithm for the evaluation of ankylosing spondylitis. In: Proceedings of the SPIE medical imaging, vol 6144, pp 58–67, San Diego, CA
46. Ibanez L, Schroeder W, Ng L, Cates J (2003) The ITK software guide. Kitware Inc., Clifton Park
47. Koenderink JJ (1990) Solid shape. The MIT Press, Cambridge
48. Tan S, Yao J, Yao L et al (2013) High precision semiautomated computed tomography measurement of lumbar disk and vertebral heights. Med Phys 40:011905
49. Besl PJ, McKay ND (1992) A method for registration of 3D shapes. IEEE Trans Pattern Anal Mach Intell 14:239–256
50. Herring JL, Dawant BM (2001) Automatic lumbar vertebral identification using surface-based registration. J Biomed Inform 34:74–84
51. Tamura Y, Sugano N, Sasama T et al (2005) Surface-based registration accuracy of CT-based image-guided spine surgery. Eur Spine J 14:291–297
52. Tan S, Yao J, Yao L (2008) Summers R M and Ward M M Vertebral surface registration using ridgelines/crestlines. Proc SPIE Med Imaging 6914:69140H
53. Friedman L, Glover GH, Krenz D et al (2006) Reducing inter-scanner variability of activation in a multicenter fMRI study: role of smoothness equalization. Neuroimage 32:1656–1668
54. Yushkevich PA, Piven J, Hazlett HC et al (2006) User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. Neuroimage 31:1116–1128
55. Zou KH, Warfield SK, Bharatha A et al (2004) Statistical validation of image segmentation quality based on a spatial overlap index. Acad Radiol 11:178–189
56. Tan S, Yao J, Flynn JA et al (2014) Quantitative measurement of syndesmophyte volume and height in ankylosing spondylitis using CT. Ann Rheum Dis. 73:544–550
57. Hothorn T, Hornik K, van de Wiel MA et al (2008) Implementing a class of permutation tests: the coin Package. J Stat Softw 28:1–23
58. Barrett JF, Keat N (2004) Artifacts in CT: recognition and avoidance. Radio Graph 24:1679–1691
59. Gluer CC, Blake G, Lu Y et al (1995) Accurate assessment of precision errors: how to measure the reproducibility of bone densitometry techniques. Osteoporos Int 5:262–270
60. Bland JM, Altman DG (1986) Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1:307–310

61. Euser AM, Dekker FW, le Cessie S (2008) A practical approach to Bland-Altman plots and variation coefficients for log transformed variables. J Clin Epidemiol 61:978–982
62. Tan S, Yao J, Flynn JA et al (2013) Quantitative syndesmophyte measurement in ankylosing spondylitis using CT: longitudinal validity and sensitivity to change over 2 years. Ann Rheum Dis. doi:10.1136/annrheumdis-2013-203946 (Published Online First December 2, 2013)
63. Braun J, Baraliakos X, Golder W et al (2004) Analysing chronic spinal changes in ankylosing spondylitis: a systematic comparison of conventional x rays with magnetic resonance imaging using established and new scoring systems. Ann Rheum Dis 63:1046–1055
64. Ostergaard M, Maksymowych WP, Pedersen SJ, et al (2009) Structural lesions detected by magnetic resonance imaging in the spine of patients with spondyloarthritis—definitions, assessment system, and reference image set. J Rheumatol 36(Suppl 84). doi:10.3899/jrheum.090617
65. Chiowchanwisawakit P, Ostergaard M, Pedersen SJ et al (2009) Validation of definitions for structural lesions detected by magnetic resonance imaging in the spine of patients with spondyloarthritis. J Rheumatol 36(Suppl 84). doi:10.3899/jrheum.090619
66. Madsen KB, Jurik AG (2010) MRI grading method for active and chronic spinal changes in spondyloarthritis. Clin Radiol 65:6–14
67. Baraliakos X, Davis J, Tsuji W et al (2005) Magnetic resonance imaging examinations of the spine in patients with ankylosing spondylitis before and after therapy with the tumor necrosis factor alpha receptor fusion protein etanercept. Arthritis Rheum 52(4):1216–1223
68. Hara AK, Paden RJ, Silva AC et al (2009) Iterative reconstruction technique for reducing body radiation dose at CT: feasibility study. AJR Am J Roentgenol 193(3):764–771
69. Winklehner A, Karlo C, Puippe G et al (2011) Raw data-based iterative reconstruction in body CTA: evaluation of radiation dose saving potential. Eur Radiol 21(12):2521–2526

Three-Dimensional Spine Reconstruction from Radiographs

Samuel Kadoury

Abstract For several musculoskeletal pathologies, single radiographic images do not offer the necessary information to portray the actual three-dimensional (3D) representation of the spine in order to assess effects such as intrinsic vertebral rotation, inter-vertebral disc wedging, spine torsion or dislocations. This limits the scope of routine diagnostic, follow-up exams, and treatment planning. Volumetric imaging modalities such as CT or MRI are on the other hand limited due to the fact that they cannot be acquired in the standing position, which is required for evaluation of posture. Biplanar radiography is still the imaging modality that is most frequently used for the 3D clinical assessment of spinal deformities. In this chapter, we present the different techniques involved for obtaining the 3D reconstruction of a spine using biplanar radiographs. First, we present different approaches (linear and non-linear) for calibrating the radiographic scene in order to configure the proper 2D-3D spatial relationship. Once the stereo-radiographic system is calibrated, anatomical landmarks or vertebral shapes constituting the spine can be identified on the radiographic images using manual identification or automated tools. Finally, using these high-level primitives located in an accurate calibrated system, a spine model can be reconstructed in 3D using a number of correspondence methods. For selected applications using reconstructed 3D spine models, we show how these techniques can help to better understand spinal pathologies such as idiopathic scoliosis, which is inherently a three-dimensional deformation of the spine.

S. Kadoury (✉)
Polytechnique Montreal, P.O. Box 6079 Succ. Centre-Ville, Montreal, QC H3C 3A7, Canada
e-mail: samuel.kadoury@polymtl.ca

S. Kadoury
CHU Sainte-Justine Research Center, 3175 Chemin de la Côte-Sainte-Catherine, Montreal, QC, Canada

© Springer International Publishing Switzerland 2015

S. Li and J. Yao (eds.), *Spinal Imaging and Image Analysis*,
Lecture Notes in Computational Vision and Biomechanics 18,
DOI 10.1007/978-3-319-12508-4_6

1 Introduction

Several clinical studies in orthopaedics have used three-dimensional (3D) models of the spine for evaluating pathologies in spinal deformities like adolescent idiopathic scoliosis (AIS). Scoliosis affects 2–3 % of the population. Every year, an estimated 30,000 children are fitted with a brace, while 38,000 patients undergo spinal fusion surgery. The 3D reconstruction of a patient's spine has been extremely useful in the undertaking of several studies such as the 3D evaluation of the immediate effect of the treatment with the Boston brace system [36], pre- and postoperative comparison of spine instrumentation surgery [37] and the 3D progression of scoliosis [61]. Established volumetric modalities such as magnetic resonance imaging (MRI) are very attractive because it is noninvasive for the patient, but it is unfortunately not suitable for a postoperative 3D evaluation due to the ringing artifacts caused by the surgical implants, as well as being quite expensive and time-consuming. On the other hand, X-ray computerized tomography (CT) is a more accurate modality than MRI in terms of 3D reconstruction of bony structures, but CT exposes the patient to unacceptable doses of ionizing radiations in order to reconstruct the entire spine geometry (all thoracic and lumbar vertebrae). More importantly, both of these modalities can not be done in the standing position. For these above mentioned reasons, biplanar radiography is still the imaging technique which is most frequently used for the 3D clinical assessment of spinal deformities since it allows the acquisition of data in the natural standing posture while exposing the patient to a low dose of radiation.

Due to the 3D nature of AIS, the 3D reconstruction of the spine geometry was also exploited with the goal of defining better indices to characterize the third dimension of scoliosis. Stokes et al. [59] introduced measures in the transverse plane to assess the effect of derotation maneuvers in surgical procedures. Understanding how to classify and quantify 3D spinal deformities remains a difficult challenge in scoliosis. Recently, the concept of 3D vertebra vector parameters has allowed for better measurements compared to 2D measurement [25, 58]. The Scoliosis Research Society (SRS) has therefore recognized the need for 3D classification and mandated the 3D Scoliosis Committee to continue their efforts towards developing a 3D scheme for characterizing scoliosis. Duong et al. [15] proposed an unsupervised fuzzy clustering technique in order to classify the 3D spine based on global shape descriptors, while Sangole et al. [55] proposed a new means to report 3D spinal deformities based on planes of maximal curvature (PMC). More recently, a multivariate analysis using manifold learning was able to identify four separate groups from the same cohort of thoracic deformities [29].

In order to generate 3D models of the patients spine from biplanar radiographic images, a framework (Fig. 1) will generally require the material components, as well as the software components which usually involves an expert in radiology to identify specific anatomical landmarks on the spine. This procedure is not only time-consuming, tedious and error-prone, but the repeatability of the procedure cannot be assured. A few methods have attempted to automate this process by using registration techniques which incorporated ad hoc criterions or by using statistical

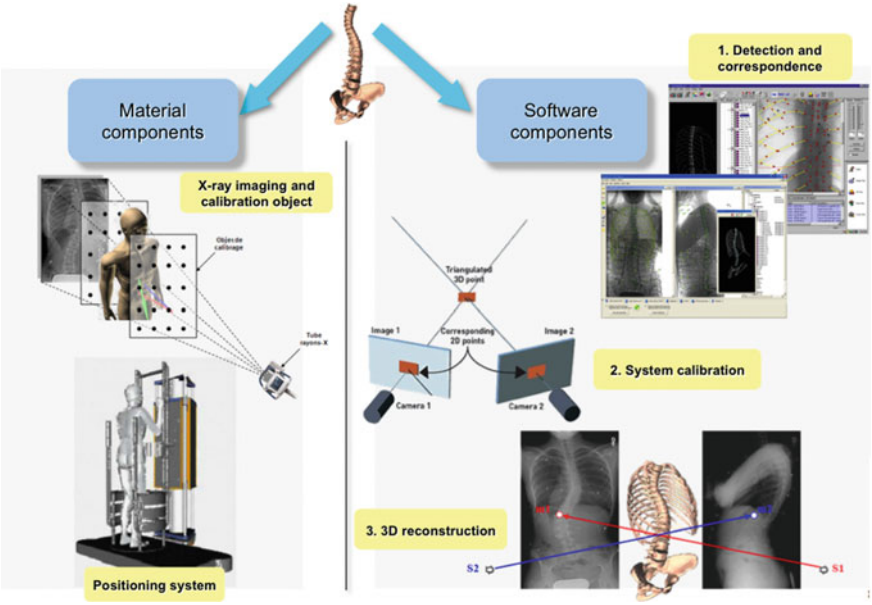


Fig. 1 Overview of the components involved for the 3D reconstruction of the spine from radiographic images

models which would reproduce the variability or the pathological distribution of the deformation [3]. However these methods depend on a pre-calibrated X-ray system, and are often highly supervised in order to identify control points. Given the problems demonstrated by the reconstruction systems adopted in the early years of this technology, it was clear that these techniques were limited on many aspects, and did not possess the necessary tools to insure adequate levels of reproducibility or versatility to be deployed in outpatient clinics.

This chapter presents the different techniques involved for generating the 3D reconstruction of a patient's spine using biplanar radiographs. The personalized 3D reconstruction of the spine enables both the visualization from calibrated biplanar radiographic images and the clinical assessment of spinal deformities in a wide range of pathologies. The remainder of this chapter is structured as follows. Section 2 presents an overview of the evolution of 3D reconstruction systems. In Sect. 3, the approaches for calibrating the X-ray scene are described, while Sect. 4 focusses on the 3D modelling of the spine using either point-based or feature-based representations. For the selected applications, results are given in Sect. 5. Finally, we conclude this chapter in Sect. 6.

2 Evolution of 3D Reconstruction Systems

2.1 First Generation

For years, research in the field of biomechanics has focused tremendously on the analysis of spinal deformities in 3D. The elaboration of complex biomechanical models required rigorous clinical experimentations, as well as quantitative evaluations of the patient's posture and movement analysis in three dimensions. In 1971, Panjabi and White illustrated the importance of studying the spine in three dimensions (3D), with 6 degrees of freedom. A number of techniques for 3D measurements, imaging and modelling of the spine were developed [1, 9, 11, 24, 34]. Since 1992, an imaging system installed at the Sainte-Justine Hospital Research Center enabled to perform the 3D reconstruction of bony structures from radiographic images (Fig. 2). The proposed system was based on the calibration principle of the Direct Linear Transform (DLT), which implicitly includes the geometric parameters in the coefficients of the projection matrices obtained linearly by an inversion of the matrix. Using a Plexiglass cage with encrusted steel pellets, the 2D/3D configuration could be determined in a linear fashion. This tool was used in over 6,000 patient visits for research purposes, in addition to other applications such as surgical simulations and biomechanical analysis of the spine. The system was also used to improve the quality of diagnosis and follow-up exams for patients with idiopathic scoliosis. Finally, this 3D imaging system was used in a number of other projects involving computer assisted surgery and personalization of models.

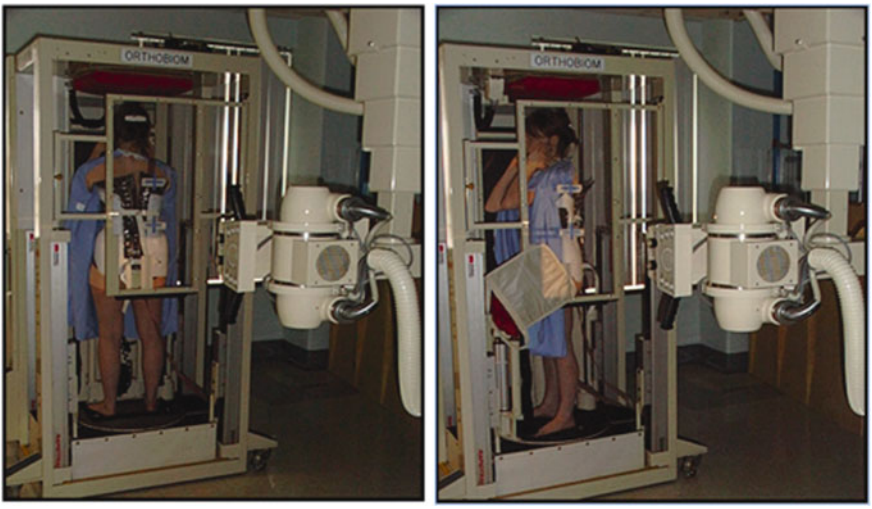


Fig. 2 One of the first 3D reconstruction systems installed at the Sainte-Justine Hospital in 1992, where a plexiglass cage with embedded pellets is used to calibrate the X-ray scene. The patient is turned from the frontal to sagittal position to obtain biplanar radiographs

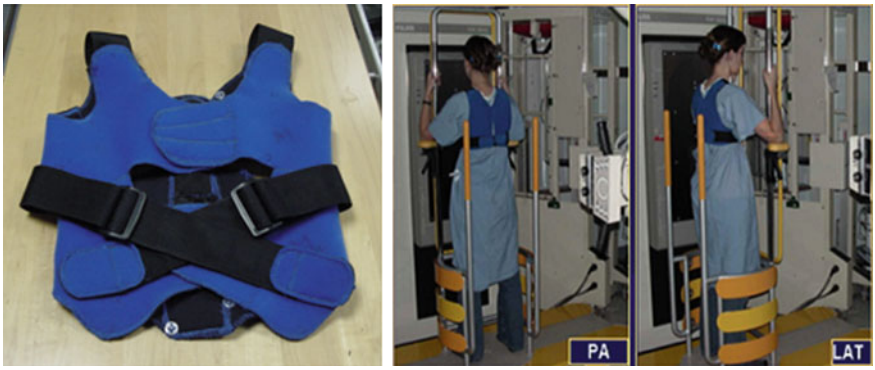


Fig. 3 *Left* Calibration vest worn by the patient and the positioning system used at the Sainte-Justine Hospital. The calibration vest has embedded radio-opaque markers visible on the biplanar to self-calibration the scene using fiducial points closer to the patient's skin. *Right* Positioning system to minimize patient motion between biplanar acquisitions

In order to facilitate the 3D reconstruction of scoliotic trunks, algorithms based on the explicit modelling of calibration matrices were used to calibrate the X-ray scene [7, 8]. However, this particular system required for the patient to be placed in a fixed positioning system while wearing a calibration vest during the radiographic acquisition (Fig. 3), necessitating calibrated images in order to obtain the 3D reconstruction of the spine. Because of the inherent limitations of the system used for research purposes, it became clear such a device could not be deployed in other clinics and thereby exploiting the benefits of 3D reconstruction of the spine. This ultimately limits the universal access to such a technology in a routine setting to assess spinal deformations in 3D.

2.2 New Generation of 3D Reconstruction Systems

The ionization dose given to the patient using conventional radiography in not negligible, and can ultimately induce a risk of cancer or leukaemia with repetitive acquisitions [53], in addition to offer images of average quality. A low-dose imaging system based on the invention of Georges Charpak (1992 Nobel prize of physics [5]) was developed by EOS imaging (Paris, France) to generate high-quality radiographic images, while considerably reducing ionization doses. A first evaluation of the prototype was conducted at the St-Vincent de Paul Hospital in Paris in 1995. Improvements to the system were introduced in order to increase the resolution and reduce the image acquisition time due in part to the new technologies in radio-sensitive detectors. The EOS system (Fig. 4) can acquire simultaneously biplanar radiographic images and includes a software (sterEOS) to perform the reconstruction of bony structures. This system is now installed in over 30 countries, including the US, Canada and Germany.

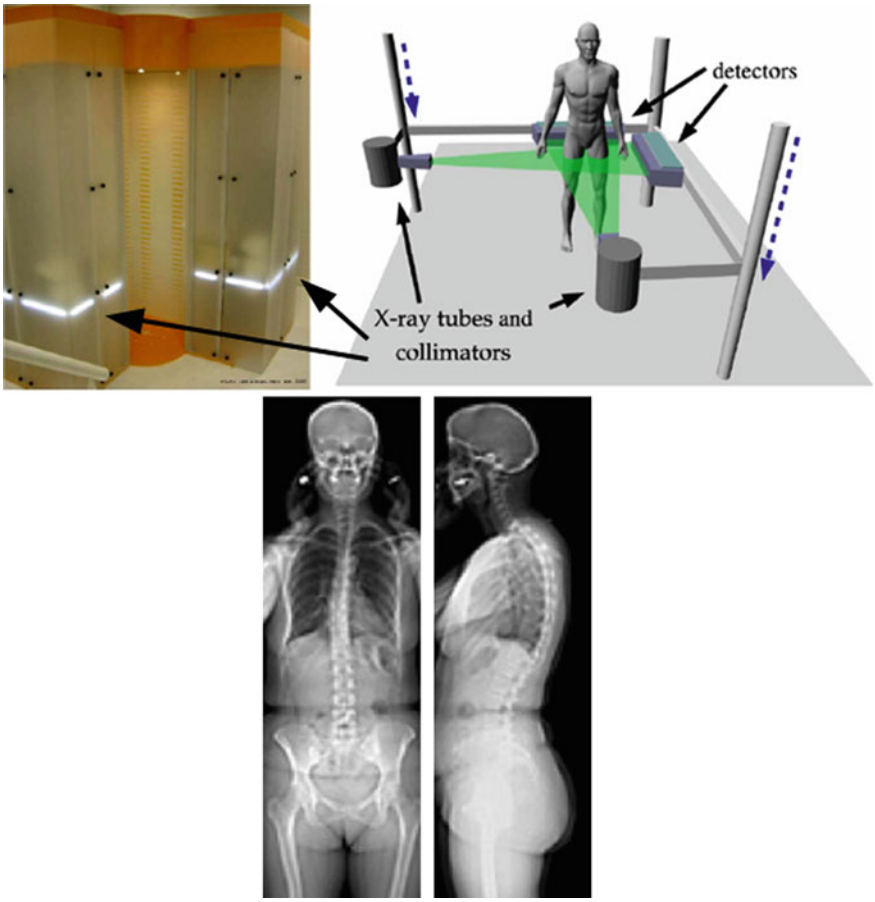


Fig. 4 EOS system (EOS imaging ©) enabling the head-to-toe reconstruction of bony structures (adapted from [13])

3 Radiographic Calibration

In order to generate three-dimensional models from stereo-radiographic images, the radiographic scene must first be calibrated so to calculate a 3D coordinate for a set of 2D projection coordinates. This section presents the various approaches that can be used to perform the calibration process.

3.1 Linear Calibration Techniques

The first 3D reconstruction techniques were based on stereo-radiography, such as the Direct Linear Transform (DLT), which have been widely used for this application [2, 10, 12, 44]. These methods require the use of a large calibration object with fiducial markers of known 3D coordinates and their projection on the 2D images to estimate the intrinsic parameters of the radiographic configuration. However, this setup makes the reconstruction algorithm vulnerable to patient motion between the exposures by creating inconsistencies in the calibration and patient stereo geometries, in addition to being cumbersome in a routine clinical setup. To overcome these limitations, Cheriet et al. [8] proposed an explicit calibration procedure to estimate the X-ray source and film geometric configuration relative to the patient's frame of reference from the content of the images (using a non-rigid calibration vest). The general idea was to adjust the geometric parameters that describe the radiographic setup in such a way as to minimize landmark retro-projection errors. The algorithm used matched calibration landmarks identified on a pair of X-rays and approximate geometric parameters to calibrate the images, while still requiring a calibration object to iteratively update the radiographic system's parameters until it converges to a stable state which reflects a valid solution. These approaches are not suitable for the new generation EOS systems (Sect. 2.2).

3.2 Non-linear Calibration Techniques

Methods have been proposed to enable the 3D reconstruction of the spine from biplanar X-ray images without requiring a bulky object or calibration vest, thus introducing the 3D evaluation of spinal deformities in a wide range of clinical setups [28]. Still, they require an expert to manually identify and match landmarks on the anatomy to calibrate and subsequently reconstruct a model in 3D. In fact, to generate a 3D model of the patient's spine from biplanar radiographs, certain points (anatomical landmarks) on the vertebrae within the image have to be located in order to obtain a 3D model of the scoliotic spine using a triangulation algorithm [12]. Typically, this identification is performed manually by an expert operator and consists of locating anatomical landmarks on a coronal and sagittal radiograph (Fig. 5). However, it is difficult to accurately identify low-level primitives such as exact points and to match them accurately on a pair of views. Thus the repeatability of this procedure cannot be assured. Furthermore this task is time-consuming, tedious and error-prone, and the quality of the 3D reconstruction is directly linked with the precision of 2D localization. Panjabi discussed in detail errors that arise when manually identifying anatomical landmarks [48].

Due to these pitfalls, clinical 3D assessment of the deformity during a patient's visit is therefore not possible. Moreover, because current self-calibration techniques rely on single point correspondences between the biplanar images which offer

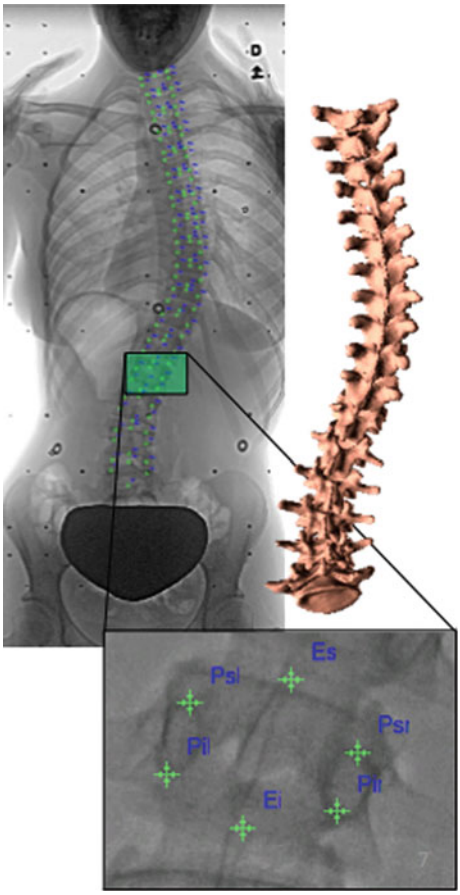


Fig. 5 Manual anatomical landmark identification on the X-ray for the personalized 3D reconstruction of the scoliotic spine. *Visible markers* are the identified landmarks on each vertebra which are used in the self-calibration of an X-ray scene

sparse data with low redundancy [7, 47], this generates multiple local minimums when optimizing the non-linear equation system describing the radiographic setup. Local correspondences also rely on the assumption that a point on an object's surface appears the same in the biplanar images in which it is visible. However due to the phenomena exhibited by the X-ray modality, punctual point matches are not necessarily a reliable matching feature for 3D bone reconstruction. For these reasons, region-based comparisons via surface integration [65] may not only improve the quality of 3D calibration results by incorporating additional data such as high-level corresponding geometrical primitives (curves, surfaces), but can reduce the number of degrees of freedom to solve the equation system.

Furthermore, the self-calibration of a biplanar X-ray system is a complex mathematical problem difficult to solve. In order to improve the quality of current self-calibration techniques [7, 28] which rely on iterative algorithms optimizing the retro-projection of sparse data points to solve a complex system of non-linear equations, it is necessary to incorporate additional data into the system and to

incorporate additional mathematical constraints (e.g. Kruppa's equations and epipolar geometry [21] which are mathematical constructs frequently used for camera self-calibration from a sequence of images). Additional constraints allow the reduction of the number of degrees of freedom in the system of equations without additional data. Using corresponding high-order geometrical primitives (line segments, ellipses, curves, etc.) instead of only point correspondences for the resolution of the self-calibration problem can drastically increase the quantity of data fed into the algorithm. Geometric parameters such as the rotation and translation components of the camera configuration can be determined based on shape information taken from the biplanar projection views, such as with mathematical high level geometrical primitives (lines or ellipses) [52] or with intrinsic properties such as tangent vectors and maximal curvature points to [16]. Although these properties have yet to be used for orthopaedic imaging, properties such as geometrical torsion which describes the 3D phenomena in AIS [51] can be used in the context of calibrating an X-ray scene to establish the 2D-3D relationships for tangential and curvature characteristics extracted from 3D spinal curves [40]. Thus by determining the 3D parameters of a Frenet-Serret frame for example, based on the 2D information collected from the projection images, this set of information can be exploited in a 2D-3D correspondence framework.

3.3 Image-Based Self-calibration

Progressing towards an automated calibration technique therefore involves segmenting anatomical shapes or high level geometrical primitives which can be accurately matched on the X-ray images. Segmentation of bony anatomical structures remains however a challenging problem due to overlapping organs in the thoracic region and low image signal-to-noise ratio. Automatic spine segmentation approaches have been sparsely explored by using machine learning approaches based on localized texture parameters, morphological descriptions in dynamic programming [19] or from Active Shape Models (ASM) using templates from learning data [57]. Based on the work of Cheng et al. [6] which presented a Bayesian approach that uses manually labelled data for parcellation applications, spatial relationships can be used, where the segmentation of the spine shape relied mostly on prior anatomical knowledge information taken from an atlas prior [32]. The core idea of using a manual training set for incorporation of prior statistics and class conditional densities can be transposed in such a work to model the variation distribution of vertebral boundaries by constraining image intensities. This approach is motivated from the fact that segmenting spine contour silhouettes from the biplanar images would offer high level geometrical primitives which could be used to establish 2D-3D correspondence metrics in the self-calibration optimization scheme. Hence, visual reconstructions can be exploited as high-level anatomical primitives, which are subsequently matched between the biplanar X-rays to

determine the 2D-3D relationship of the radiographic scene by self-calibration [27]. Hence, two distinct correspondence features were developed for the discrete non-linear optimization procedure and are described below.

3.3.1 Visual Hull Reconstruction

This section exposes an approach which has been extensively investigated in visual robotics, by adapting an algorithm derived from the shape-from-silhouette principle for a visual hull reconstruction of medical images. First, the spine shapes are first segmented on the images using a prior knowledge Bayesian framework [32] to capture vertebral contour, width, and rotation information in the frontal and sagittal planes. This facilitates the partitioning problem for complex pathological deformations. The silhouettes of the segmented anterior portion of the spine on the biplanar images are then used for the visual hull reconstruction of the global shape of the spine as illustrated in Fig. 6. A visual hull depends both on the spine silhouettes s^i with $i = \{1, 2\}$ and on the camera viewing parameters Π^i such as $\Pi^i : \Re^3 \rightarrow \Re^2$, which remain to be determined. Once the silhouettes are obtained from the images [32], the reconstruction is based on the concept of visual hulls. Formally, the visual hull of the 3D spine S with respect to the viewing plane Ω, denoted by $VH(S, \Omega)$, is a volume in space such that for each point P in $VH(S, \Omega)$ and each camera viewpoint V^i in Ω, the half-line from V^i through P contains at least one point of S [39]. This definition states that the visual hull consists of all points in space whose images lie within all silhouettes viewed from the viewing region.

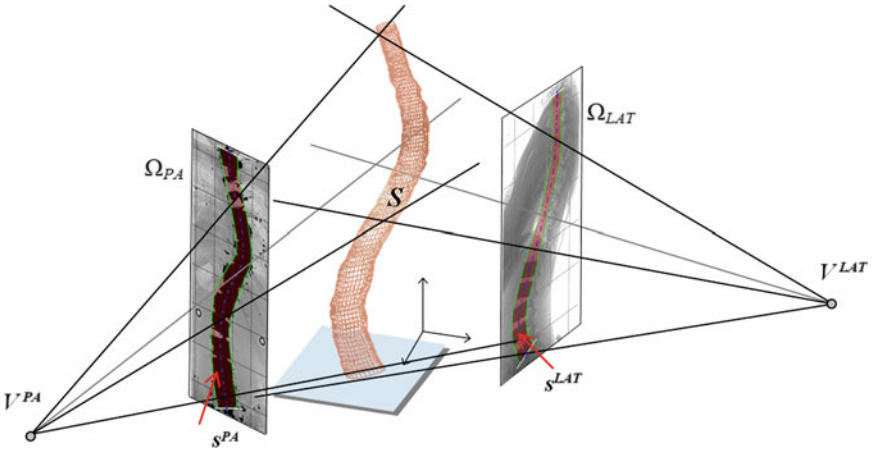


Fig. 6 Principal of the visual hull 3D reconstruction of the global spine shape based on the projected silhouette s^{PA} and s^{LAT} on the biplanar radiographic views. The shape of the object S is estimated by the intersection of both visual cones issued from viewpoints V, offering a sparse and approximate representation of the global spine shape

Hence, the visual hull is the maximal object that has the same silhouettes as the original object, as viewed from the viewing region. The segmented object and spine silhouettes are projected into the 3D space by conical visual hulls in the projective projection model, and its projection should coincide with its silhouette on Ω such that $\Pi^i(S) = s^i$. By computing the intersection of the visual hulls projected from both images (i.e. biplanar viewing directions), the estimation for the shape of the spine is obtained according to the current projection parameters Π^i.

3.3.2 Geometric Torsion of the Scoliotic Spine

Due to the 3D nature of idiopathic scoliosis, the natural curvature properties of the spinal curve can also exploited in the refined optimization of the radiographic parameters, by using the geometrical torsion of the spine measuring the amount of deviation (divergence) of the curved line from the plane determined by the tangent t and normal n vectors [51]. In scoliosis, geometric torsion is related to the amount of helicoidal deformity in the spine. It can be defined as a local geometric property of the 3D curved line passing through thoracic and lumbar vertebrae that measures the amount of helicoidal deviation of the vertebrae, without specific relation to the rotation and deformation of the vertebrae themselves. The continuous parametric 3D B-spline spinal curve $C_k(u) \in \Re^3$ that passes through the center of the spine was represented by Frenet's formulas in order to calculate the geometric torsion as portrayed in Fig. 7. Ultimately, 2D-3D correspondences and a derived relationship which uses Frenet frames to extract 2D information from the image curves are used, and compute their 3D measurements. First, let $\alpha_i(u)$ be a regular curve parameterized

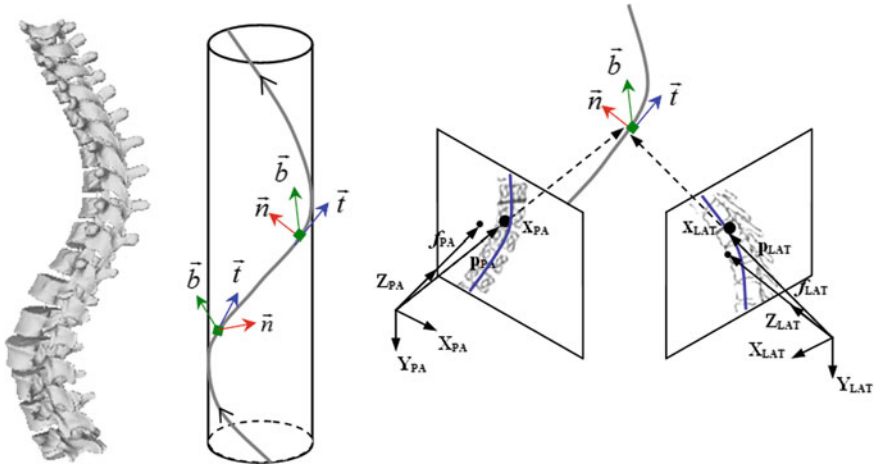


Fig. 7 The scoliotic spine represented as a *helical line* which can be uniquely determined by its geometric torsional quantity. The concept of geometric torsion (τ) is illustrated by the moving trihedron formed by the tangential t, normal n, and binormal b vectors

by arc length in $\Re^2$. If the Frenet (tangent, normal, binormal) frame F = {**T**, **N**, **B**} and the curvature are known at $\alpha_i(0)$, a local approximation is obtained as follows:

$$\alpha_i(u) = \alpha_i(0) + s\mathbf{T}_0 + \frac{s^2}{2}\kappa_0\mathbf{N}_0\frac{s^3}{6}\kappa_0\tau_0\mathbf{B}_0. \tag{1}$$

Hence, $\alpha_{PA}(u)$ and $\alpha_{LAT}(u)$ are defined as the curves representing the spine centerline derived from the segmented silhouettes s^i on the PA and LAT images respectively. Therefore for each value of u such that $u = [0, 2\tau, 3\tau, \ldots, 1]$ and τ is the equidistant step-size, a node in the tangent space $n = (x_{PA}, y_{PA}, x_{LAT}, y_{LAT}, \theta_{PA}, \theta_{LAT}, \kappa_{PA}, \kappa_{LAT})$ is defined, where x_i and y_i are the image projection coordinates of **X**, θ_i the orientation of projected tangent in the image planes, and κ_i the image curvatures, with i representing the biplanar X-rays. Here, the 3D position **X** and tangent **T** are computed using standard methods [22]. To determine the normal **N** and the 3D curvature κ at a 3D space curve point on $C_k(u)$ from biplanar views, the mathematical relationship proposed by Li and Zucker [40] can be used where the 3D normal **N**, the curvature κ and the parameters from the viewing geometry is formulated as:

$$(\mathbf{u}_{PA} \times \mathbf{T}) \cdot \mathbf{N}_\kappa = \frac{f(1 - (\mathbf{u}_{PA} \cdot \mathbf{T})^2)^{\frac{3}{2}}}{\delta(1 - (\mathbf{u}_{PA} \cdot t_{PA})^2)^{\frac{3}{2}}}\kappa_{PA} \tag{2}$$

where $\mathbf{u}_{PA} = \mathbf{p}_{PA}/\|\mathbf{p}_{PA}\|$, **p** is determined from the vector pointing to the projection in the image plane and f is the vector pointing to the image center, both in the camera coordinate system. The δ parameter is the depth computed from calibration while t is the local tangent vector on the 2D image. The curvature can then be computed as $\kappa = \|\mathbf{N}\kappa\|$, given the constraint $\|\mathbf{N}\| = 1$. Hence, the normal vector can be determined by $\mathbf{N} = \mathbf{N}\kappa/\kappa$.

3.3.3 Self-calibration by Means of Optimization of the Radiographic Parameters

The proposed self-calibration algorithm involves explicit use of the description of the calibration matrices M_i in order to estimate the geometrical parameters of the radiographic setup leading to an optimal 3D reconstruction of all the spine's vertebrae [7, 8]. The projective matrices M_i used for the stereo-reconstruction of 3D landmarks are modeled as:

$$M_i(\xi_i) = \begin{bmatrix} c_{x_i} & 0 & x_{p_i} & 0 \\ 0 & c_{y_i} & y_{p_i} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \left[\begin{array}{c:c} R_i & T_i \\ \hdashline 0^T & 1 \end{array}\right] \tag{3}$$

where R_i is the rotation matrix defined by angular $(\alpha_i, \beta_i, \gamma_i)$, T_i is the translation vector $(X_{S_i}, Y_{S_i}, Z_{S_i})$. Intrinsic parameters are modeled by the x_{p_i}, y_{p_i} coordinates of

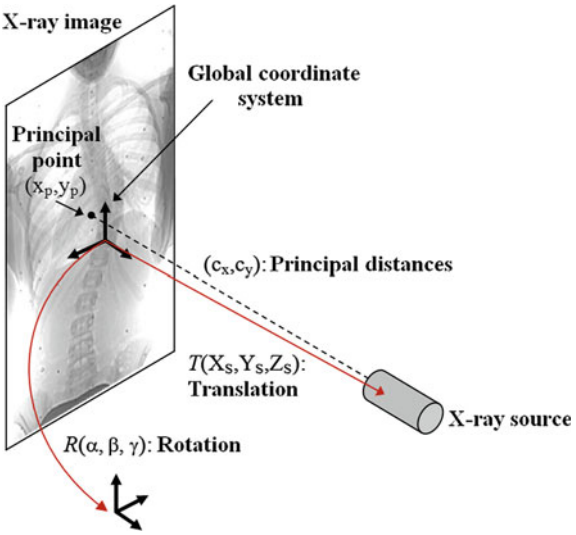


Fig. 8 Illustration of the ten geometric parameters described within the context of the X-rays

the principal point and c_{x_i}, c_{y_i} as the principal distances. These parameters are described schematically in Fig. 8. A rough estimation of these parameters is extrapolated from a small object of known dimensions taken in the radiographic scene [28]. Hence the geometrical parameters $\xi_i = (x_{p_i},\ y_{p_i}, c_{x_i}, c_{y_i}, \alpha_i, \beta_i, \gamma_i, X_{S_i}, Y_{S_i}, Z_{S_i})$ are subsequently updated based on an iterative nonlinear optimization process with regards to the global shape of the spine, following the objective function:

$$E_{\text{global}}(\xi) = \sum_{i=1}^{2} E_{\text{visualhull}}(\xi_i) + \beta E_{\text{torsion}}(\xi_i). \tag{4}$$

This cost function combines two image-based criterions. The first component maximizes the intersected region between the segmented silhouette and the projection of the shape of the spine computed by the visual hull 3D reconstruction, and minimizes isolated regions such that:

$$E_{\text{visualhull}}(\xi_i) = \iint_{\Omega_i - \Pi^i(S)} s^i(u_i, v_i) du_i dv_i - \iint_{\Pi^i(S)} s^i(u_i, v_i) du_i dv_i \tag{5}$$

where $s^i(u, v)$ is the segmented silhouette on image plane i, $\Pi^i(S)$ is the projection of the global visual hull shape S, and Ω_i the image plane domain defined in the (u, v) space. The second component evaluates the difference between the back-projection of the equidistant 3D Frenet frames taken at j/N intervals along the 3D spinal curve $C_k(u)$ and the 2D curves $\alpha_i(s)$ of the X-ray images:

$$E_{\text{torsion}}(\xi_i) = \sum_{j=0}^{N} [\| x_{\alpha_i(j/N)} - \Pi^i(\mathbf{X}_{C_k(j/N)}) \|^2 + \| t_{\alpha_i(j/N)} - \Pi^i(\mathbf{T}_{C_k(j/N)}) \|^2 + \| \kappa_{\alpha_i(j/N)} - \Pi^i(\mathbf{K}_{C_k(j/N)}) \|^2] \quad (6)$$

where N is the number of Frenet frames along the spinal curve. The first two terms evaluate the Euclidean distance between the analytical projection of $\mathbf{X}$ and $\mathbf{T}$ from standard perspective transformation formulae using the current estimate of the geometrical parameters and the image measures x and t respectively. The third term measures the difference in curvature values κ using (14). The method uses a bundle adjustment approach based on an iterative nonlinear optimization process. The Levenberg-Marquardt algorithm is used for optimization, iterating until the correction to the geometric parameters becomes negligible [8]. The set of parameters and projection matrices is therefore regenerated and this procedure is repeated until the system reaches a steady state, where the distance between the observed and computed projection falls to a minimum. To avoid local minima, a directional optimization approach is used to obtain a first, coarse solution which is accurate enough to be used as an initial guess. Moreover, higher reliability of the torsion parameters is ensured in the optimization scheme by enforcing regularity in the Tikhonov sense with the term β. This helps to compensate the instability of the curvature parameters for patients with strong deformations which can affect the convergence of the algorithm. The regularization term acts as a dampening factor, controlling the quality of these parameters by penalizing terms exhibiting very high tangential and torsional values.

4 3D Reconstruction of the Spine

In this section, we present the different methods for generating a 3D model of a vertebra or spine from radiographic images. These methods are categorized in the following classes: point-based, contour-based and statistical methods. We then present a hybrid statistical and image-based approach to generate personalized 3D reconstructions based on geometrical properties in Sect. 4.4.

4.1 Point-Based Methods

The 3D reconstruction of point-based models is usually performed manually by an expert operator and consists of locating six corresponding anatomical landmarks (2 endplate midpoints + 4 pedicle extremities) on each vertebra from T1 (first thoracic vertebra) to L5 (last lumbar vertebra) on a coronal and sagittal X-ray (Fig. 5). Additional non-stereo corresponding point (NSCP) landmarks on the

spinous processes or on the corners of the vertebral body may be added to obtain a more refined and detailed geometry of the 3D vertebrae by deforming generic models using an epipolar geometry [45]. However, it is difficult to identify with precision low-level primitives such as exact points and to match them accurately on a pair of views. Thus the repeatability of this procedure can not be assured. As discussed in Sect. 3.2, the manual identification of landmarks is a long and complex process, which cannot ensure repeatability over multiple raters.

4.2 Contour-Based Methods

Local correspondence also relies on the assumption that a point on an object surface appears the same in the biplanar images in which it is visible. However due to the intrinsic properties exhibited by the X-ray modality, local correspondence is not necessarily a reliable feature for 3D bone reconstruction. Non-stereo corresponding contours (NSCC) methods have been proposed for the 3D reconstruction of anatomical objects demonstrating few corresponding features on the biplanar X-rays, such as for long bones (femur) or the pelvis [38]. The approach optimizes 3D deformations of prior models by minimizing the object's projection from manually identified 2D contours. These techniques demonstrated promising results but are still limited to the manual identification of curves along the edges of long bony structures. For these reasons, variational methods with a region-based component have been applied to multi-view stereo reconstruction as an alternative to local correspondence [65]. Unlike local correspondence, there is no matching of points between pairs of images for consistency, but instead the comparison is integrated over regions. This can not only improve the precision of reconstruction results by incorporating additional data such as high-level corresponding geometrical primitives (curves, surfaces), but reduces manual intervention required to identify specific anatomical landmarks on the X-rays.

4.3 Statistical Methods

In order to reduce inaccuracies on the 2D localization of landmarks and to be a clinically useful procedure, previous studies were conducted to propose more automated methods. Initial attempts were based on vertebral template matching [46, 56], and feature-based by using active shape models (ASM) [57] or Hough transforms [23, 64], to detect dominant characteristics (corners, edges) from the vertebral shape body. Still, these techniques were ineffective towards noise and varying appearance in shape. Statistical shape models, and more recently 2D-3D registration methods, have been the focus of increased attention for the 3D reconstruction of the human spine. A variety of methods have been proposed in the previous years for image to physical space (patient) registration. While some have

used preoperative 3D models from CT or MR images to register with 2D X-ray or fluoroscopic images from gradient amplitudes [20, 41, 60], Fleute and Lavallee have used statistical a priori knowledge of the 3D geometric shapes in order to model the 2D vertebral shapes by applying point distribution models (PDM) [17]. Similar approaches introduced by Lorenz et al. [43] and Vrtovec et al. [63] have used PDM methods from training statistical shape models, thus automatically capturing the geometrical knowledge of the principal modes of variation to isolate 3D vertebrae from tomography images. A method proposed to use a priori knowledge of the vertebral shape using eight morphologic descriptors of the vertebral body to accurately estimate the geometrical model [50]. The obtained model would be manually refined by projecting the spine's silhouette on the X-rays. Inference-based optimization refinements were subsequently presented to obtain an accurate estimate of the vertebra's orientation and 3D locations [14]. Still, these approaches remain highly supervised by an operator to manually identify landmarks. Benameur et al. [3] proposed a 3D-2D registration method for vertebrae of the scoliotic spine. In this case, the geometric knowledge of isolated normal vertebrae is captured by a statistical deformable template integrating a set of admissible deformations, and expressed by the first modes of variation in a Karhunen-Loeve expansion. However, none of these methods have attempted to integrate a statistical model taking into account the set of admissible deformations for the whole scoliotic spine shape. Another drawback from most methods is that each vertebra is treated individually instead of as a whole articulated model which may include the global 3D deformation of the spine. Hence in order to account for the global geometrical representation of scoliotic deformities, a variability model (mean and dispersion) of the whole spine allowed increasing the accuracy of the 2D-3D registration algorithm by incorporating knowledge-based inter-vertebral constraints [4]. Klinder et al. [35] has transposed these 3D inter-vertebral transformations to accomplish the segmentation of the spinal cord from CT-scan images. In fact, 3D spinal curve analysis where a model of the curvature of the vertebral column describing the relationship between vertebrae has been particularly useful for 3D medical image analysis of the spine. Because of the intricate and tortuous 3D nature of scoliosis, automated curved planar reformation (CPR) techniques have been presented [62] in order to increase visualization of the deformity by transforming the orthogonal and transverse references to a spinal coordinate system. Furthermore, CPR has been used to assist in the spine segmentation problem using a reformed 3D spinal centerline [33] or by exploiting the approximate proximity of vertebrae along the centerline [18].

4.4 Personalized 3D Reconstruction of the Spine

Given the significant challenges in the reconstruction of scoliotic spines, which exhibit high variability not only on the global pose due to changing inter-vertebral transformations, but also within the local appearance of scoliotic vertebrae (rotation,

wedging), we present here the components of a hybrid statistical and image-based biplanar reconstruction method [26]. The spine centerlines extracted from the pre-operative images are used to map the 3D reconstruction of the spinal curve in a low-dimensional representation of a scoliotic database, and perform a statistical modeling of the anatomy based on an analytical regression. The model is refined locally at each vertebral level via a segmentation method based on a level set surface evolution paradigm.

4.4.1 Training Data

The statistical model was built from a 3D database containing 711 scoliotic spines demonstrating several types of deformities. Each scoliotic spine in the database was obtained from biplanar stereo-reconstructions. It is modeled with 12 thoracic and 5 lumbar vertebrae (17 in total), represented by 6 landmarks on each vertebra (4 pedicle extremities and 2 endplate center points).

Segmentation of the scoliotic vertebrae on the X-ray images was performed by using generic vertebra priors obtained from serial CT-scan reconstruction of a cadaver specimen (Fig. 9). Models were segmented using a connecting cube algorithm [42] with 1-mm-thick CT-scan slices taken at 1-mm steps throughout the dry spine. The atlas is composed of 17 cadaver vertebrae (12 thoracic and 5 lumbar). The atlas is divided into 3 levels of polygonal mesh catalogues of increasing complexity, to adopt the widely used multi-resolution registration approach where coarse-to-fine geometrical models are applied for optimal convergence, where models were composed between 3,831 and 6,942 vertices depending on the vertebra level. The same six precise anatomical landmarks (4 pedicle tips and 2 on the vertebral body) were annotated on each individual model.

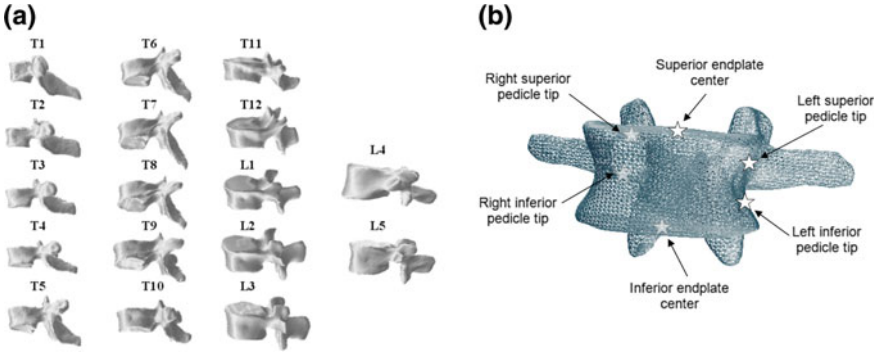


Fig. 9 **a** Atlas of 17 vertebral models obtained from computer tomography (*CT*) and represented with 3D Fourier descriptors. **b** Annotated landmarks of the first lumbar vertebra

4.4.2 Manifold-Driven Model Generation

The input of the method consists of calibrated coronal and sagittal X-ray images $I_{i=\{1,2\}}$ (defined is space Ω_i) of the patient's pre-operative spine acquired. The personalized 3D model is achieved by means of a reconstruction method merging statistical and image-based models [26]. The 3D spine centerline $C_i(u)$ is obtained from cubic B-splines extracted from the images. The centerline is first embedded onto a non-linear manifold $\mathcal{M}$ containing M scoliotic spines ($M = 711$) and used to predict an initial spine. The manifold establishes the patterns of legal variations of spine shape changes in a low-dimensional sub-space based on locally linear embeddings as illustrated in Fig. 10. To map the high-dimensional 3D curve assumed to lie on a non-linear manifold into a low-dimensional subspace, the first step consists of selecting the K closest neighbors for each data point using the Euclidean distance between centerlines as a closeness measure. The manifold reconstruction weights W are then found to reconstruct point i from it's K closest neighbors using the reconstruction errors as measured by:

$$\varepsilon(W) = \min_{W} \sum_{i=1}^{M} \left\| C_i(u) - \sum_{j=1}^{K} W_{ij} C_j(u) \right\|^2 \tag{7}$$

where $C_i(u)$ is a data B-spline described above and $\varepsilon(W)$ sums the squared distances between all data points and their corresponding reconstructed points. The minimum of $\varepsilon(W)$ which describes the optimal weight matrix W is then determined by solving a least-square problem. The weights W_{ij} represent the importance of the

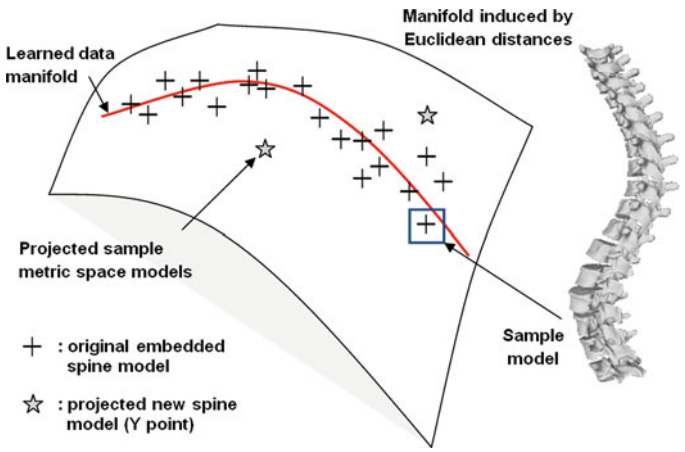


Fig. 10 Illustration of spine distribution embedded onto on a low-dimensional manifold. *Stars* represent new sample point models which were unseen in the training set and projected onto the manifold. *Pluses* are original model points from the training data which creates the manifold distribution

*j*th data point to the reconstruction of the *i*th element. The last step consists of mapping each high-dimensional $C_i(u)$ to a low-dimensional point Y_i, representing each data point in the global coordinate system in d-space, using a cost function which minimizes the reconstruction error:

$$\Phi(Y) = \sum_{i=1}^{M} \|Y_i - \sum_{j=1}^{K} W_{ij} Y_j\|^2. \tag{8}$$

The coordinates Y_i can be translated by a constant displacement without affecting the overall cost $\Phi(Y)$. This degree of freedom is removed by requiring the coordinates to be centered at the origin, such that $\sum Y_i = 0$. The optimal embedding, up to a global rotation of the embedding space, is obtained from the bottom $d+1$ eigenvectors of the sparse and symmetric $M \times M$ matrix enclosing the reconstruction weights W_{ij}. This helps to minimize the cost function $\Phi(Y)$ as a simple eigenvalue problem. The d eigenvectors form the d embedding coordinates in $\mathcal{M}$ [54]. Hence, a new model point can be determined in the embedded d-space as a low-dimensional data point by finding its optimal manifold coordinates Y_i.

Given a new projection point Y_n, an appropriately scaled model is generated from an analytical method based on nonlinear regression using a Radial Basis Function kernel function f to perform the inverse mapping. Formally, the model is described such that $\mathbf{S} = [f_1(Y_n), \ldots, f_D(Y_n)]$ with $\mathbf{S} = (s_1, s_2, \ldots, s_{17})$, where s_i is a vertebra model defined by $s_i = (p_1, p_2, \ldots, p_6)$, and $p_i \in \Re^3$ is a 3D vertebral landmark. Details can be found in reference [26].

This crude statistical 3D model is refined with an individual scoliotic vertebra segmentation approach. This is achieved by extending 2D geodesic active regions in 3D, in order to evolve prior deformable 3D surfaces by level sets optimization. Each component model S_i from the atlas of vertebral triangulated meshes, represented by the triangular mesh vertices $\{\mathbf{v}_j^i | j = 1, \ldots, V\}$, are initially positioned and oriented from their respective 6 precise landmarks s_i composing $\mathbf{S}$. The surface evolution is then regulated by the gradient map and image intensity distributions [49], where $E_{\mathrm{RAG}} = \lambda E_{\mathrm{CAG}}(S) + (1-\lambda) E_{\mathrm{R}}(S)$ is the energy function with the edge and region-based components controlled by λ are defined as:

$$\begin{aligned} E_{\mathrm{CAG}} &= \sum_{i=1}^{2} \oint_{S_i} \frac{1}{1 + |\nabla I_i(\mathrm{u}_i)|^{\alpha}} d\mathrm{u}_i \\ E_{\mathrm{R}} &= -\sum_{i=1}^{2} \iint_{\Pi_i(S_i)} \log(p_R(I_i(\mathrm{u}_i))) d\mathrm{u}_i \\ &\quad -\sum_{i=1}^{2} \iint_{\Omega_i - \Pi_i(S_i)} \log(p_{Rc}(I_i(\mathrm{u}_i))) d\mathrm{u}_i \end{aligned} \tag{9}$$

with Π_i as the perspective projection parameters, while p_R and p_{Rc} are Gaussian distributions for bone and background regions respectively. The projected

silhouettes of the morphed 3D models would therefore match the 2D information on the biplanar X-rays in the image domain u_i, replicating the specifics of a particular scoliotic deformity.

4.4.3 Bundle Adjustment of the 3D Vertebral Landmarks

The crude statistical 3D model of the personalized spine is subsequently refined by adjusting the 3D coordinates of the vertebrae. For a bundle adjustment of the 3D landmark coordinates, a non-linear optimization method minimizes the cost function $E(s_l)$ and updates the 4 pedicle extremities and 2 endplate centers (6 anatomical landmarks) in s_l at each vertebral level l (starting from L5 and progressing to T1), based on the measures taken on the biplanar images.

Cost Function

The Powell-Brent optimization method minimizes a cost function combining image edge alignment from the 3D surface model, epipolar geometry correspondence and morphological constraints formulated by Eqs. (11), (13) and (14) which are described below:

$$E(\bar{s}_l) = \omega_1 D_{edges} + \omega_2 D_{epipolar} - \omega_3 D_{morphology} \tag{10}$$

where $\bar{s}_l = R(\bar{s}_{l-1})[\bar{s}_l] + T(\bar{s}_{l-1})$ takes into account the previous updated vertebra model and (R, T) is the rigid displacement of landmarks p_i at the previous vertebra level s_{l-1} before/after optimization. The weights ω are dynamically assigned on a vertebral level basis with ω_1 representing the image-based criterion of the cost function regulated by a pixel coherence factor [30], ω_2 represents the epipolar geometry constraint regulated by the calibration accuracy obtained in Sect. 3, while ω_3 enforces the criterion such that $\omega_1 + \omega_2 + \omega_3 = 1$. The set of 3D landmarks p_i for each vertebra s_i are globally adjusted based on the following measures.

Image Gradient Edge Alignment

In order to integrate image-based information in the optimization process, a similarity measure estimates the distance of the projection of a 3D deformed model to the computed gradient of the X-rays. The approach would: (1) deform the prior generic high resolution 3D vertebra model obtained from CT acquisitions, using the level set surface evolution technique with the set of landmarks p_i evolving with the same deformation Eq. (9); (2) project the triangulated mesh and the distance measure $\varphi(x, y, z; t)$ of the 3D model using the projection parameters of the 3D radiographic scene to create a silhouette onto the images; (3) compute a 2D distance map for these edges and; (4) sum over the distance map values at the locations

indicated by the edges of the gradient image. Given the binary gradient X-ray image, the distance of an image point q to the projected edge structures $V = \{v_j\}$ is $d(q) = \min_j |q - v_j|$. However due to the poor quality of the images, a precise edge information cannot be obtained and the gradient images may not correspond to the edge templates. The proximity to edges can be defined by using a Gaussian expression controlled by the parameter σ^2 and weighted by the projected 2D distance measure of the surface evolution:

$$D_{edges} = \sum_{i=1}^{2} \sum_{j} p_{ij} \exp \frac{((q - v_{ij})\varphi_i(x, y, z))^2}{\sigma^2} \tag{11}$$

where p_{ij} is the probability for pixel v_j in image i of being an edge. The distance measure $\varphi(x, y, z; t)$ of the surface evolution in 3D is projected on image plane i using the rotation component θ of the projection parameters ξ_i (Sect. 3):

$$\varphi_i(x, y, z) = \varphi(x, y, z) \sin \left(\frac{\frac{\partial \varphi(x,y,z)}{\partial x} \cos \xi_i(\theta) + \frac{\partial \varphi(x,y,z)}{\partial y} \sin \xi_i(\theta)}{|\nabla \varphi(x, y, z)|} \right). \tag{12}$$

To determine the values of p_{ij}, a two-dimensional proximity function can be computed by convoluting the image with a large Gaussian kernel.

Epipolar Geometry Constraint

The calibration of the 3D radiographic viewing geometry was also used to constrain the landmark correspondence between the biplanar images. An iterative retro-projection method helps to refine landmark position, by taking the current 3D landmark location, project it in 2D onto the coronal (PA)/sagittal (SAG) views and measure the perpendicular distance of the projected coordinate on both views to its corresponding epipolar line. The distance error for the L landmark points ($L = 6$ representing the 4 pedicle extremities and 2 endplates) is defined as:

$$\begin{aligned} D_{epipolar} = \sum_{i=1}^{L} [&Eucl(\hat{w}_i^{SAG}, \mathbf{F}^T \hat{w}_i^{PA})^2 \\ &+ Eucl(\hat{w}_i^{PA}, \mathbf{F}^T \hat{w}_i^{SAG})^2] \end{aligned} \tag{13}$$

where $Eucl(*)$ denotes the Euclidean distance of a point to a line, $\hat{w}_i$ is the analytical projection of the 3D object point p_i obtained from standard perspective transformation formulae. $\mathbf{F}^T \hat{w}_i$ is the corresponding epipolar line on one image based on point p_i from the other image, and $\mathbf{F}$ is the 4×4 fundamental matrix integrating the geometrical parameters ξ which describes the projective 3D structure of the scene.

Likelihood Estimation

A likelihood estimation model integrating 2D morphological and feature information was included in order to measure the error given from the current data. This estimate expresses the measure of similarity between the current model points $\hat{w}_i$ and an estimate $w_i(x)$ which encodes expert morphological knowledge of the relationships between the L landmarks [31]. Each landmark i is assigned to a specific function of $w_i(x)$ depending on the landmark type (i.e. pedicle tip), and is based on local vertebral height, width, orientation and relative distances between landmarks. The model also measures the similarity response of a rotation and scale invariant wavelet coefficient feature $c_{msd}(\hat{w}_i)$ specific to the landmark type, at location $\hat{w}_i$ on the image. The probability of this likelihood estimate is:

$$D_{morphology} \propto \prod_{i=1}^{L} \left\{ \exp\left[-\frac{1}{2} \left(\frac{\psi_{PA}(\hat{w}_i^{PA}) + \psi_{SAG}(\hat{w}_i^{SAG})}{2\sigma} \right)^2 \right] \Delta w \right\} \tag{14}$$

where $\psi_{PA}(\hat{w}_i^{PA})$ and $\psi_{SAG}(\hat{w}_i^{SAG})$ are the similarity measures for the landmark coordinates on the coronal and sagittal plane defined as:

$$\psi_{PA}(\hat{w}_i^{PA}) = (\hat{w}_i^{PA} - w_i^{PA}(x)) * c_{msd}(\hat{w}_i^{PA}) \tag{15}$$

$$\psi_{SAG}(\hat{w}_i^{SAG}) = (\hat{w}_i^{SAG} - w_i^{SAG}(x)) * c_{msd}(\hat{w}_i^{SAG}). \tag{16}$$

In Eqs. (15) and (16), $w_i^{PA}(x) = \sum_j \delta_j f_j(x)$ and $w_i^{SAG}(x) = \sum_j \delta_j s_j(x)$ are the estimates of the landmark coordinates on the coronal and sagittal plane respectively based on morphological distribution of the neighbouring j landmarks. Each landmark j is assigned with prior knowledge distances $f_j(x)$ and $s_j(x)$ for the PA and SAG views respectively, pondered by the predefined weights δ.

5 Results

5.1 Self-calibration of the Radiographic Scene

5.1.1 Validation Methodology

A clinical validation using real data has assessed the clinical validity of the presented self-calibration algorithm. A comparison between a previously validated system using an explicit calibration based on manually identified landmarks in a fixed radiographic setup [8], and the proposed system based on uncalibrated X-rays was made by generating a 3D model of the spine using both techniques. The data used for the clinical study consisted of 60 pairs of digitized X-rays of adolescents

with AIS. The inclusion criteria for this study was adolescent subjects who had their X-rays taken during a scoliosis clinic consultation for either diagnosis or follow-up, and had a calibration object placed during radiographic acquisition in order to compare results. For each patient in the data set, a 3D reconstruction of the spine was obtained from both methods. A series of clinical 2D and 3D geometrical parameters were subsequently computed from these models and compared between both techniques.

5.1.2 In Vivo Clinical Validation

Table 1 presents the results from this validation. Retro-projection errors are significantly lower ($p \leq 0.05$) when using the proposed system based on self-calibration with uncalibrated X-rays. It is somewhat difficult to evaluate this finding because there is no gold standard to compare these results with. However because of the intrinsic effect of 2D errors on the 3D model, it can be deduced that the better the geometrical epipolar matching is in the stereoscopic vision, the better is the resulting 3D model. The value of the computerized Cobb angle in the frontal plane (C_{PA}^{PT}, C_{PA}^{MT}, C_{PA}^{L}) with both systems is very similar, with slightly higher differences in the sagittal plane (C_{LAT}^{T4-T12}, C_{LAT}^{L1-L5}). The orientation of the planes of maximum curvature (θ_{PMC}^{PT}, θ_{PMC}^{MT}, θ_{PMC}^{L}) offers very acceptable differences, with insignificant differences set at $p \leq 0.05$. Balance (y_{T1-L5}, x_{T1-L5}), however, gives a greater difference compared to the previous linear appraoch, which is explained by the fact that the reference planes are different. While the calibrated X-rays use the external calibration plate, the uncalibrated X-rays use the images coordinate system as a reference.

Table 1 RMS difference and Wilcoxon test results of the geometrical indices measured on 60 3D reconstructions of the spine obtained from the proposed system and a previous method [28]

Parameter	Symbol	Unit	Mean diff.	*p*-value
Epipolar error	ε	mm	0.97 ± 0.61	<0.001 (SD)
Cobb angle (PT)	C_{PA}^{PT}	deg	0.31 ± 0.26	0.69 (NS)
Cobb angle (MT)	C_{PA}^{MT}	deg	0.19 ± 0.17	0.59 (NS)
Cobb angle (L)	C_{PA}^{L}	deg	0.29 ± 0.27	0.34 (NS)
Kyphosis	C_{LAT}^{T4-T12}	deg	0.52 ± 0.41	0.16 (NS)
Lordosis	C_{LAT}^{L1-L5}	deg	0.63 ± 0.37	0.21 (NS)
Max. deformity (PT)	θ_{PMC}^{PT}	deg	0.57 ± 0.54	0.17 (NS)
Max. deformity (MT)	θ_{PMC}^{MT}	deg	0.56 ± 0.49	0.16 (NS)
Max. deformity (L)	θ_{PMC}^{L}	deg	0.54 ± 0.36	0.55 (NS)
Axial rotation	θ_{APEX}^{MT}	deg	0.87 ± 0.78	0.15 (NS)
Frontal balance	y_{T1-L5}	deg	0.97 ± 0.53	0.04 (SD)
Sagittal balance	x_{T1-L5}	deg	1.88 ± 1.15	<0.01 (SD)

SD: significant difference; NS: non-significant difference

From the clinical results shown in this section, it can be observed that the image-based self-calibration system used for the 3D reconstruction of the spine compares very well with models obtained with the previous system using low-level primitives (landmarks) which were identified manually by an expert operator. From a clinical perspective, it demonstrates that the image-based self-calibration method yields results with insignificant differences ($p < 0.05$) with regards to a set of standardized 3D measurements used for the clinical evaluation and the assessment of adolescent scoliosis, without relying on manual landmark identification.

5.2 Personalized 3D Reconstruction of the Spine

5.2.1 Clinical Validation

The proposed 3D spine reconstruction method was applied to scoliotic patients recruited at the scoliosis clinics of Sainte-Justine Hospital (Montreal, Canada). The selection of the patients included in this group was based on the availability of the images needed to compute 3D reconstructions of the spine, and that all patients had 12 thoracic and 5 lumbar vertebrae. Twenty pairs of biplanar X-ray images taken from scoliotic patients with mild deformities (Cobb angle range 15°–40°) were used to evaluate the 2D and 3D differences of the proposed method. For each case, comparisons between results obtained with the proposed method and those from a radiology expert were established.

To assess the precision of the image-based similarity measure used for the optimization procedure, Fig. 11 shows results with the retro-projection of the deformed 3D vertebra contours (high-level primitive) fitting adequately to the bony edges of the corresponding vertebra in the coronal and sagittal X-ray image. The qualitative evaluation of the global method also shows the projected anatomical landmarks obtained from the optimized 3D model, and yield better accuracy in terms of epipolar geometry to the 2D locations manually identified by a radiology expert on each vertebra. Figure 12 presents a box-whisker diagram with the overall representation of differences and errors for the group of patients. The overall sum of squared differences (method vs. observer) for the selected cases was of 0.9 ± 0.7 mm for the 2D point landmarks and of 1.8 ± 0.9 mm for the vertebral contours. However the root-mean-square (RMS) epipolar geometry error (distance of the landmarks to the epipolar line) yields significantly lower errors ($p < 0.05$) for the proposed method compared to a manual technique (1.5 ± 1.2 vs. 4.7 ± 3.2 mm). The point-to-point mean difference between the 3D spine models issued from the proposed technique and from a manual identification yielded a 3D mean difference of 1.8 ± 1.5 mm for lumbar vertebra and 2.2 ± 1.6 mm for thoracic vertebra. Differences are slightly higher in the thoracic region due to extrapolation errors and lower visibility, thus offering less image-based information on the X-ray images.

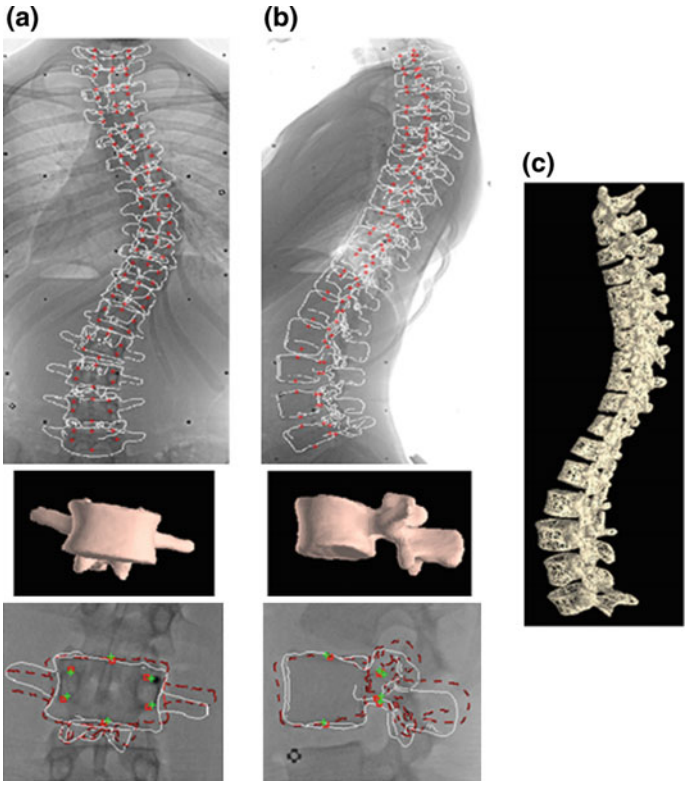


Fig. 11 Results obtained on a **a** coronal and **b** sagittal X-ray image. Comparison of landmark and projected contours results from the manual technique (*square/dashed line*) and the proposed method using 3D deformable vertebra models (*cross/solid line*). **c** Final 3D reconstruction from the proposed method

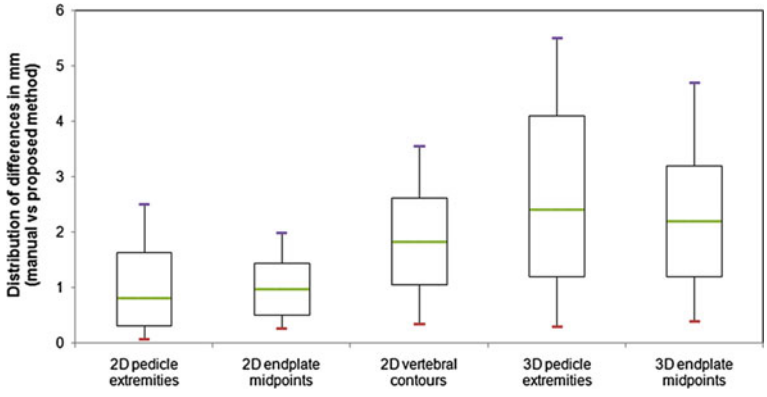


Fig. 12 Distribution of 2D and 3D landmark differences, obtained from the manual technique and the proposed method

Table 2 Results on point-to-surface comparisons with mean, root-mean-square (*RMS*) and maximum errors of 8 scoliotic vertebrae with 3D models obtained from MRI data

Vertebra	*N*	Mean (mm)	RMS (mm)	Max. (mm)
T10	1	1.0 ± 0.7	1.3 ± 0.4	4.2 ± 0.9
T11	2	0.9 ± 0.8	1.3 ± 0.3	3.4 ± 1.0
T12	2	1.3 ± 0.9	1.8 ± 0.3	4.4 ± 1.2
L2	2	1.1 ± 1.0	1.4 ± 0.2	4.3 ± 0.9
L3	1	1.3 ± 1.2	1.7 ± 0.3	4.5 ± 1.1

N Denotes the total number of vertebrae at different vertebral levels

5.2.2 Ground-Truth Comparison to MRI Data

To evaluate the overall accuracy of the 3D reconstruction system on real patient data, 8 vertebrae (4 lumbar and 4 thoracic) from two scoliotic patient who were scanned with an MRI device (AVATO, Siemens Medical Solutions, Germany) were used for comparison. Both patients also had their biplanar X-rays taken prior to surgery. Each slice of 1 mm thickness was taken with a resolution of 256 × 256 pixels and 12 bits per pixel with no interspacing. The results of the comparisons are expressed as *point-to-surface* distances, i.e., each point of reconstructed vertebra is projected onto the surface on the corresponding scanned vertebra and the point-to-surface Euclidean distance is computed. For an appropriate clinical comparison, the same six landmarks from the personalized 3D reconstruction (endplate midpoints, pedicle extremities) were identified on each vertebra using an interactive graphical computer tool in order to rigidly register both 3D vertebra models.

The overall point-to-surface comparison results (mean, RMS, and maximum) between the reconstructed 3D vertebral models issued from the proposed image-based stereo-radiography method and from MRI scans are presented in Table 2. The mean point-to-surface errors are 1.2 ± 1.1 mm for lumbar vertebra and 1.1 ± 0.8 mm for thoracic vertebra. Visual comparisons between the 3D reconstruction using the proposed reconstruction technique and the reference model (MRI) are presented for thoracic vertebrae in Fig. 13. The results show the patient-specific vertebral models obtained from stereo-radiography offers an adequate correspondence with the ground-truth 3D representation given by MRI, specifically in the vertebral bodies and pedicle regions. The validation results presented above show that the accuracy of the statistical image-based 3D reconstruction method is comparable to ground-truth 3D reconstructions obtained from MRI data.

5.3 Classification of Spinal Deformities

Using the proposed methods for calibrating and generating a 3D spine model from X-ray image, a cohort of 170 AIS preoperative patients with right thoracic deformations was processed, classified as Lenke Type-1 by members of the 3D Scoliosis Committee of the SRS, in order to uncover potential 3D subclasses within this

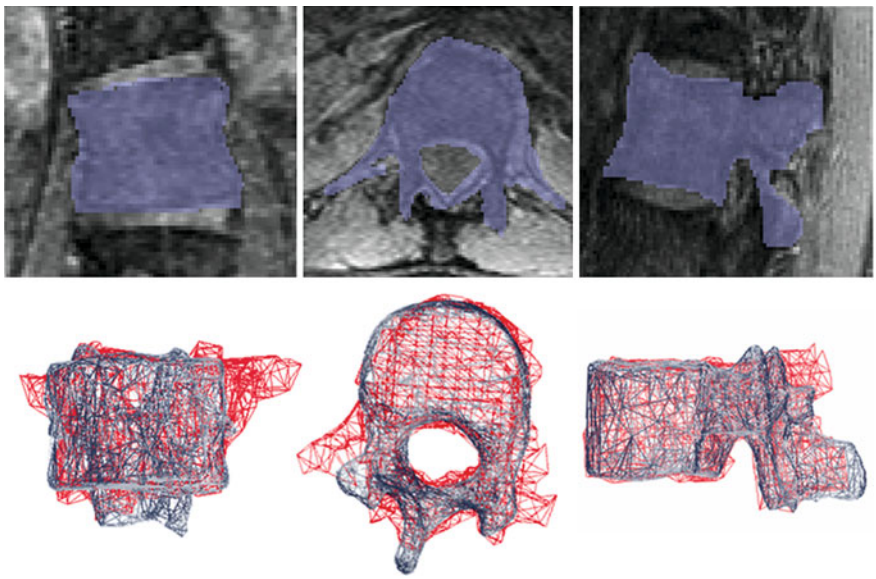


Fig. 13 Corresponding segmented cuts from MRI scans, along with 3D polygonal mesh superimposed with 3D reconstructions obtained from the proposed method

group. The proposed non-linear manifold embedding algorithm presented in Sect. 4 was able to reduce the high-dimensionality of the 3D data, using statistical properties of neighbouring spine models to infer a global representation of the sub-population.

Four clusters were detected from the low-dimensional manifold of 3D models based on inter- and intra-cluster measures. The result from this classified embedding is presented in Fig. 14. The first group consisted in 37 patients with normal thoracic kyphosis profiles with hyper-lordosis and the highest levels in main thoracic Cobb angles for all groups. The second group had 55 patients, mostly non-surgical (minor curves), with low thoracic kyphosis and normal lumbar lordosis values, but with the highest degree of rotation of the PMC from the sagittal plane. The third group had 21 cases with hypokyphotic and hyper-lordotic profiles. Finally, the fourth and last group included 57 patients with hyper-kyphotic thoracic profiles, with major surgical curves and demonstrating very high axial rotations in the apical vertebrae. These results show an additional group in contrast to the study by Sangole et al. which found hypo-kyphotic subgroups in a 3D analysis which used measures such as planes of maximal curvature and kyphosis as influential parameters that split the cases. Coronal and sagittal spine profiles, as well as the da Vinci schemas [55] of the representative cases that were identified as the cluster centers are illustrated in Fig. 15. The manifold representation can potentially be useful for classification of 3D spinal pathologies such as AIS and serve as a tool for understanding the progression of deformities in longitudinal studies.

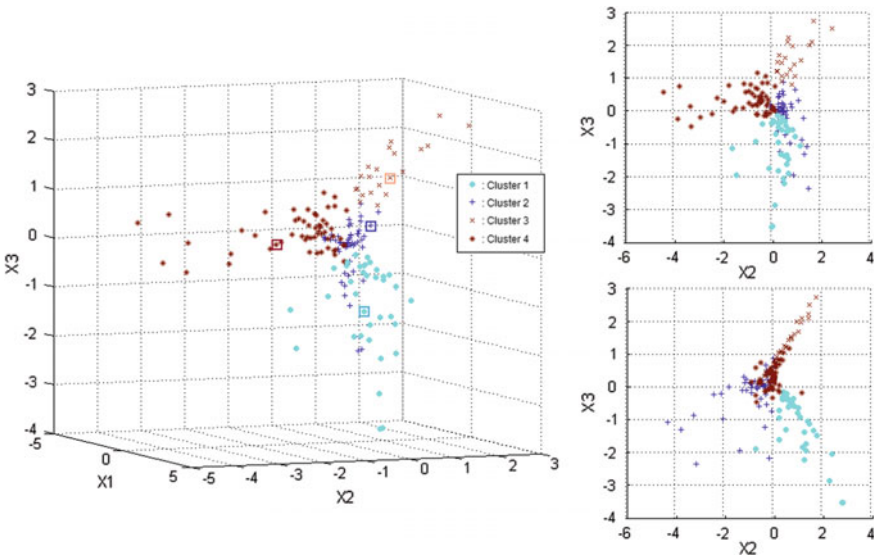


Fig. 14 Result of the low-dimensional manifold of 3D spine models from 170 Lenke type-1 patients, clustering into four sub-groups. Cluster center points from these groups show (*1*) normal kyphosis with hyperlordosis, (*2*) low kyphosis and normal lordosis, with high rotation of plane of maximum curvature, (*3*) hypo-kyphotic and hyper-lordosis and (*4*) hyper-kyphotic cases

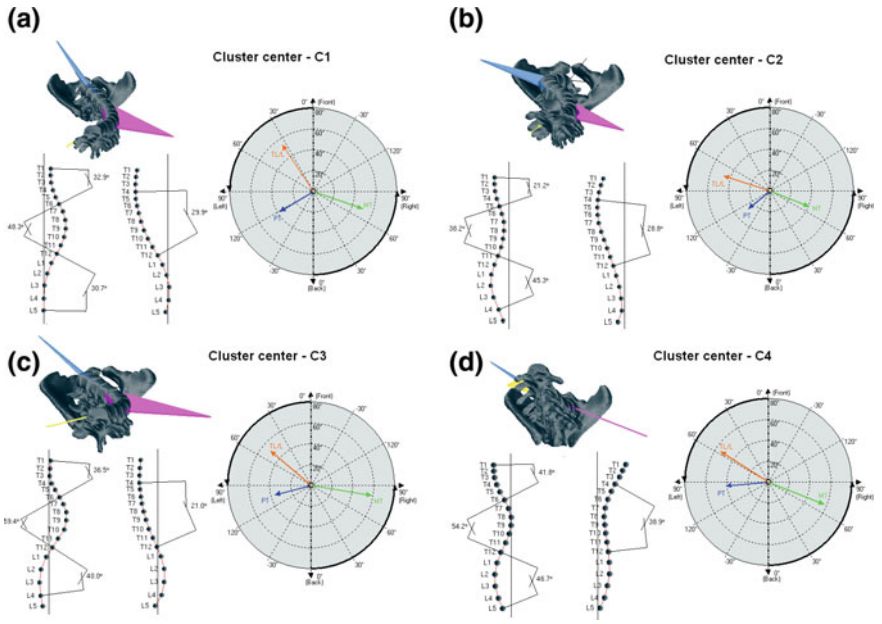


Fig. 15 Frontal, lateral and top view profiles with planes of maximal deformity (*PMC*) of the cluster centers for the four detected clusters. The respective da Vinci representations are also shown with corresponding cases. **a** Cluster center from *C1*. **b** Cluster center from *C2*. **c** Cluster center from *C3*. **d** Cluster center from *C4*

6 Conclusion

In this chapter, we presented a comprehensive framework for the three dimensional reconstruction of spines from radiographs. This first section included the calibration of the X-ray scene using a data-driven approach which extracted matched features between the biplanar images to estimate the geometric parameters of the setup. This was followed by a hybrid statistical and image-based approach to generate a 3D model of the spine from these calibrated X-rays.

The self-calibration approach uses high level shape primitives extracted from the natural content of the images, such as the spine silhouettes, to establish a reliable correspondence between the pair of X-ray images, and furthermore determine the 2D-3D relationship of the radiographic scene. Geometrical-based features were proposed to optimize the spine correspondences on the biplanar radiographs in order to obtain a global calibration of the acquisition system. The results confirm that using intensity, surface and geometrical-based components correlated with prior knowledge information enables the segmentation of the spine's global shape on the X-ray images. Results have shown that these high-level primitives help to automatically self-calibrate the radiographic setup by using representative shape related features such as the visual hull reconstruction, and are a viable and accurate procedure for the 3D reconstruction of the spine. The proposed automatic technique allows generating more accurate 3D vertebra shapes compared to the manual identification and matching of landmarks performed by an operator based on epipolar geometry. Furthermore, an image-based calibration technique incorporates information on orientation features which were previously unavailable. While the accuracy of the method is promising for the extraction of meaningful 3D clinical data, errors may be propagated from the segmentation of the spine silhouettes or from patient motion between the sequential biplanar acquisitions which can affect the convergence and final accuracy of the geometrical parameters. Still, the approach allows the automatic X-ray calibration for the 3D reconstruction of scoliotic spines, which was difficult to perform with previous methods that require a calibration object and manually identified landmarks.

The hybrid statistical and image-based 3D reconstruction approach presented in this chapter was anchored on the statistical distribution of a scoliotic population and automatically segment scoliotic vertebrae using 3D level set surface evolution techniques from enhanced biplanar images. The proposed method offers a more reliable approach to this problem by integrating statistical, image-based and morphological knowledge, and therefore becomes a suitable tool for clinical assessment of spinal deformities. The proposed approach presented a sufficient level reliability to correctly detect the rotation and location of scoliotic vertebrae so it can be used in clinical trials. The method presented in this chapter generates models similar to those obtained from manual identification. However, the manual approach is a tedious and error prone procedure and does not guaranty 100 % accuracy. Therefore the differences exhibited in the experiment may come from the identification errors provided from the manual landmarking. Furthermore, the proposed framework uses

locally linear embeddings (LLE) to map the high-dimensional observation data from the spine model database, that are presumed to lie on a non-linear manifold, onto a single global coordinate system of lower dimensionality. LLE preserves neighbourhood relationships of similar spine geometries, thereby revealing the underlying structure of the data such as spine classification. Dimensionality reduction by LLE succeeds in recovering the underlying manifold, whereas linear embedding methods, such as Principal Component Analysis (PCA) or Multi-Dimensional Scaling (MDS), would map various data points to nearby points in the plane, creating distortions both in the local and global geometry.

Future work in the field will look at extending these frameworks by enforcing shape, texture, spatial and neighbourhood relations between the structures to increase the reliability and repeatability of the segmentation approaches. Other efforts are made to extend these techniques to post-operative cases. The methods presented in this chapter can also be extended to other medical reconstruction applications such as for the pelvis or femur, when a sufficient amount of prior data is available to adequately model various types pathologies.

Acknowledgments We would like to acknowledge the contributions of F. Cheriet and H. Labelle in this research. Research funding was supported in part by the Fonds Quebecois de la Recherche sur la Nature et les Technologies grants, the MENTOR program from the Canadian Institutes of Health Research and the Canada Research Chairs.

References

1. Andre B, Dansereau J, Labelle H (1994) Optimized vertical stereo base radiographic setup for the clinical three-dimensional reconstruction of the human spine. J Biomech 27(8):1023–1035
2. Aubin CE, Dansereau J, Parent S, Labelle H, de Guise JA (1997) Morphometric evaluations of personalised 3D reconstructions and geometric models of the human spine. Med Biol Eng Comput 35:611–618
3. Benameur S, Mignotte M, Labelle H, de Guise JA (2005) A hierarchical statistical modeling approach for the unsupervised 3-D biplanar reconstruction of the scoliotic spine. IEEE Trans Biomed Eng 52:2041–2057
4. Boisvert J, Cheriet F, Pennec X, Labelle H, Ayache N (2008) Geometric variability of the scoliotic spine using statistics on articulated shape models. IEEE Trans Med Imag 27:557–568
5. Charpak G (1981) Particle detection. Recherche 12(128):1384–1396
6. Cheng H, Bouman CA (2001) Multiscale Bayesian segmentation using a trainable context model. IEEE Trans Imag Process 10:511–525
7. Cheriet F, Dansereau J, Petit Y, Aubin CE, Labelle H, de Guise JA (1999) Towards the self-calibration of a multiview radiographic imaging system for the 3D reconstruction of the human spine and rib cage. J Pattern Recog Artif Intell 13:761–779
8. Cheriet F, Laporte C, Kadoury S, Labelle H, Dansereau J (2007) A novel system for the 3-D reconstruction of the human spine and rib cage from biplanar X-ray images. IEEE Trans Biomed Eng 54:1356–1358
9. Cook L, De Smet A, Tarlton MA, Fritz S (1981) Assessment of scoliosis using three-dimensional analysis. IEEE Trans Biomed Eng 28(4):366–371
10. Dansereau J, Stokes IA (1988) Measurements of the three-dimensional shape of the rib cage. J Biomech 21:893–901

11. Dansereau J, Beauchamp A, Guise JD (1990) Three-dimensional reconstruction of the spine and rib cage from stereoradiographic and imaging techniques. In: 16th conference of the Canadian society of mechanical engineering
12. Delorme S, Petit Y, de Guise JA, Labelle H, Aubin CE, Dansereau J (2003) Assessment of the 3-D reconstruction and high-resolution geometrical modeling of the human skeletal trunk from 2-D radiographic images. IEEE Trans Biomed Eng 50:989–998
13. Despres P, Beaudoin G, Gravel P (2005) Evaluation of a full-scale gas microstrip detector for lowdose X-rayimaging. Nucl Instrum Methods Phys Res A 536:52–60
14. Dumas R, Blanchard B, Carlier R, de Loubresse CG, Huec JCL, Marty C, Moinard M, Vital JM (2008) A semi-automated method using interpolation and optimisation for the 3D reconstruction of the spine from bi-planar radiography: a precision and accuracy study. Med Biol Eng Comput 46:85–92
15. Duong L, Cheriet F, Labelle H (2006) Three-dimensional classification of spinal deformities using fuzzy clustering. Spine 31:923–930
16. Fallavollita P, Cheriet F (2008) Optimal 3D reconstruction of coronary arteries for 3D clinical assessment. Comput Med Imag Graph 32:476–487
17. Fleute M, Lavallee S (1999) Nonrigid 3-D/2-D registration of images using statistical models. In: Proceedings of MICCAI, vol 38, pp 138–147
18. Ghebreab S, Smeulders AWM (2004) Combining strings and necklaces for interactive three-dimensional segmentation of spinal images using an integral deformable spine model. IEEE Trans Biomed Eng 51:1821–1829
19. Guo-Qing W, JianZhong Q, Schramm H (2001) Generalized dynamic programming approaches for object detection: detecting spine boundaries and vertebra endplates. In: Proceedings of IEEE conference on computer vision pattern recognition, vol 1, pp 954–959
20. Hamadeh A, Cinquin P (1997) Kinematic study of lumbar spine using functional radiographies and 3D/2D registration. Lect Notes Comput Sci 1205:109–118
21. Hartley RI (1997) Self-calibration of stationary cameras. Int J Comput Vision 22:5–23
22. Hartley R, Zisserman A (2000) Three-dimensional computer vision. Cambridge University Press, Cambridge
23. Howe B, Gururajan A, Sari-Sarraf H, Long LR (2004) Hierarchical segmentation of cervical and lumbar vertebrae using a customized generalized Hough transform and extensions to active appearance models. In: Proceedings of IEEE symposium image analysis interpretation, pp 182–186
24. Huynh T, Dansereau J, Maurais G (1997) Development of a vertebral endplate 3-D reconstruction technique. IEEE Trans Med Imaging 16(5):689–696
25. Illés T, Somoskeoy S (2013) Comparison of scoliosis measurements based on three-dimensional vertebra vectors and conventional two-dimensional measurements: advantages in evaluation of prognosis and surgical results. Eur Spine J 22:1255–1263
26. Kadoury S, Cheriet F, Labelle H (2009) Personalized X-ray 3D reconstruction of the scoliotic spine from statistical and image models. IEEE Trans Med Imag 28:1422–1435
27. Kadoury S, Cheriet F, Labelle H (2010) Self-calibration of biplanar radiographic images through geometric spine shape descriptors. IEEE Trans Biomed Eng 57(7):1663–1675
28. Kadoury S, Cheriet F, Laporte C, Labelle H (2007) A versatile 3-D reconstruction system of the spine and pelvis for clinical assessment of spinal deformities. Med Biol Eng Comput 45:591–602
29. Kadoury S, Labelle H (2012) Classification of three-dimensional thoracic deformities in adolescent idiopathic scoliosis from a multivariate analysis. Eur Spine J 21:40–49
30. Kadoury S, Cheriet F (2006) X-ray image restoration with adaptive PDE filter for an accurate 3D reconstruction of the human spine. In: Proceedings of international conference on computer assisted radiology surgery, vol 1, p 470
31. Kadoury S, Cheriet F, Labelle H (2008) A statistical image-based approach for the 3D reconstruction of the scoliotic spine from biplanar radiographs. In: Proceeding of IEEE international symposium on biomedical imaging, vol 1, pp 660–663

32. Kadoury S, Cheriet F, Labelle H (2009) Segmentation of scoliotic spine silhouettes from enhanced biplanar X-rays using a prior knowledge Bayesian framework. In: Proceedings of IEEE international symposium on biomedical imaging, vol 1, pp 478–481
33. Kaminsky J, Klinge P, Rodt T, Bokemeyer M, Luedemann W, Samii M (2004) Specially adapted interactive tools for an improved 3D-segmentation of the spine. Comput Med Imag Graph 28:119–127
34. King H, Moe J, Bradford D et al (1983) The selection of fusion levels in thoracic idiopathic scoliosis. J Bone Joint Surg Am 65:1302–1313
35. Klinder T, Wolz R, Lorenz C, Franz A, Ostermann J (2008) Spine segmentation using articulated shape models. In: Proceedings of MICCAI, vol 11, pp 227–234
36. Labelle H, Dansereau J, Bellefleur C, Poitras B (1996) Three-dimensional effect of the boston brace on the thoracic spine and rib cage. Spine 21:59–64
37. Labelle H, Dansereau J, Bellefleur C, Poitras B, Rivard CH, Stokes IA, de Guise JA (1995) Comparison between preoperative and postoperative three-dimensional reconstructions of idiopathic scoliosis with the cotrel-dubousset procedure. Spine 20:2487–2492
38. Laporte S, Skalli W, de Guise JA, Lavaste F, Mitton D (2003) A biplanar reconstruction method based on 2D and 3D contours: application to the distal femur. Comp Meth Biomech Biomed Eng 6:1–6
39. Laurentini A (1994) The visual hull concept for silhouette-based image understanding. IEEE Trans Pattern Anal Mach Intell 16:150–162
40. Li G, Zucker S (2006) Contextual inference in contour-based stereo correspondence. J Comput Vis 69:59–75
41. Livyatan H, Yaniv Z, Joskowicz L (2003) Gradient-based 2-D/3-D rigid registration of fluoroscopic X-ray to CT. IEEE Trans Med Imag 22:1395–1406
42. Lorensen W, Cline H (1988) Marching cubes: a high resolution 3-D surface construction algorithm. Comput Graph 4:163–169
43. Lorenz C, Krahnstoever N (2000) Generation of point-based 3D statistical shape models for anatomical objects. Comput Vis Image Underst 77:175–191
44. Marzan GT (1976) Rational design for close-range photogrammetry. Ph.D. thesis, University of Illinois, Department of Civil Engineering, Urbana-Champaign
45. Mitton D, Landry C, Veron S, Skalli W, Lavaste F, de Guise JA (2000) 3D reconstruction method from biplanar radiography using non-stereocorresponding points and elastic deformable meshes. Med Biol Eng Comput 38:133–139
46. Muggleton JM, Allen R (1997) Automatic location of vertebrae in digitized videofluoroscopic images of the lumbar spine. Med Eng Phys 19:77–89
47. Novosad J, Cheriet F, Delorme S, Poirier S, Beausejour M, Labelle H (2002) Self-calibration of biplanar radiographs for a retrospective comparative study of the 3D correction of adolescent idiopathic scoliosis. Stud Health Technol Inform 91:272–275
48. Panjabi M, Chang D, Dvorak J (1992) An analysis of errors in kinematic parameters associated with in vivo functional radiographs. Spine 17:200–205
49. Paragios N, Deriche R (2002) Geodesic active regions: new paradigm to deal with frame partition problems in CV. Vis Comm Image Repre 13:249–268
50. Pomero V, Mitton D, Laporte S, de Guise JA, Skalli W (2004) Fast accurate stereoradiographic 3D-reconstruction of the spine using a combined geometric and statistic model. Clin Biomech 19:240–247
51. Poncet P, Dansereau J, Labelle H (2001) Geometric torsion in idiopathic scoliosis: three-dimensional analysis and proposal for a new classification. Spine 26:2235–2243
52. Qiang J, Costa MS, Haralick RM, Shapiro LG (1999) An integrated linear technique for pose estimation from different geometric features. In: Proceedings of vision interface, pp 705–733
53. Rillardon L, Campana S, Mitton D, Skalli W, Feydy A (2005) Evaluation of the intervertebral disc spaces with a low dose radiographic system. J Radiol 86(3):311–319
54. Roweis ST, Saul LK (2000) Nonlinear dimensionality reduction by locally linear embedding. Science 290:2323–2326

55. Sangole A, Aubin C, Labelle H et al (2009) Three-dimensional classification of thoracic scoliotic curves. Spine 34:91–99
56. Simonis C, Allen R (1993) Calculation of planar spine kinematic parameters using videofluoroscopic images and parallel computation. In: Proceedings of conference on engineering in medicine and biology, pp 1087–1088
57. Smyth PP, Taylor CJ, Adams JE (1997) Automatic measurement of vertebral shape using active shape models. Image Vis Comput 15:575–581
58. Somoskeoy S, Tunyogi-Csapo M, Bogyo C et al (2012) Clinical validation of coronal and sagittal spinal curve measurements based on three-dimensional vertebra vector parameter. Spine 12:960–968
59. Stokes I, Bigalow L, Moreland M (1987) Three-dimensional spinal curvature in idiopathic scoliosis. J Orthop Res 5:102–113
60. Tomazevic D, Likar B, Slivnik T, Pernus F (2003) 3-D/2-D registration of CT and MR to X-ray images. IEEE Trans Med Imag 22:1407–1416
61. Villemure I, Aubin CE, Grimard G, Dansereau J, Labelle H (2001) Progression of vertebral and spinal three-dimensional deformities in adolescent idiopathic scoliosis: a longitudinal study. Spine 26:2244–2250
62. Vrtovec T, Ourselin S, Gomes L, Likar B, Pernus F (2007) Automated generation of curved planar reformations from MR images of the spine. Phy Med Biol 52:2865–2878
63. Vrtovec T, Tomazevic D, Likar B, Travnik L, Pernus F (2004) Automated construction of 3D statistical shape. Image Anal Stereo 23:111–120
64. Yalin Z, Nixon MS, Allen R (2004) Automated segmentation of lumbar vertebrae in digital videofluoroscopic images. IEEE Trans Med Imag 23:45–52
65. Yezzi AJ, Soatto S (2003) Stereoscopic segmentation. J Comput Vis 53:31–43

Vertebral Column Localization, Labeling, and Segmentation

Raja S. Alomari, Subarna Ghosh, Jaehan Koh and Vipin Chaudhary

Abstract The vertebral column consists of interconnected bone structures that extend from the neck down to the pelvis. In addition to its crucial functionality in spinal cord protection, it provides the necessary flexibility and support for the whole body. Worldwide interest in spine related research has been increasing due to the widely spread of related abnormalities in the developed countries which accounts for over $100 billion annually in the diagnosis, treatment, and associated loss of wages. Our specific interest in this chapter is in the medical image analysis of the vertebral column. In this chapter, we aim at providing a broader review of the available literature in vertebral column image analysis. Moreover, we focus on providing an understanding of the localization, labeling, and segmentation problems for the various vertebral column structures from the available medical imaging modalities. Additionally, we describe the general challenges facing the various solutions for these problems. Our taxonomy is based on the target imaging modality to simplify the understanding of the broad research in this area.

R.S. Alomari (✉)
The University of Jordan, Amman, Jordan
e-mail: ralomari@buffalo.edu

S. Ghosh · V. Chaudhary
The University at Buffalo SUNY, Buffalo, NY, USA
e-mail: sghosh7@buffalo.edu

V. Chaudhary
e-mail: vipin@buffalo.edu

J. Koh
Samsung Inc., Suwon, South Korea
e-mail: jkoh@buffalo.edu

© Springer International Publishing Switzerland 2015

S. Li and J. Yao (eds.), *Spinal Imaging and Image Analysis*,
Lecture Notes in Computational Vision and Biomechanics 18,
DOI 10.1007/978-3-319-12508-4_7

1 Introduction

Localization, labeling, and segmentation of the vertebrae and the intervertebral discs are essential tasks that have been attracting an increasing number of research groups worldwide. The accuracy and robustness of these imaging tasks are crucial for subsequent abnormality diagnosis. Moreover, accurate results of these tasks are critical for radiologists to perform an accurate diagnosis from various imaging modalities including X-ray radiography, Computed Tomography (CT) scans, and Magnetic Resonance Imaging (MRI). Furthermore, surgeons demand accurate reporting of these results when overlaid on a computer guided surgery system or a computer assisted surgery system.

Whilst the localization task is to locate an anatomical structure (e.g. locating the intervertebral discs by a point within or a bounding box around the discs), the segmentation task is to provide a fine contour that accurately delineates that structure (e.g. a contour around the vertebra). Labeling, on the other hand, is to identify the anatomical nomenclature of each structure (e.g. labeling each of the five lumbar vertebrae as L1, L2, L3, L4 and L5). Figure 1 shows an example of localization and labeling for the six intervertebral discs connected to the five lumbar vertebrae on a sagittal MRI [5].

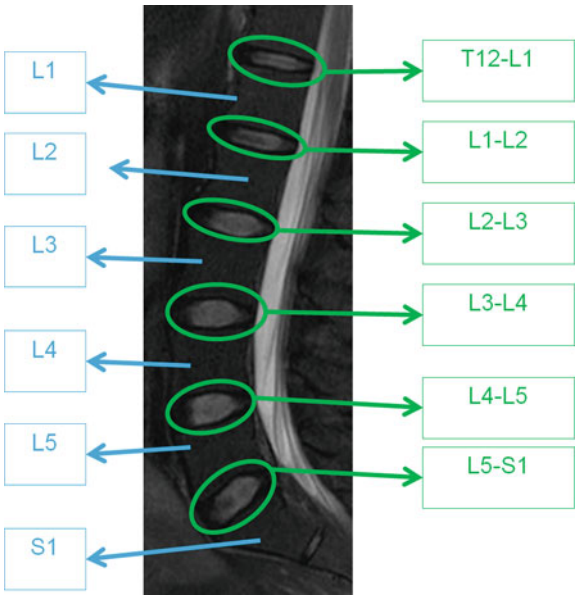


Fig. 1 Localization and labeling of a sagittal lumbar T2-weighted MRI. Lumbar area is the second area to the last of the vertebral column. It is the main part of the vertebral column that is responsible for bearing the major body weight. The lowest lumbar vertebra is *L5* and the highest is *L1*. Intervertebral discs are labeled based on the enclosing vertebrae [5]

The essential structures within the vertebral column that have been attracting researchers for localization, labeling, and segmentation are the intervertebral discs, the vertebrae and the *Dural Sac*.

The lumbar vertebrae are the five vertebrae between the rib cage and the pelvis which are designated as L1–L5, starting at the top. The intervertebral discs are fibrocartilaginous cushions which are named upon the vertebral bodies that sandwich a particular disc, e.g., the disc in between L1 and L2 is named L1-L2. In clinical practice, the radiologist reports the diagnosis at each disc level and at each vertebra level. Hence, the first requirement of any lumbar Computer Aided Diagnosis (CAD) system is to localize and label the lumbar discs and vertebrae as shown in Fig. 1 [5]. Specifically, localization refers to providing centroids or bounding boxes for each of the lumbar discs, and labeling refers to identifying each localized disc as one of the six lumbar discs (T12-L1, L1-L2, L2-L3, L3-L4, L4-L5, L5-S1). While some researchers have discussed methods to provide a point within each lumbar disc [5, 76], there are also methods [32] that provide a bounding box around every visible disc in clinical lumbar MRIs as illustrated in Fig. 2.

Another important tissue structure in the lumbar spine is the *Dural Sac*. It is the membranous sac that encases the spinal cord within the bony structure of the

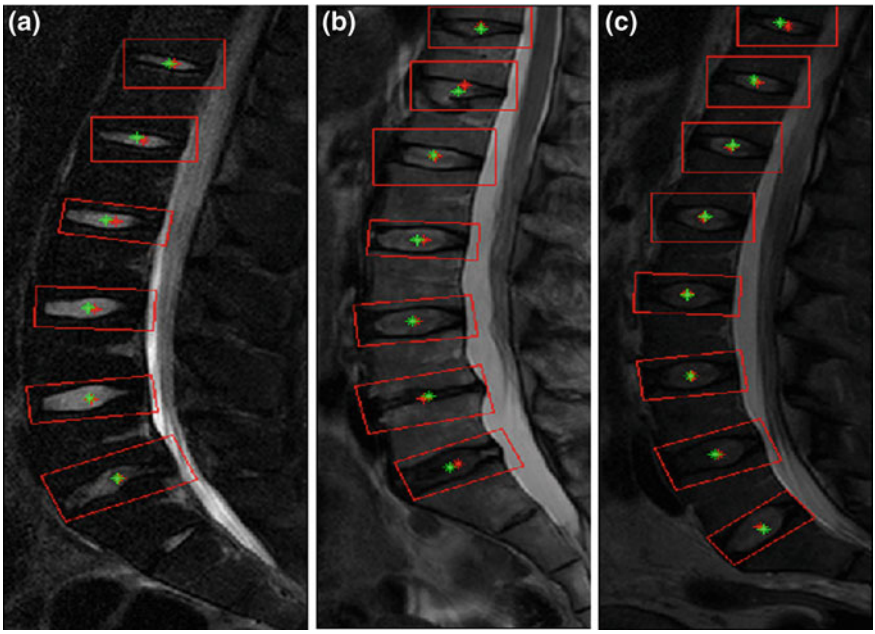


Fig. 2 This figure illustrates the results of an automatic lumbar disc localization method [32] which detects all the visible discs in a lumbar MRI. In case more than six discs are detected, the lower most six discs are identified as the lumbar discs. The *red boxes* are the bounding boxes provided for each of the visible discs in the clinical MRI. The *red stars* show the automatic disc centers, while the *green stars* show the true centers

vertebral column. The human spinal cord extends from the foramen magnum and continues through to the conus medullaris near the second lumbar vertebra, terminating in a fibrous extension known as the filum terminale. The *Dural Sac* usually ends at the vertebral level of the second sacral vertebra. Intensity inhomogeneity within the sac due to varying amounts of white and gray matter makes the segmentation of the *Dural Sac* and the spinal cord very challenging. Moreover, automatic segmentation in clinical MRIs is even harder due to variations in appearance and a lack of bright spinal fluid in cases with certain abnormalities such as stenosis.

After localization and labeling, comes the challenging task of tissue segmentation. Segmentation of discs is quite difficult due to extreme variability in shape, size and appearance of intervertebral discs in lumbar MRI. Moreover, discs with abnormalities can be very fuzzy and difficult to segment manually leading to

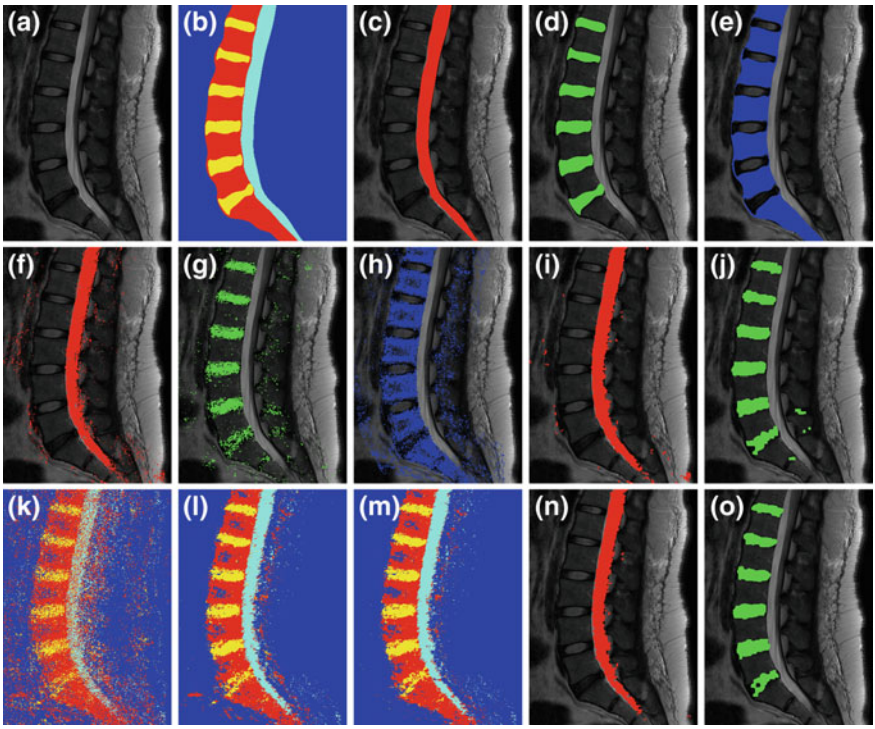


Fig. 3 Illustration of automated lumbar tissue segmentation [34]: **a** shows the original mid-sagittal MRI, **b–e** show the manual segmentation (ground truth), **f–h** show the label maps for the dural sac, disc and vertebra respectively using method 1 (probability map + HOG features), while **i** and **j** show the dural sac and disc segmentation after morphological post processing. **k–m** Show the label maps generated at the end of iteration number 1, 6 and 200 respectively using method 2 (probability map + HOG features + neighborhood labels via Gibbs sampling), while **n** and **o** show the dural sac and disc segmentation after morphological post processing

significant inter-observer variability. At the same time, segmentation of intervertebral discs is a very important part of lumbar CAD systems in order to diagnose and quantify abnormalities such as herniation, desiccation and degeneration.

Requirements for CAD systems of the lumbar region are unique since we need to segment the *Dural Sac* and localize, label and segment the lumbar intervertebral discs before we can initiate the diagnosis. Figure 3 shows an illustration of automated segmentation [34] of the discs, vertebrae and the *Dural Sac* of a clinical MRI using two methods, the first using a probability map and HOG features, while the second method uses neighborhood label information as well, in a Gibbs Sampling approach.

2 The Vertebral Column

This section is dedicated to present the anatomy of the vertebral column in general with focus on the lumbar area. It also provides the standardized nomenclature of the various abnormalities in the vertebral column as endorsed by the North American Spine Society (NASS), the American Society of Spine Radiology (ASSR), and the American Society of Neuroradiology (ASNR) [24].

The vertebral column, also known as the backbone or the spinal column, is typically made up of (33) individual bones called vertebra (*plural*: vertebrae) that interlocks with each other. These vertebrae are classified into five areas from top to bottom: *Cervical* (7), *Thoracic* (12), *Lumbar* (5), *Sacral* (5), and *Coccyx* (4). Among these (33) vertebrae, only the top (24) are movable due to which clinicians often state that the vertebral column consists only of (26) vertebrae counting the Sacral vertebrae as one and the Coccyx as one. In each of these four regions, the vertebrae have unique features that allows certain functionality [88].

There are five distinct regions in the vertebral column. The top most region is the *Cervical* region which consists of seven vertebrae anatomically named from top to bottom as C1 (also called Atlas), C2 (also called Axis), C3, C4, C5, C6, and C7. The main function of this region is to support the weight of the head [normally weighs about 10 pounds (4.5 kg)]. The cervical has the most range of motion due to the first two specialized vertebrae that connect to the skull.

The *Thoracic* region comes next and consists of twelve vertebrae named from top to bottom as T1, T2, T3, T4, T5, T6, T7, T8, T9, T10, T11, and T12. The major role of the thoracic spine is to protect the organs that lie in the chest by supporting the rib cage. The motion is limited due to the nature of the chest.

The *Lumbar* region comes next and has the largest vertebrae. This region is responsible for the whole flexibility of the back as well as bearing the weight of the body. Five vertebrae exist in this area that are named (from top to bottom) as L1, L2, L3, L4, and L5. The spinal cord stops, typically at the L1-L2 area where the nerves hang down inside the *Pachymeninx* (*Dural Mater*) which is the tough and inflexible outermost of the three layers of the meninges surrounding the brain and spinal cord. The *Sacral* and the *Coccyx* regions have less functionality and they are

barely movable. The sacral one fused vertebra (named as S or S1) provides attachment for the *Ilium* (hip) bones and protects the pelvic organs while the coccyx region fused vertebra are not recognized for main functionality.

2.1 Lumbar Spine Region

Since our focus is on the lumbar region, we present more details on the anatomy of the lumbar region as well as details of the MRI. There are two common anatomic terms that relate to the low back: anterior and posterior. Anterior refers to the front of the spine while posterior refers to back of the spine as shown in Figs. 4 and 5. The section of the spine that makes up the low back is called the lumbar spine. Lumbar spine includes these main structures: Intervertebral discs, vertebrae, and other structures.

2.1.1 Intervertebral Discs

Intervertebral discs are unique structures that absorb shocks between adjacent vertebrae. They act as the ligaments that connect the vertebrae together and the pivot point which allows the spine mobility by bending and rotating. They make about one fourth of the spinal column length [88].

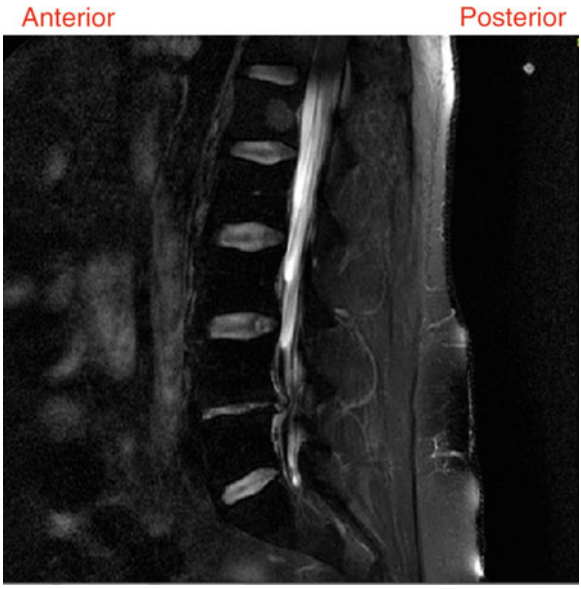


Fig. 4 Sagittal T2-SPIR-weighted MRI showing anterior and posterior terms [5]

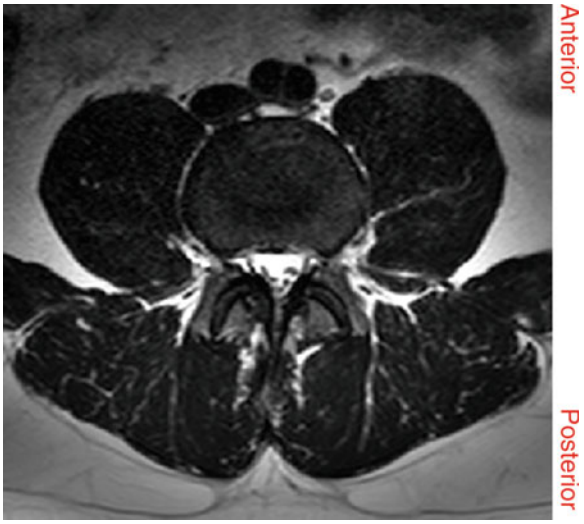


Fig. 5 Axial T1-weighted MRI showing anterior and posterior terms [5]

An intervertebral disc is composed of two parts: an outer strong ring called *Annulus Fibrosus* and a soft gel-like inner called *Nucleus Pulposus*. The nucleus pulposus consists of 80–85 % water in normal cases. By aging the disc dehydrates limiting its ability to absorb shocks. The outer rings gets weaker as well and start having tears that causes various abnormalities. The bottom up view is the anterior while top down is the posterior direction. Figures 6 and 7 show T1- and T2-weighted MRI for the same lumbar disc from our dataset, respectively.

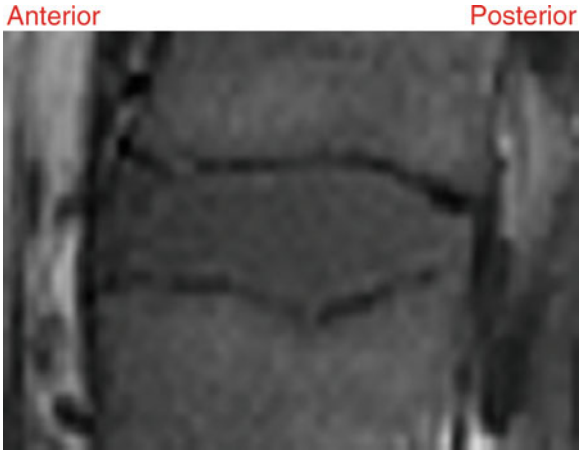


Fig. 6 Sagittal T1 MRI for L1-L2 disc [5]

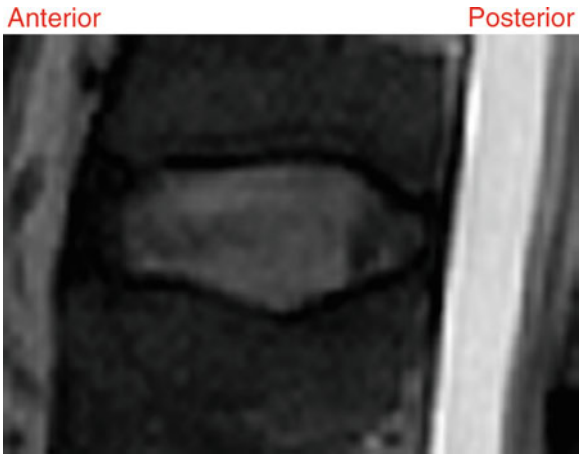


Fig. 7 Sagittal T2 MRI for L1-L2 disc [5]

2.1.2 Vertebrae

A vertebra (*plural*: vertebrae) is a bone with specific structure that supports and protects the spinal cord. A typical vertebra consists of two segments: Anterior (front) and posterior (back). The anterior part of the vertebra is the body while the posterior, which is known also as the vertebral (neural) arch, includes: the vertebral foramen, a pair of pedicles and a pair of laminae, and supports seven processes [four articular, two transverse, and one spinous (*aka* neural spine)].

2.1.3 Other Structures

In addition to the main structures within the vertebral column, there are few other structures including.

Nerves: The spinal cord hangs inside a bony ring through the vertebral column that is made up of millions of nerve fibers. The spinal cord extends down to the L2 vertebra. Below L2, a bundle of nerves named as *Cauda Equina* hangs down in what is known as the *Thecal Sac*. Two large nerves branch off the spinal cord, one from each side passing through the neural foramina of each vertebra. These spinal nerves group together to form the main nerves that go to the organs and limbs. The nerves of the lumbar spine (*Cauda Equina*) go to the pelvic organs and lower limbs [88].

Connective tissues: They are the fibrous connections that hold the cells of the body together. The ligaments are strong connective tissues that attach bones together. There are many long ligaments that connect on the front and back sections of the vertebrae. The anterior longitudinal ligament runs lengthwise down the front of the vertebral bodies. Two other ligaments run full-length within the spinal canal. The posterior longitudinal ligament attaches on the back of the vertebral bodies. The *Ligamentum Flavum* is a long elastic band that connects to the front surface of

the lamina bones (just behind the spinal cord). Thick ligaments also connect the bones of the lumbar spine to the sacrum (the bone below L5) and pelvis.

Muscles in the lower back are arranged in layers. The superficial layer, the closest to the skin, is covered by a thick tissue called *Fascia*. The middle layer, called the *Erector Spinae*, has strap-shaped muscles that run up and down over the lower ribs, chest, and low back. They join in the lumbar spine to form a thick tendon that binds the bones of the low back, pelvis, and sacrum. The deepest layer of muscles attaches along the back surface of the spine bones, connecting the low back, pelvis, and sacrum. These deepest muscles coordinate their actions with the muscles of the abdomen to help hold the spine steady during activity [95].

Spinal segments is a notion that includes two vertebrae separated by an intervertebral disc, the nerves that leave the spinal column at each vertebra, and the small facet joints that link each level of the spinal column. The intervertebral disc separates the two vertebral bodies of the spinal segment. The disc normally works like a shock absorber. It protects the spine against the daily pull of gravity. It also protects the spine during heavy activities that put strong force on the spine, such as jumping, running, and lifting. The spinal segment is connected by two facet joints described earlier. When the facet joints of the lumbar spine move together, they bend and turn the low back [95].

3 Popular Lumbar Imaging Modalities

Most image based research literature focuses on X-ray radiography, Dual-energy X-ray Absorptiometry (DEXA or DXA), CT, and MRI. X-ray radiography and DEXA are cheaper and widely popular modalities as an initial diagnostic tool. Hence, the availability of the data provided researchers with great opportunities to investigate labeling, localization, and even diagnosis problems.

On the other hand, MRI (Fig. 8) and CT (Fig. 9) are more expensive and less available for researchers. Hence, fewer researchers obtained access to such data and were able to investigate localization, segmentation, and diagnosis problems on the various anatomical structures. Few efforts utilized other modalities such as ultrasound, especially, for fetal spine detection and abnormality detection.

Both X-ray and DEXA (*aka* DXA) radiography consist only from one 2D slice that shows the area of interest. On the other hand, CT scans show a full 3D volume for the area of interest. Clinical CT spans the whole area in slice-by-slice fashion that can be directly used to produce a full 3D volume. Usually CT consists of a set of axial slices with specific thickness depending on the available technology.

Moreover, clinical MRI consists of few protocols that vary depending on the available technology. The current standard in MRI in North America, for low back, is the 3 T MRI. Most of the current MRI radiology centers produce: (1) T1-Weighted sagittal (T1 W-sagittal), (2) T2-Weighted sagittal (T2 W-sagittal), (3) T2-Weighted axial (T2 W-axial) for a set of selected discs, (4) T2-Weighted axial (T2 W-axial) and (5) Myelo MR images (Fig. 8). While the sagittal views span the

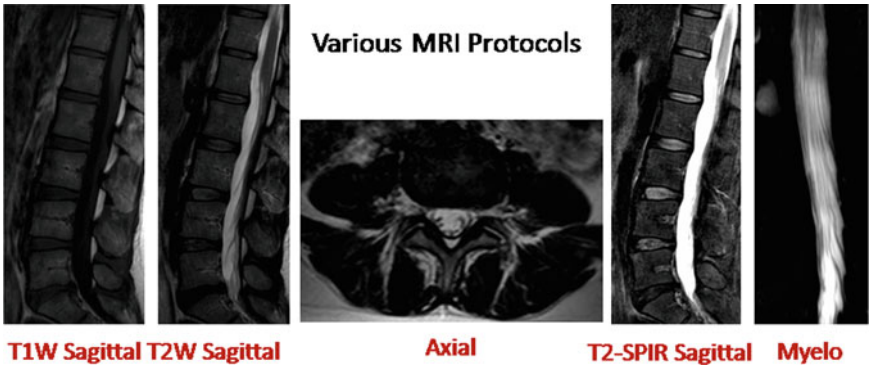


Fig. 8 This figure illustrates a sample image from each of the five popular clinical MRI protocols —T1 weighted sagittal, T2 weighted sagittal, T2 weighted axial, T2-SPIR sagittal and Myelo [6]

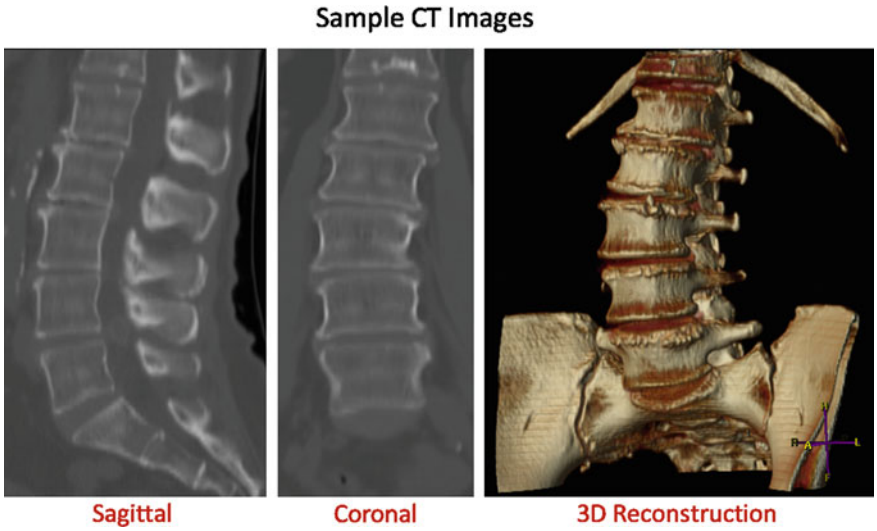


Fig. 9 This figure shows samples of a lumbar CT scan. The *first two images* show sagittal and coronal slices from a CT volume, while the *last image* shows a 3D reconstruction [2]

side-to-side dimension of the body, the axial views are acquisitions of each intervertebral disc within the area of interest. Each disc has a set of axial slices that are aligned with the dimension of the major axis of the disc. MRI acquisition technician spends a manual effort for planning the acquisition to make sure that each disc volume is acquired correctly and that all acquired protocols are manually co-registered. The patient is not allowed to move during the whole acquisition period.

Many clinical MRI protocols exist that have trade-offs in diagnosis of various backbone abnormalities. The technician has four main parameters to tune before MRI acquisition that control the appearance (intensity) of the resulting image: (1) proton density, (2) longitudinal relaxation time (T1), (3) transverse relaxation time (T2), and (4) the flow. The proton density refers to the concentration of protons in the tissue in the form of water and macromolecules (proteins, fat, etc.). Both T1 and T2 relaxation times define the way that the protons revert back to their resting states after the initial RF pulse. The most common effect of the flow is the loss of signal from rapidly flowing arterial blood.

Two common pulse sequences for MR imaging are widely used: T1- and T2-weighted spin-echo sequences. The T1-weighted sequence uses a short Repetition Time (TR) and a short TE (Echo Time) (TR $\leq$ 1,000 ms, TE $\leq$ 30 ms). The T2-weighted sequence uses a long TR and long TE (TR $\geq$ 2,000 ms, TE $\geq$ 80 ms). Moreover, two major techniques are used for suppression of fat signals in MRI: Short Tau Inversion Recovery (STIR) and selective partial inversion recovery (SPIR). The STIR sequences suppress fat signal by using an initial 180° radiofrequency pulse to invert the longitudinal magnetization. Image acquisitions are then performed with the inversion time equivalent to the known null point for fat (approximately 0.69 $\times$ T1) [92]. SPIR is a more recent fat-suppression technique that is based on the use of frequency-specific pulse sequences [93]. Only the fat magnetization pulse is inverted leaving water resonances as is. This technique is useful for suppressing any tissue-specific pulse given the known-frequency of that tissue. However, the SPIR technique is extremely sensitive to the magnetic field inhomogeneity. SPIR is used with both T1- and T2-weighted MRI [44].

Furthermore, another important MRI sequence generator that is related to common current clinical MRI is called MR Myelography [57] (Myelo is a new Latin word, from Greek muelos, which means spinal cord). In this method, the background signal is suppressed by using heavily T2-weighted fast spin-echo pulse sequences and obliterating fat signal by pre-saturation. The resulting slices are then projected into a composite image using a standard maximum intensity projection (MIP) algorithm [57].

It is worth mentioning that inter-observer variability exist in lumbar diagnosis similar to many diagnosis tasks from various imaging modalities including X-ray radiographs, MRI, CT, Single-Photon Emission Computed Tomography (SPECT), and High Resolution (HR). However, MRI shows high inter-observer reliability compared to plain radiographs in lumbar area diagnosis (e.g., [62]). Mulconrey et al. [70] showed that abnormality detection for degenerative disc and Spondylolisthesis with MRI has $\kappa = 0.773$ and $\kappa = 0.728$, respectively, which is considered high in showing inter-observer reliability where this reliability is considered perfect when $0.8 \leq \kappa \leq 1$.

4 Challenges

Automatic detection of abnormalities from MRI or CT scans has been studied by researchers for quite some time. The challenges are manifold—ranging from variations in scanner specifications, parameter settings, modalities, differences in body structure and composition and last but not the least the task of segmentation which is a big challenge in computer vision.

In general, the segmentation of CT and MRI scans is difficult due to three main reasons.

(1) **Partial Volume Effect**: It is a scenario where multiple tissues contribute to pixels and blurs intensity across boundaries as illustrated in Fig. 10.
(2) **Intensity Inhomogeneity**: It is defined as non-anatomic intensity variations of the same tissue over the image, and may be caused by the imaging instrumentation (RF non-uniformity, static field inhomogeneity) or due to patient movement as seen in Fig. 11.
(3) **Intensity Similarity**: Very often two or more tissues have the same intensities in MRI scans as illustrated in Fig. 12.

All these factors contribute to the fact that segmentation of a lumbar MRI is a very challenging task.

Real world clinical MRIs are even more challenging since patients very often suffer from one or more lumbar abnormalities such as vertebral fractures, Spondylolysis, Spondylolisthesis, Scoliosis, intervertebral disc abnormalities (degeneration, desiccation, herniation, bulge and annular tears) and spinal Stenosis. In addition, there is lumbar variability due to patient age, height and structure leading to diverse images. Figure 13 shows a sample set of clinical lumbar MRIs showing some of the variability [5].

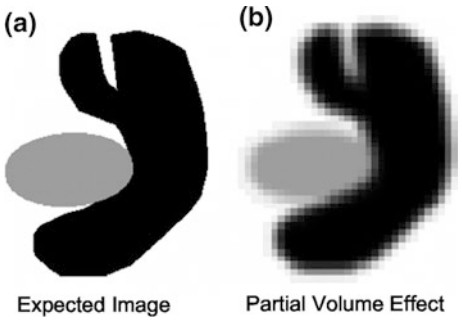


Fig. 10 This figure shows an illustration of partial volume effect in an imaginary scan consisting of two different kinds of tissues. While the *first image* shows the expected image the *second one* shows the actual image with fuzzy boundaries due to partial volume effect [32]

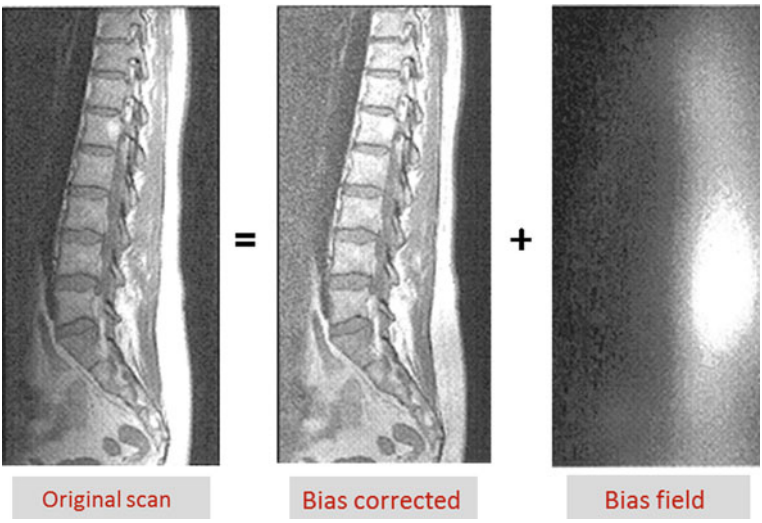


Fig. 11 This figure illustrates intensity inhomogeneiy in a Lumbar MRI [32]

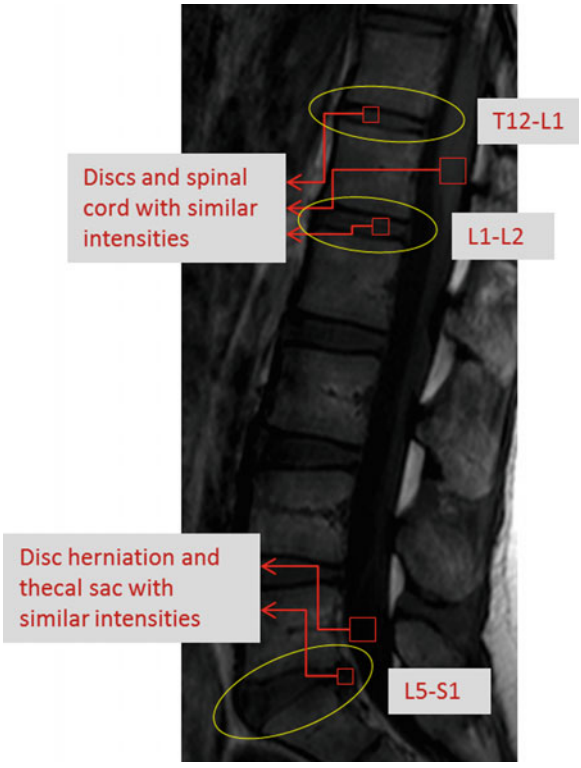


Fig. 12 This figure shows intensity similarity of different tissues in a T1 weighted lumbar MRI [32]

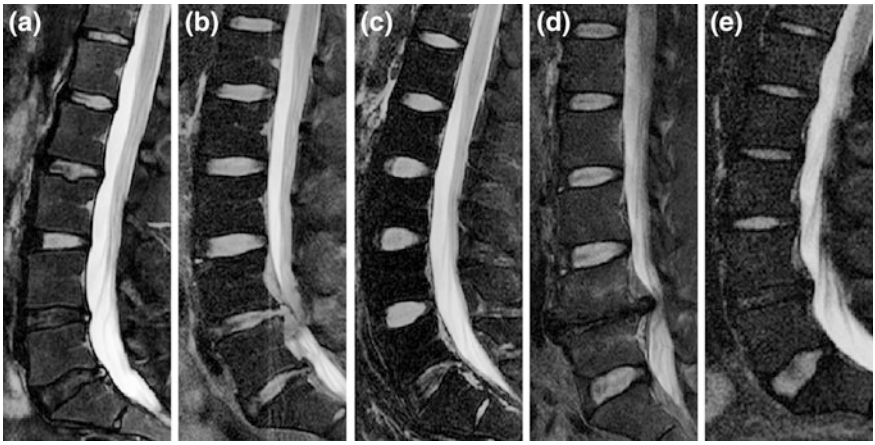


Fig. 13 Variability in disc appearances, shapes, locations, and sizes in different abnormal cases. **a** Shows variability in appearance of discs. The lower two discs (L4-L5 and L5-S1) have less intensity levels due to abnormalities (Herniation, Stenosis and Desiccation). **b** Shows variability in shape of discs with close intensity levels due to abnormalities in the lower two discs (Herniation and Stenosis). **c** Shows clear difference at the lowest disc level (L5-S1) as well as the difference in bending of the lumbar vertebral column which results in variability in location. **d** Shows variability in location of discs from other figures, sizes of discs, and the missing disc at L4-L5 disc. **e** Shows variability in disc sizes between the upper four discs and the lowest disc L5-S1. Ages of these patients are 35, 36, 29, 47, and 27, respectively from **a** to **e**. All images have been edited by cropping and contrast enhancement for better visualization

5 Advances in Localization, Labeling, and Segmentation

There are three main steps for the proper diagnosis in medical imaging: (1) Localization and labeling of anatomic structures, (2) Segmentation and (3) Diagnosis and quantification of abnormalities. There has been an extensive amount of work done in the area of vertebral body localization and segmentation from X-ray radiographs and CT scans in the past two decades. On the other hand, localization of soft tissues in MRI and diagnosis of disc abnormalities is comparatively more recent and has been of central focus for low back research in the last decade.

In this section, we review in detail, the current literature in the context of spinal tissue localization, segmentation and abnormality diagnosis. We classify the literature based on the medical imaging modality.

5.1 X-ray Radiography

Since X-ray radiographs use ionizing radiation and show better detailing of the bony tissues, there has been plenty of research in the direction of vertebrae segmentation from X-ray scans. Moreover, the availability of X-ray radiographs data helped boost the related amount of research.

5.1.1 Vertebrae

More than two decades ago, Hedlund and Gallagher [39] performed vertebral morphometry on lateral thoracic and lumbar X-ray radiographs of 153 women with a preliminary diagnosis of *Spinal Osteoporosis*. Measurements included anterior and posterior vertebral height, width, area, wedge angle, percent reduction of Anterior to Posterior Height (PRH) and Percent Difference in Anterior Height between adjoining vertebrae (PDAH). They showed that among individuals with mild Osteoporosis (0–2 fractures) PDAH identified 86 % of the fractures and 95 % of the individuals with fractures.

Manual selection of anatomical points for vertebral abnormality diagnosis is time consuming, imprecise and subjective. To obtain a more objective and accurate description of the vertebral body shape, semi-automatic methods were proposed that were based on statistical models of vertebral bodies in the sagittal view. Very early on, in 1993, a computerized quantifying technique for vertebral morphometry on lateral radiographs of the spine was proposed by Nicholson et al. [73]. Although fracture detection was improved by expanding the description of the vertebral body shape from six points to a contour, the amount of traumatic spinal injury or latent vertebral fracture was often underestimated. The main reason for the wrong diagnosis originated from the limited measurement possibility caused by the lack of depth perception in X-ray radiographs. Later in 1997, Smyth et al. [87] described how Active Shape Models (ASM) could be used to locate both normal and fractured vertebrae from Dual energy X-ray Absoptiometry (DXA) images of the spine. However, three initialization points have to be manually selected. To overcome the lack of depth perception in X-ray images, Benameur et al. [8] performed projection of a three-dimensional (3D) statistical shape model of a vertebra to a pair of orthogonal 2D X-ray radiographs. They validated this method on 57 scoliotic vertebrae images. However, the proposed segmentation was highly dependent on model initialization.

Various efforts that target the diagnosis of certain vertebra conditions involved localization and segmentation. In 2000, Long and Thoma [60] investigated the segmentation of C2 and C3 vertebrae from the cervical area using an ASM as a first step for building an image based retrieval system for a dataset consisting of 7,000 lumbar X-ray radiographs and 10,000 cervical spine X-ray radiographs. They built the Web-based Medical Information Retrieval System (WebMIRS) based on the National Health and Nutrition Examination Surveys (NHANES). Later,

Cherukuri et al. [16] proposed image processing techniques for computing size-invariant, convex hull-based features to highlight anterior Osteophytes. Feature evaluation of 714 lumbar spine vertebrae using a multi-layer perceptron yielded normal and abnormal average correct discrimination of 90.5 and 86.6 %, respectively. Despite that the main purpose of this work is diagnosis, a great portion of this effort was to identify the vertebra.

Another work was presented by de Bruijne and Nielsen [11] who used Shape Particle Filter [23] and k-nearest neighbor pixel-level classification for a semi-automatic segmentation of the lumbar vertebrae. Around the same time, Kaminsky et al. [47] presented a standardized protocol that combined newly developed interactive tools (rotation transformation, warped dissection plane) with standard segmentation tools to provide both a fast and accurate 3D spine segmentation procedure. Howe et al. [42] proposed a multi-level segmentation technique for vertebrae from cervical and lumbar X-ray radiographs using an ASM and a generalized Hough transform. Their validation is based on a leave-one-out test with an error of 2 mm for 57 % of the cervical cases and less than 4 mm for 68 % of the lumbar cases. However, their method needs manual intervention for initialization.

Later, Crimi et al. [22] presented a Bayesian approach and used prior information to estimate the covariance matrix from a small number of samples in a high dimensional shape to segment the vertebrae from X-ray radiographs. Moreover, Zewail et al. [103] segmented the vertebrae from X-ray radiographs using a contourlet-based salient point matching and a localized multi-scale shape prior. They tested their work on 100 X-ray radiographs and obtained an average segmentation error of 1.2 mm. Later, Crimi et al. [22] presented a Bayesian approach and used prior information to estimate the covariance matrix from a small number of samples in a high dimensional shape to segment the vertebrae from X-ray radiographs. Moreover, Zewail et al. [103] segmented the vertebrae from X-ray radiographs using a Contourlet-based salient point matching and a localized multi-scale shape prior. They tested their work on 100 X-ray radiographs and obtained an average segmentation error of 1.2 mm.

More recently, Lecron et al. [59] presented a fully automated vertebrae detection. They used an edge polygonal approximation to detect vertebral edges and a SIFT descriptor to train an SVM-model. They achieved a corner detection rate of 90.4 % and a vertebra detection rate from 81.6 to 86.5 % on 250 cervical cases.

5.1.2 Intervertebral Discs and Dural Sac

Despite that X-ray radiographs are mainly used for bone visualization, few efforts have been performed in the analysis of the intervertebral discs and the spinal canal. Chamarthy et al. [14] introduced image analysis techniques, including scale-invariant, distance transform-based features to characterize the disc space narrowing (with four grades zero to three) in cervical vertebrae. For a data set of 294 vertebrae X-ray images, experimental results yielded average correct grade assignment of

greater than 82.10 %. However, the testing images had to be manually labeled with the boundary points.

Later on, Koompairojn et al. [56] described a fully automatic *Spinal Stenosis* diagnosis system via vertebral morphometry [73], using an Active Appearance Model (AAM) for segmentation and a Bayesian framework for classification. Experimental results on 86 lumbar spine X-ray images from the NHANES II database showed accuracy ranging from 75 to 80 %. Moreover, Stanley et al. [89] investigated new size-invariant features (claw and traction) for the detection of anterior Osteophytes for efficient Content Based Image Retrieval (CBIR). Using a K-means clustering and nearest neighbor classification approach, average correct classification rates of 85.80, 86.04 and 84.44 % were obtained for claw, traction and anterior Osteophytes, respectively, on 390 cervical vertebrae.

5.2 Computed Tomography (CT)

CT imaging technique has become indispensable for diagnosis of spine abnormalities by providing a detailed 3D representation of the anatomy. Compared to X-ray radiography and MR imaging, CT proved to have higher sensitivity and specificity in the visualization of the bone structures.

5.2.1 Vertebrae

A number of automatic and semi-automatic methods for segmentation of vertebrae and vertebral structures in CT have been proposed [27, 50, 52, 64, 80, 84, 91, 96, 98, 102] over the last decade. On one hand some researchers proposed techniques that segment each vertebra separately [50, 61, 64, 84, 98], which might lead to mis-segmentation due to absence of a clear boundary between vertebrae. To overcome this issue, some authors proposed techniques for simultaneous segmentation of all vertebrae [52, 80].

Early last decade, Hahn [36], proposed a fully automated approach to evaluate rotation of the cervical vertebrae in 3D using a multidimensional Powell minimization algorithm for spiral CT scans. Later, in 2004, the same research group [37] presented a method for determination of the planes separating the individual vertebrae of the spine from CT volumes using a Balloon based model. This model requires careful initialization similar to the 2D active contours (snakes) besides its high dependency on the edge detector.

Meanwhile, Ghebreab and Smeulders [27] presented a combination of Strings [26] and Necklaces [28] to model the spine in the lumbar area using both a priori knowledge about natural variation and anatomical saliency in the visual appearance of the spine. The Strings model focuses on learning the most relevant biological variation in the visual appearance of the spine as a whole, and Necklaces aims at exploiting inhomogeneities in multiple continuous shape and gray-level features of

vertebrae. Thus they were able to use both a priori knowledge about natural variation and anatomical saliency in the visual appearance of the spine. However, they tested their method on only six CT cases and with minimal spinal and vertebral deformations. Furthermore, manual intervention is used for initialization of their model. On the other hand, Vrtovec et al. [96] detected the spine curve from CT using a polynomial model to provide a Curved Planar Reformation (CPR) of the 3D spinal column. They fit the spinal curve to a set of points extracted from a distance map that emphasized the vertebral bodies and tested the method on five cases, including one Scoliotic case achieving mean positional errors between 2 and 6 mm.

Furthermore, Yao et al. [102] presented a systematic algorithm for segmenting the spinal column from chest and/or abdominal CT scans, without labeling of vertebrae. Their method is based on thresholding, Watershed, and directed graph search besides modeling the vertebral bony tissue as a four-part model. They showed correct segmentation of 69 cases out of the 71 total cases. Meanwhile, Mastmeyer et al. [64] segmented the lumbar vertebral bodies in CT images by combining viscous deformable models with the geometrical shape of the vertebral body, starting from a point in the center of each vertebral body. Tan et al. [91] presented a level set-based segmentation algorithm for the vertebrae and validated their work on synthetic 3D vertebrae volumes. After parameter selection, they tested the algorithm on 50 vertebrae (from ten subjects), obtaining 90 % success rate. Later, Shen et al. [84] presented a segmentation technique for vertebrae from 3D CT scans using prior knowledge with a set of high level features to form a surface model. However, they did not perform labeling. They tested their model on 150 vertebrae with a comparative segmentation to two experts' segmentation. In the same year, Klinder et al. [51] presented a two-scale framework for modeling and segmenting the spine from thoracic CT scans, achieving a segmentation accuracy of 1.0 mm in average for ten thoracic CT volumes. By applying statistical models of shape, gradient and appearance of spinal structures in 3D, the same research group [52] detected, identified and segmented the vertebrae in CT volumes. However, their identification algorithm is based on vertebrae Active Appearance Model for spatial registration and matching which is very computationally expensive (20–30 min per case). Their framework was tested on 64 CT images including pathologies like Scoliosis, Kyphosis and collapsed vertebrae. Later, Kim and Kim [50] automatically segmented the vertebrae by a region growing algorithm inside a volume limited by a 3D fence that was obtained from a deformable model. They obtain 80 % success on a 50 patient dataset. More recently, Ma and Lu [61] proposed a method for segmentation and identification of thoracic vertebrae in CT images by training an edge detector to bone structures via steerable gradient features and using a deformable surface model in a two-stage coarse-to-fine scheme. They achieve point-to-surface error 0.95 ± 0.91 mm on 40 volumes.

Segmentation that is not based on deformable models [50, 91, 102], generally do not provide any quantitative information of vertebral deformations for CAD systems, while the segmentation based on deformable models is mathematically too abstract for describing deformations in clinical practice [52, 61, 64]. For example, Štern et al. [98] proposed a parametric method for quantitative description of

vertebral body deformations; evaluated from the parameters of a 3D super-quadratic model, which is initialized as an elliptical cylinder and then gradually deformed by introducing transformations that yield a more detailed representation of the vertebral body shape. Their method was validated on 75 CT and 75 MRI vertebrae extracted from none normal and ten abnormal subjects; showing a success rate of 94.5 and 88.6 %, respectively.

Meanwhile, Kadoury et al. [46] proposed a method for inferring articulated spine models from pre-operative X-ray to intra-operative CT images. This approach automatically segments the entire spinal column with annotated landmarks by modeling complex, non-linear patterns of prior deformations from a Riemannian manifold embedding, showing an accuracy of 0.7 ± 1.8 mm for thoracic vertebra and 2.1 ± 2.5 mm for lumbar vertebra based on the localization of surgical landmarks. Recently, Rasoulian et al. [80] developed a statistical multi-vertebrae shape and pose model and proposed a registration-based technique to segment the CT images of spine using a reduced number of registration parameters. Validation on lumbar vertebrae of 32 subjects shows a mean error less than 2 mm, which the authors argue, is sufficient for many spinal needle injection procedures, such as facet joint injections.

In another recent approach that avoids an explicit parametric model of appearance, Glocker et al. [35] proposed a vertebrae localization and identification algorithm which builds upon supervised classification forests. They overcome the tedious requirement for dense annotations by a semi-automatic labeling strategy. Extensive evaluation on a dataset of 224 spine CT scans of patients with pathologies (including high-grade Scoliosis, Kyphosis, and presence of surgical implants) shows a mean localization error of 12.4 mm and 70 % identification rates on pathological spines, which outperforms a parametric approach using Regression Forests and Hidden Markov Models (HMM).

One major effort in vertebrae segmentation was part of a semi-automated vertebra fracture detection system [2]. In the segmentation of vertebra, they started with the CT volume and select the middle slice as a starting point for segmentation. There are two main steps to train their model: (1) Inter vertebral disc localization (that leads to vertebra localization as illustrated below). (2) ASMs for each vertebra level.

For the first training task, they trained the proposed model in [5] by allowing a radiologist to place a point inside each disc for the six discs enclosing the five lumbar vertebrae. Then they saved this data with the corresponding images to train the model for the disc localization step (a point inside each disc).

The second training task is the selection of a fixed set of points (16 points) for each vertebra. Then produced a separate model for each vertebra level and prepare the training data required for an ASM (x-, y-coordinates and the image itself). Figure 14 shows a sample image with the 16 points on the edges of each vertebra as selected by the expert radiologist.

The steps for the segmentation of the lumbar vertebrae from CT are explained below in three sub steps: vertebrae localization, vertebra point distribution

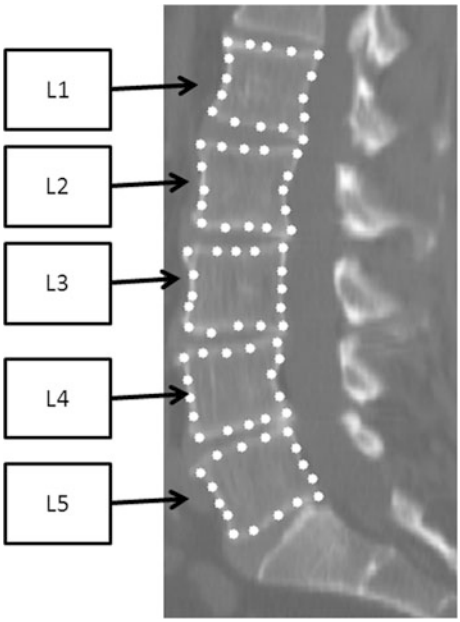


Fig. 14 Preparation of training data [2]. An expert radiologist manually selects a set of 16 points based on a predefined model for locating these points

modeling by the ASM, and vertebra boundary delineation by a Gradient Vector Flow (GVF-)snake [1, 2].

The vertebrae localization step provides a point inside each vertebra. This step utilized an earlier work on disc localization from clinical MRI [5]. After producing a point inside each disc, they take the average point between each two discs and consider this as the vertebra localization point as shown in Fig. 15.

The next step is to model the vertebra point distribution by an ASM [20]. In this work, they produced a separate model for each vertebra level. A radiologist prepares the training data where he manually marks 16 landmark points for each vertebra as shown in Fig. 14. These points are named from k_1 to k_{16}. Similar to [20], they initially calculated the mean shape $\bar{x} = \frac{1}{N}\sum_1^N x$ where N is the size of the training data. Then each vertebra shape x_i, where $i \in \{1, \ldots, N\}$, is recursively aligned to the mean shape $\bar{x}$ using generalized Procrustes analysis to remove translational, rotational, and isotropic scaling from the shape.

Then, they model the remaining variance around the mean shape for each vertebra with principal components analysis (PCA) to extract the Eigen vectors of the covariance matrix associated with 98 % of the remaining point position variance according to the standard method for deriving the ASM's linear shape representation.

However, they did not use the original CT image for training the ASM of each vertebra. Rather, they applied the range filter R first on the image to obtain a better edge enhancement for vertebrae. R is the range filter operator where the intensity levels in each 3 × 3 window are replaced by the range value (maximum–minimum)

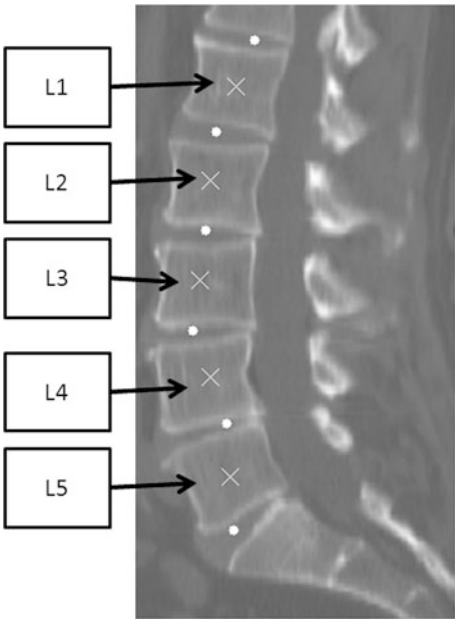


Fig. 15 Automated vertebrae localization. *Filled circles* are disc labels from Alomari et al. [5]. *Crosses* are the average location between each two disc labels

in that window. This operator R has high values in abrupt-change regions and small values in smooth regions as shown in Fig. 16.

To apply the ASM for detection of the point distribution of the vertebra body boundary, they applied the mean shape $\bar{x}$ around the vertebra point produced by the localization step (cross inside each vertebra). Then, allowed the ASM to converge and obtain the boundary. They then fed this boundary to the GVF-snake in the next step.

The ASM can capture the rough boundary of the vertebra as a point distribution model. However, fine detailed delineation of the vertebra body need a more refining model. They [1] selected the GVF-snake proposed by Xu and Prince [101] because it has been proven to move toward desired image properties such as edges including concavities. GVF-snake is the parametric curve that solves:

$$\mathbf{x}_t(s,t) = \alpha \mathbf{x}''(s,t) - \beta \mathbf{x}''''(s,t) + \mathbf{v} \tag{1}$$

where α and β are weighting parameters that control the contour's tension and rigidity, respectively. x'' and x'''' are the second and fourth derivatives, respectively, of x. $\mathbf{v}(x,y)$ is the Gradient Vector Flow (GVF), $s \in [0,1]$, and t is time component to make a dynamic snake curve from $x(s)$ yielding $x(s, t)$.

GVF-snake requires an edge map that is a binary image highlighting the desired features (edges) of the image. Most researchers use Canny edge detector or Sobel operator on the original image. They presented the GVF-snake with a canny edge map applied on the range-filtered image I.

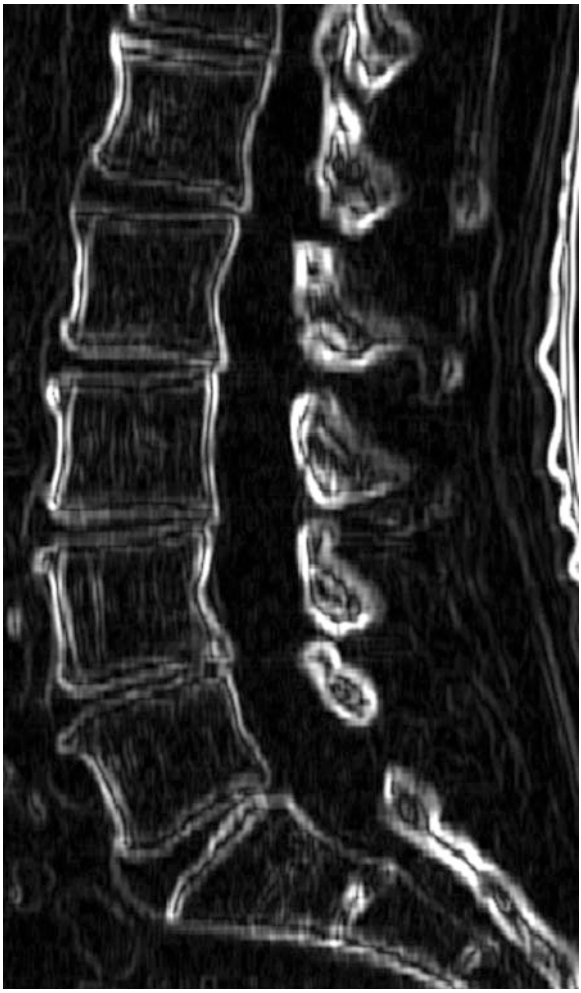


Fig. 16 Range filter 3 × 3 window on the CT image [1]

They then applied the GVF-snake by initializing its contour to the contour produced by the ASM, that is the points k_1 to k_{16}. Figure 15 shows the same example after the convergence of the GVF-snake.

Figures 17 and 18 show four cases selected from the data set to show the robustness of the final contour despite the various abnormalities in various lumbar levels. They performed qualitative measure where a radiologist visually and carefully examined each vertebra contour and approved the automated segmentation contour for all cases.

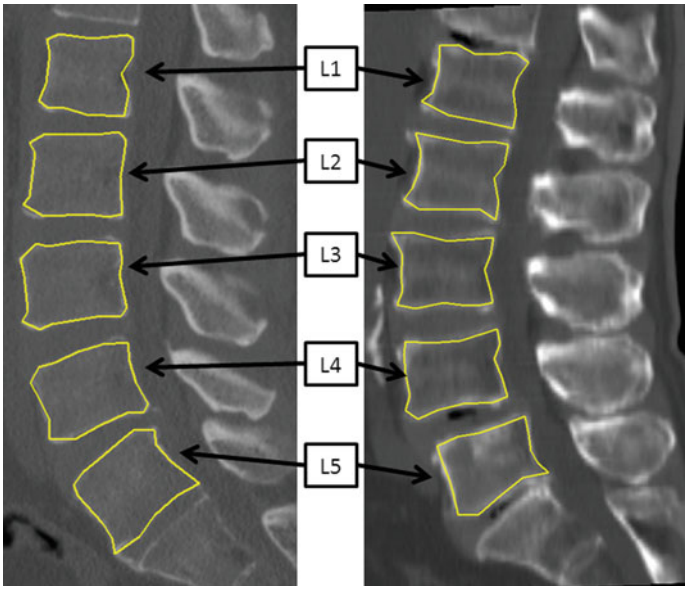


Fig. 17 Final contour for two cases. Images are contrast-enhanced for visual convenience [1]

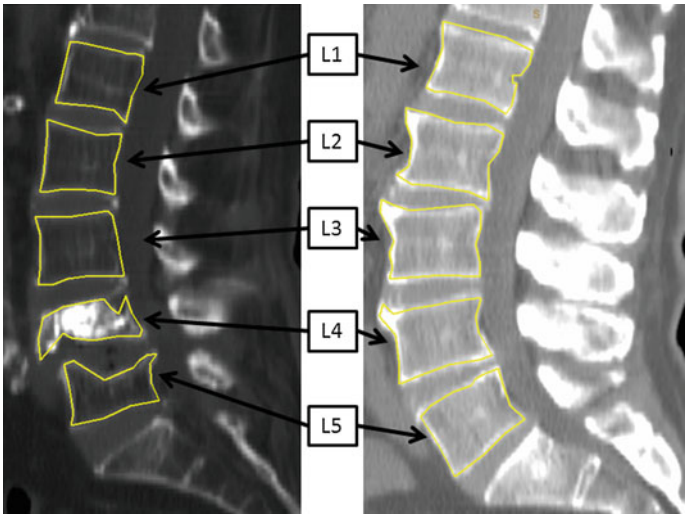


Fig. 18 Final contour for two cases. *Left* severely abnormal L4 vertebra. Images are contrast-enhanced for visual convenience [1]

5.2.2 Spinal Cord and Canal

Many research groups have focused on the segmentation of the spinal cord and the spinal canal in CT. Early on, Karangelis and Zimeras [48] introduced a semi-automatic 3D method for segmenting the spinal cord and tested that on 14 CT volumes. On each slice image, they used a boundary tracking method along with linear interpolation in the z-direction. However, proper selection of the seed point and the threshold limits its applicability. Meanwhile, Archip et al. [7] presented a top-down knowledge-based technique that identified the spinal cord in CT images. This approach used an Anatomical Structures Map and a task-oriented architecture plan solver. They claimed that the method was flexible enough to handle inter-patient variation and transparent to the radiologist ensuring that the experts can take control of undesirable results by image analysis. On 23 cases, the spinal canal was localized with an accuracy of 92 %, the spinal cord with an accuracy of 85 % and the lamina with an accuracy of 72 %. Couple years later, Burnett et al. [12] developed a semi-automatic algorithm for spinal canal segmentation of CT scans. The spinal canal was partially delineated by wavelet-based edge detection and fitted to a deformable model. Later, the template was aligned manually to fit more accurately to the spinal canal. Experiments on 557 axial images showed that automatic delineation of the spinal canal was successful on 91 %, unsuccessful on 2 % and requiring further editing on the rest 7 % of the images. Around same time, Nyúl et al. [75] proposed a semi-automatic method using 2D snakes for segmenting the spinal cord in a slice-by-slice manner testing that on 27 CT images for the Thoracic region. The 3D volume is then generated by interpolation. Snakes [49] are highly sensitive for the initialization which is usually performed manually.

On the other hand, because CT scans are better than X-rays and MRIs in terms of boney structure visualization, there has been great efforts toward building a CAD system for detection of various abnormalities such as Syndesmophytes (abnormal bone structures at the vertebral end plates) [90], spine Metastases [38] and vertebral fractures [2, 29]. Most of these efforts include localization, labeling, or segmentation work.

Mid last decade, Tan et al. [90] provided a quantitative measure of the Syndesmophytes using high resolution CT images. They first segmented the whole vertebra using a cascade of successive level sets, and then used curvature information to segment and quantify Syndesmophytes achieving 0.898 Pearson correlation between manual (medical expert) and the automated diagnosis which a high positive correlation level.

More recently, Hammon et al. [38] proposed a method of automatic detection of Lytic and Blastic Thoracolumbar spine Metastases (malignant tumors) from 3D CT images. They first detected the vertebral bodies using iterative marginal space learning and then use a cascade detector consisting of three random forest-based discriminative models to detect Metastases. Evaluation on 20 patients with 42 Lytic and on 30 patients with 172 Blastic Metastases (where the CAD system was trained using CT images of 114 subjects with 102 Lytic and 308 Blastic spinal Metastases) showed a sensitivity of 88 % for Lytic and 83 % for Blastic Metastases.

In vertebral fracture detection, Ghosh et al. [29] developed an unsupervised and non-parametric approach for vertebral segmentation using Hough lines and morphological operations. They also proposed a set of clinically motivated features including vertebral height features for automatic fracture detection using a Support Vector Machine (SVM). On 50 clinical cases they showed a segmentation error of 1.5 mm and a wedge fracture detection accuracy of 97 %. More recently, Al-helo et al. [2] proposed another method using ASM and a GVF-snake for vertebra segmentation and clinically motivated features for wedge fracture detection resulting in 98 % accuracy (specificity of 87.5 % and sensitivity over 99 %) using an unsupervised learner.

5.3 Magnetic Resonance Imaging (MRI)

While CT proved to have higher sensitivity and specificity in the visualization of the bone structures, MRI provides superior contrast in visualizing the soft tissue that surrounds the vertebrae, without ionizing radiation associated with CT or X-ray imaging. Moreover, MRI does not subject the patient for harmful radiations of the X-ray radiography and CT. It is important to highlight that research efforts in the literature have not been focused on distinct problems. There are many overlaps in research papers that may target localization, labeling, segmentation, and even diagnosis. We provide approximate categorization below for the literature based on the target problem.

5.3.1 Localization and Labeling

As early as 1989, Chwialkowski et al. [19] studied the localization of discs, vertebrae and spinal cord in one MRI case using intensity profiles and edge detectors. A decade later, Booth et al. [10] used an algorithm based on symmetry, active contours and edge detection to identify the vertebral body edges from cross-sectional vertebral MRI. However, the unavailability of data prevented these efforts from robust validation. Later in the last decade, Vrtovec et al. [97] detected the spine curve from MRI using a polynomial model to provide a Curved Planar Reformation (CPR) of the 3D spinal column. Their optimization framework is based on the automatic image analysis of MR spine images that exploits some basic anatomical properties of the spine. They tested the method on 21 axial MR scans of the spine from twelve subjects, achieving mean errors of 2.5 mm and 1.7° for the position of the 3D spine curve and axial rotation of vertebrae, respectively.

Mid last decade, Peng et al. [78] performed vertebra and disc labeling on five whole spine MRIs, by extracting intensity profiles of discs and use a convolution operation to match a template of the disc. Later, Masaki et al. [63] proposed a method for automated geometry planning based on intensity and a Hough transform to localize the spine and the discs. They only used ten MRI normal cases for

validation. The dependency on static values (thresholds) limits the capability of segmentation methods when they are tested on different datasets. Furthermore, performing many sequential steps to achieve the segmentation task increases the error rate due to propagation of the error from each step to the next. Another study by Weiss et al. [100] proposed a semi-automatic technique for disc labeling. The upper and lower halves of the spine are separately labeled after histogram processing, filters and the use of threshold values. They tested their algorithm on fifty MRI cases.

In surgery planning, Pekar et al. [77] developed a labeling method for the whole spine. Initially, a set of disc candidates are located by a filter using eigenvalues analysis of the Hessian matrix. Then using prior structural knowledge of the spine, they picked the disc centers from the candidates. After that labeling takes place starting from the first spine point and moving upward/downward. They also used a distance constraint for locating the next disc, otherwise a new point is introduced and that disc is considered missing due to abnormality. They used 15 subjects for validation producing 60 image volumes for lumbar and cervical areas with two poses for each subject.

Bhole et al. [9] presented a method for automatic detection and labeling of lumbar vertebrae and discs from clinical MRI by combining tissue property and geometric information from T1-Weighted (T1 W) sagittal, T2-Weighted (T2 W) sagittal and T2 W axial MRI protocols. They achieved 98.8 % accuracy for disc labeling on 67 sagittal images. However, they relied on specific threshold values extracted from the dataset which prevents the extension for their method to another dataset with variable parameter settings.

Schmidt et al. [82] introduced a probabilistic inference method using a part-based model achieving up to 97 % disc detection rate on 30 cases. In another similar approach, Oktay and Akgul [76] proposed a method using Pyramidal Histogram of Oriented Gradients (PHOG) based on SVM and a probabilistic graphical model and achieved 95 % accuracy on forty cases.

Localization and labeling has been better understood in the literature. The author's research group developed and tested a myriad of techniques. Koh et al. [54] proposed a joint attention and active contour models to segment the low back spine and subsequently label discs in later research efforts. However, the initial contour is highly sensitive to the inhomogeneous MRI signal intensity. Furthermore, [5] proposed a novel probabilistic model of the lumbar discs. This model adequately insulates the localization variables from the pixel intensities while at the same time modeling the exact disc geometry rather than solely pixel-level labels. Let $\mathcal{D} = \{d_0, d_1, \ldots, d_6\}$. be the set of disc variables with each $d_i = (x_i, y_i)^{\mathrm{T}}, i \in [1, 6]$ representing the disc center (it could also include disc angle, boundary, etc.), d_0 is a label for non-disc pixels. Inferring $\mathcal{D}$ from an image is our ultimate goal, but we avoid doing it directly due to its large computational complexity. We thus introduce a set of auxiliary variables, called disc-label variables and denoted by $\mathcal{L} = \{l_i, \forall i \in \Lambda\}$. Each disc-label variable can take a value of $\{-1, +1\}$ for non-disc or disc, respectively. The disc-labels make it plausible to separate the disc

variables from the image intensities, i.e., the disc-label variables will capture the local pixel-level intensity models while the disc variables will capture the high-level geometric and contextual models of the full set of discs. This approach is simpler and more robust than the model by Schmidt et al. [82] where they had a particular label for each disc. Next, we present more details about this highly cited work. This approach marginalizes over the possible disc-labelings since these are auxiliary variables giving the following optimization function:

$$\mathcal{D}^* = \arg\max_{\mathcal{D}} \sum_{\mathcal{L}} P(\mathcal{L}, \mathcal{D}|I) \tag{2}$$

$$= \arg\max_{\mathcal{D}} \sum_{\mathcal{L}} \frac{P(I|\mathcal{D}, \mathcal{L})P(\mathcal{D}, \mathcal{L})}{P(I)} \tag{3}$$

$$= \arg\max_{\mathcal{D}} \sum_{\mathcal{L}} P(I, \mathcal{L})P(\mathcal{L}|\mathcal{D})P(\mathcal{D}) \tag{4}$$

where the second equality follows from the multi-level nature of the model (the disc variables are assumed independent of the intensities). Note the summation is over a very large set of possible assignments $\left(2^{|\mathcal{A}|}\right)$. Then, the authors model it as a Gibbs distribution:

$$P(I, \mathcal{L}) = \frac{1}{Z}\exp[-\beta_1 \sum_{s\in\mathcal{A}} U_{\mathrm{I}}(l_s, I(s))] \tag{5}$$

$$P(\mathcal{L}|\mathcal{D}) = \frac{1}{Z}\exp[-\beta_2 \sum_{s\in\mathcal{A}} U_{\mathrm{D}}(l_s, \mathcal{D})] \tag{6}$$

$$P(\mathcal{D}) = \frac{1}{Z}\exp[-\beta_3 \sum_{d_i\in\mathcal{D}} U_{\mathrm{L}}(d_i) - \beta_4 \sum_{(i\sim j)} V_{\mathrm{D}}(d_i, d_j)] \tag{7}$$

where $\beta_k \geq 0, k = \{1, \ldots, 4\}$ are tunable parameters and $Z[\cdot]$ are the partition functions. The $(\cdot \sim \cdot)$ notation denotes the set of neighboring elements on the disc chain. The potentials U_{I} and U_{D} model the pixel (low)-level intensity and spatial models, respectively. The potentials U_{L} and V_{D} model the object (high)-level location and context, respectively.

The exact inference is infeasible for this model because of the dependencies of $\mathcal{D}$ on all $\mathcal{L}$ despite that $\mathcal{D}$ is a Markov chain. They used the generalized Expectation Maximization (gEM) algorithm to optimize Eq. (4). Whereas an EM algorithm requires maximization in the M step, a generalized EM algorithm only requires an improvement over the current state. This particular method has a high disc localization rate. However, the spatial information assumes higher locality in low spine area within the MRI. Moreover, it does not incorporate other aspects such as angles within the disc chain and most importantly, it does not take into consideration the

meta-data of the patient such as weight, height, and history. Patient's low back structures vary based on their weight, height, and history.

While most of the literature in localization provides disc centroids [5, 9, 76, 82]. Ghosh et al. [32] presented an approach using heuristics and machine learning methods to provide tight bounding boxes for each disc achieving 99 % localization accuracy on 53 cases. This method can by-pass complicated segmentation algorithms and directly feed the detected disc region to a CAD system that extracts relevant features and automatically provides diagnostic results [30, 31].

5.3.2 Segmentation

Few research efforts have been conducted on segmentation of vertebrae from MRI despite that bones are better outlined in CT scans. In 2004, Carballido-Gamio et al. [13] discussed the segmentation of vertebral bodies from sagittal T1-Weighted (T1 W) MRI using normalized cuts [85] with Nyström approximation method [25]. T1 W MRI were first preprocessed by Anisotropic Diffusion algorithm [79] that smooths the image without distorting the edges. However, they test their work on only six subjects for lumbar area. Five years later, Huang et al. [43] proposed a statistical learning approach based on an improved AdaBoost algorithm for efficient vertebra detection from MRI with a success of 98 % on less than 25 cases.

As early as 1997, Roberts et al. [81] proposed a method based on watershed algorithm to segment the five lumbar level discs from MRI. However, they required major user intervention by carefully selecting an ROI. Their work studies the relation between patient age and disc height. They concluded that the disc height increases with aging and that it increases from L1-L2 level and decreases at L5-S level. Later on, Hoad and Martel [40] presented a technique to segment the bone and soft tissues from MRI. However, their method requires sensitive initialization by the user to locate four points on each vertebrae. Wachter et al. [99] used various image segmentation techniques including shape model, Hough transforms, and edge detectors to segment the 3D spine and discs in the cervical area from full 3D MRI. They did not report the number of validation cases except stating that they are several T1 W and T2 W cases. Couple years later, Chevrefils et al. [17] proposed a method to segment the discs based on Watershed and many image processing techniques including opening and erosion. This method, however, encounters an over-segmentation issue. To overcome this problem, the same group [18] also presented a framework for automatic segmentation of intervertebral discs of Scoliotic spines from 2D and 3D MRI. Twenty two texture features (18 statistical and four spectral) were extracted from every closed region obtained from their earlier segmentation procedure [17]; followed by PCA and clustering which resulted in an overall accuracy of 85 %, specificity of 83 % and sensitivity of 87 % on 505 images derived from only three patients.

A Hough transform based approach was presented by Shi et al. [86] which showed success on 48 out of 50 cases but no quantitative evaluation was discussed. Moreover, the first disc has to be hand labeled for initialization. Another approach

was proposed by Michopoulou et al. [68] based on three variations of atlas based segmentation. However, they start from a manually input point for the center of each disc. Evaluation on 42 normal and 78 degenerated discs showed best performance by the atlas-robust-fuzzy C-Means approach which combines prior anatomical knowledge with fuzzy clustering techniques.

More recently, Neubert et al. [71] presented a method for the 3D segmentation of Vertebral Bodies (VBs) and Intervertebral Discs (IVDs) from the thoracolumbar region using statistical shape analysis and registration of gray-level intensity profiles. Validation on a dataset of high resolution 3D MR SPACE scans from 28 asymptomatic volunteers resulted in Dice values of 0.89 and 0.88 (lumbar and thoracic IVDs, respectively). Furthermore, Law et al. [58] proposed an unsupervised disc segmentation method that employs an Anisotropic Oriented Flux detection scheme to distinguish the discs from the neighboring structures with similar intensity, recognize ambiguous disc boundaries, and handle the shape and intensity variation of the discs. However, they require two user provided points for initialization. Evaluation on mid-sagittal slices of 69 cases (110 normal vertebrae) showed an average of 0.92 Dice similarity coefficient.

Most of the methods presented to date for the segmentation of the spinal cord from MRI, has been semi-automatic [21, 41, 65, 74]. They include various approaches such as B-spline active surface optimization [21], watershed segmentation [74] and deformable models [65]. Horsfield et al. [41] proposed a semi-automatic method utilizing a constrained active surface model of the cord surface assess multiple Sclerosis.

In the past few years there has been few research efforts towards the fully automated spinal cord segmentation. Koh et al. [53] developed an approach using Gradient Vector Flow Field which achieved a similarity index of 0.7 on 52 cases. They estimated the spinal cord using the magnitude of the gradient vector flow edge map, followed by a connected component analysis to remove holes in the segmentation. The same research group [54] proposed an unsupervised and fully automatic method based on an active contour model based on saliency maps, achieving 0.71 Dice Similarity Index on 60 cases. Similarly, Mukherjee et al. [69] applied an active contour approach, which evolved an image gradient based, open-ended contour using dynamic programming-based energy-minimization. Evaluation on MRI scans of cat showed a mean positive correlation of 0.94. More recently, Chen et al. [15] proposed a deformable atlas-based registration combined with a topology preserving classification to robustly segment the spinal cord and the CeroSpinal Fluid (CSF).

In a knowledge-based approach to reconstruct the cervical tissues of the cervical spine, Seifert et al. [83] used the Hough transform and knowledge about spine curvature to find initial seed points for discs which are then refined by clustering by considering the center of gravity of the cluster as the disc center. Disc centers are then used to segment the soft tissues (spinal cord, trachea and discs) from nine cervical MRIs resulting in 91 % accuracy. However, due to the use of a number of rules and heuristics, it is not clear if this approach will work for pathological cases.

In most of the previous work, segmentation of the *Dural Sac*, vertebrae and intervertebral discs have been handled separately which might lead to overlapping tissue regions. Moreover, some techniques depend on shape models giving rise to errors in case of high variability in appearance. Recently Ghosh et al. [34], used a Gibbs sampling approach to simultaneously label all tissues in the lumbar MRI. This method uses both neighborhood intensity information and label information for each update. Experimental results on 53 cases showed an average Similarity Index of 0.77 and 0.66 for the *Dural Sac* and Intervertebral discs respectively. Within the same research group, Alomari et al. [6] presented a coordinated joint model to accurately segment the lumbar discs from clinical MRIs in addition to their diagnosis work.

On the other hand, due to better discrimination of soft tissues in MRI, there has been a growing interest in the research community for automatic diagnosis of InterVertebral Discs (IVD) abnormalities such as Herniation, Degeneration, Desiccation, as well as *Spinal Stenosis* and Spinal Scoliosis from 2D and 3D MRIs. Most of these efforts include steps for localization and segmentation of the target structure.

Early last decade, Tsai et al. [94] detected Herniation from 3D MRI and CT volumes of the discs by using geometric features such as shape, size and location. However, it is a computationally expensive method and served better for visualization.

Clinical MRIs are, however, mostly 2D due to the high cost and acquisition time involved. Michopoulou et al. [67] presented the classification of the Intervertebral Discs (IVDs) into normal or degenerated, by using fuzzy C-Means to perform semi-automatic atlas-based disc segmentation and then used a Bayesian classifier. They achieved 86–88 % accuracy on 34 cases. They also reported 94 % accuracy using texture features [66] for 50 manually segmented discs.

A reasonable amount of research involving the use of real clinical MRIs on large dataset from the same research group [3, 4, 31, 30, 55] and diagnostic reports has also been reported. Alomari et al. [4] presented a fully automated herniation detection system using GVF-snake for an initial disc contour and then trained a Bayesian classifier on the resulting shape features. They achieved 92.5 % accuracy on 65 clinical MRI cases but a low sensitivity of 86.4 %. Alomari et al. [3] also presented a desiccation diagnosis system in lumbar discs from clinical MRI using a probabilistic model and achieving over 96 % accuracy. Ghosh et al. [31, 30] presented a comprehensive comparison of features, dimensionality reduction techniques and classifiers for herniation detection resulting in high specificity and sensitivity. They were however evaluated on only 35 clinical cases. Koh et al. [55] developed a computer-aided diagnosis framework for lumbar spine with a two-level classification scheme using heterogeneous classifiers. They used clinical MR image data from 70 subjects in T1 and T2-weighted sagittal view for evaluation of the system achieving 99 % herniation detection accuracy along with a speedup factor of 30 times in comparison with radiologist's diagnosis.

Jäger et al. [45] presented a complete system for computer-aided assessment of anomalies in 3-D MRI images of Scoliotic spine which provided an orthogonal

view onto every vertebra. First the spinal cord is segmented using a manual seed point and an iterative process where the segmentation is updated by an energy based scheme derived from Markov random field (MRF) theory. Then the vertebrae are labeled using an intensity profile and finally using parametric approximation MPRs (Multi-planar reformatting) are computed that are orthogonal to the backbone for every position of the spinal cord. Evaluation on 20 clinical 3-D MRI SPACE datasets, results in a mean angle difference of less than six degrees.

Along with proposing a method for the 3D segmentation of vertebral bodies and IVDs, Neubert et al. [71] showed that the shape parameters describing the extracted 3D volumes of lumbar IVDs allowed successful identification (100 % sensitivity, 98.3 % specificity) of IVDs with early degenerative changes. They also noted that the 28 subjects used were asymptomatic, and that the shape features seemed to work well for early detection of degeneration. Recently, the same group [72] evaluated the performance of 3D shape parameters, intensity features, and planar measurements of lumbar IVDs to detect degeneration in 28 asymptomatic and 11 symptomatic patients, concluding that intensity features are the most relevant in symptomatic patients.

In another exploratory work, Ghosh et al. [33] showed the utility of axial lumbar MRI for automatic diagnosis of abnormal discs using Convolutional Neural Network for dynamic feature extraction and classification. They achieved 80.81 % accuracy (specificity of 85.29 % and sensitivity of 75.56 %) on 86 clinical cases (391 discs) using only an axial slice for each disc.

6 Summary

We provided a detailed description of the challenges and the current status towards a fully automated lumbar diagnostic system. Not only is there variability in scans due to varying modalities and parameter settings, there is also extreme inter-patient vulnerability due to patient structure, age, gender and abnormalities. In addition, medical scans suffer from problems like partial volume effects and intensity inhomogeneity which makes segmentation, labeling and diagnosis from medical imaging scans a very challenging problem. While CT uses harmful radiation, it is cheaper than MRI. However, MRIs are better in terms of soft tissue details and is a preferred modality to diagnose underlying causes of back pain. There has been significant efforts in the past few decades towards automatic labeling, segmentation and diagnosis via vertebral column CT and MRI scans. Approaches suggested in the current literature use various image processing, machine learning and computer vision techniques. However, in the direction of automatic diagnosis using real clinical MRI data, work has been rather limited due to the unavailability of data and the fact that clinical data are relatively more challenging.

Acknowledgments This work was partially supported by grants from NYSTAR and NSF.

References

1. Al-Helo S, Alomari R, Chaudhary V, Al-Zoubi M (2011) Segmentation of lumbar vertebrae from clinical CT using active shape models and GVF-snake. In: Annual international conference of the IEEE, Engineering in Medicine and Biology Society, EMBC 2011, pp 8033–8036. doi:10.1109/IEMBS.2011.6091981
2. Al-Helo S, Alomari R, Ghosh S, Chaudhary V, Dhillon G, Al-Zoubi M, Hiary H, Hamtini T (2013) Compression fracture diagnosis in lumbar: a clinical CAD system. Int J Comput Assist Radiol Surg 8(3):461–469
3. Alomari R, Corso J, Chaudhary V, Dhillon G (2010) Computer-aided diagnosis of lumbar disc pathology from clinical lower spine MRI. Int J Comput Aided Radiol Surg 5(3):287–293
4. Alomari R, Corso J, Chaudhary V, Dhillon G (2010) Toward a clinical lumbar CAD: herniation diagnosis. Int J Comput Aided Radiol Surg 6(1):119–126
5. Alomari R, Corso J, Chaudhary V (2011) Labeling of lumbar discs using both pixel- and object-level features with a two-level probabilistic model. IEEE Trans Med Imaging 30 (1):1–10
6. Alomari R, Chaudhary V, Corso J, Dhillon G (2013) Lumbar spine disc herniation diagnosis with a joint shape model. In: Proceedings of MICCAI computational spine imaging workshop, to appear
7. Archip N, Erard P, Egmont-Petersen M, Haefliger J, Germond J (2002) A knowledge-based approach to automatic detection of the spinal cord in CT images. Med Imaging, IEEE Trans 21(12):1504–1516
8. Benameur S, Mignotte M, Parent S, Labelle H, Skalli W, de Guise J (2003) 3D/2D registration and segmentation of scoliotic vertebrae using statistical models. Comput Med Imaging Graph 27(5):321–337
9. Bhole C, Kompalli S, Chaudhary V (2009) Context sensitive labeling of spinal structure in MR images. In: Proceedings of the SPIE medical imaging conference, vol 7260, pp 72603P–72603P–9
10. Booth S, Clausi DA (2001) Image segmentation using MRI vertebral cross-sections. In: Proceedings of Canadian conference on electrical and computer engineering, vol 2, pp 1303–1307
11. de Bruijne M, Nielsen M (2004) Image segmentation by shape particle filtering. In: Proceedings of 17th international conference on pattern recognition (ICPR), pp 722–725
12. Burnett S, Starkschall G, Stevens CW, Liao Z (2004) A deformable-model approach to semi-automatic segmentation of CT images demonstrated by application to the spinal canal. Med Phys 31(2):251–263
13. Carballido-Gamio J, Belongie S, Majumdar S (2004) Normalized cuts in 3D for spinal MRI segmentation. IEEE Trans Med Imaging 23(1):36–44
14. Chamarthy P, Stanley RJ, Cizek G, Long R, Antani S, Thoma G (2004) Image analysis techniques for characterizing disc space narrowing in cervical vertebrae interfaces. Comput Med Imaging Graph 28:39–50
15. Chen M, Carass A, Oh J, Nair G, Pham DL, Reich DS, Prince JL (2013) Automatic magnetic resonance spinal cord segmentation with topology constraints for variable fields of view. NeuroImage 83:1051–1062
16. Cherukuri M, Stanley RJ, Long R, Antani S, Thoma G (2004) Anterior osteophyte discrimination in lumbar vertebrae using size-invariant features. Comput Med Imaging Graph 28(12):99–108
17. Chevrefils C, Chriet F, Grimard G, Aubin C (2007) Watershed segmentation of intervertebral disk and spinal canal from MRI images. In: Lecture notes in computer science: image analysis and recognition, pp 1017–1027
18. Chevrefils C, Cheriet F, Aubin C, Grimard G (2009) Texture analysis for automatic segmentation of intervertebral disks of scoliotic spines from MR images. Inf Technol Biomed, IEEE Trans 13(4):608–620

19. Chwialkowski MP, Shile PE, Peshock RM, Pfeifer D, Parkey RW (1989) Automated detection and evaluation of lumbar discs in MR images. In: Proceedings of IEEE EMBS
20. Cootes TF, Taylor CJ (2001) Statistical models of appearance for medical image analysis and computer vision. In: Proceedings of SPIE medical imaging
21. Coulon O, Hickman SJ, Parker GJ, Barker G, Miller D, Arridge S (2002) Quantification of spinal cord atrophy from magnetic resonance images via a B-spline active surface model. Magn Reson Med 47(6):1176–1185
22. Crimi A, Ghosh A, Sporring J, Nielsen M (2009) Bayes estimation of shape model with application to vertebrae boundaries. In: Pluim J, Dawant BM (eds) Medical imaging 2009: image processing, vol 7259. SPIE, Bellingham, p 72590A
23. Doucet A, de Freitas N, Gordon N (2001) Sequential Monte Carlo methods in practice. Springer, Berlin
24. Fardon DF, Milette P (2001) Nomenclature and classification of lumbar disc pathology. Spine 26(5):E93–E113
25. Fowlkes C, Shan Q, Belongie S, Malik J (2002) Extracting global structure from gene expression profiles. In: Methods of microarray data analysis II. Springer, Berlin, pp 888–905
26. Ghebreab S, Smeulders A (2003) Strings: variational deformable models of multivariate ordered features. IEEE Trans Pattern Anal Mach Intell 25:1399–1410
27. Ghebreab S, Smeulders A (2004) Combining strings and necklaces for interactive three-dimensional segmentation of spinal images using an integral deformable spine model. Biomed Eng IEEE Trans 51(10):1821–1829
28. Ghebreab S, Pfluger PR, Smeulders AWM (2002) Necklaces: inhomogeneous and point-enhanced deformable models. Comput Vis Image Underst 86:96–117
29. Ghosh S, Alomari R, Chaudhary V, Dhillon G (2011a) Automatic lumbar vertebra segmentation from clinical CT for wedge compression fracture diagnosis. In: Proceedings of SPIE, vol 7963, pp 796303–796309
30. Ghosh S, Alomari R, Chaudhary V, Dhillon G (2011b) Composite features for automatic diagnosis of intervertebral disc herniation from lumbar MRI. In: Proceedings of the 33rd annual international conference of the IEEE Engineering in Medicine and Biology Society, EMBC, pp 5068–5071
31. Ghosh S, Alomari R, Chaudhary V, Dhillon G (2011c) Computer-aided diagnosis for lumbar MRI using heterogeneous classifiers. In: Proceedings of the 8th IEEE international symposium on biomedical imaging: from nano to macro, ISBI, pp 1179–1182
32. Ghosh S, Malgireddy MR, Chaudhary V, Dhillon G (2012) A new approach to automatic disc localization in clinical lumbar MRI: combining machine learning with heuristics. In: Proceedings of IEEE international symposium on biomedical imaging, ISBI, pp 114–117
33. Ghosh S, Chaudhary V, Dhillon G (2013a) Exploring the utility of axial lumbar MRI for automatic diagnosis of intervertebral disc abnormalities. In: Proceedings of SPIE medical imaging
34. Ghosh S, Malgireddy MR, Chaudhary V, Dhillon G (2013b) A supervised approach towards segmentation of clinical MRI for automatic lumbar diagnosis. In: Proceedings of MICCAI computational spine imaging workshop, to appear
35. Glocker B, Zikic D, Konukoglu E, Haynor DR, Criminisi A (2013) Vertebrae localization in pathological spine CT via dense classification from sparse annotations. In: Mori K, Sakuma I, Sato Y, Barillot C, Navab N (eds) Medical image computing and computer-assisted intervention MICCAI 2013, vol 8150., Lecture notes in computer scienceSpringer, Berlin, pp 262–270
36. Hahn M (2001) New approach to evaluate rotation of cervical vertebrae. In: Medical imaging 2001: image processing, SPIE, vol 4322, pp 1696–1704
37. Hahn M, Beth T (2004) Balloon based vertebra separation in CT images. In: IEEE symposium on computer-based medical systems 2004, p 310
38. Hammon M, Dankerl P, Tsymbal A, Wels M, Kelm M, May M, Suehling M, Uder M, Cavallaro A (2013) Automatic detection of lytic and blastic thoracolumbar spine metastases on computed tomography. Eur Radiol 23(7):1862–1870

39. Hedlund L, Gallagher J (1988) Vertebral morphometry in diagnosis of spinal fractures. Bone and Mineral 5(1):59–67
40. Hoad CL, Martel AL (2002) Segmentation of MR images for computer-assisted surgery of the lumbar spine. Phys Med Biol 47(19):3503–3517
41. Horsfield MA, Sala S, Neema M, Absinta M, Bakshi A, Sormani MP, Rocca MA, Bakshi R, Filippi M (2010) Rapid semi-automatic segmentation of the spinal cord from magnetic resonance images: application in multiple sclerosis. NeuroImage 50(2):446–455
42. Howe B, Gururajan A, Sari-Sarraf H, Long L (2004) Hierarchical segmentation of cervical and lumbar vertebrae using a customized generalized Hough transform and extensions to active appearance models. In: 6th IEEE southwest symposium on image analysis and interpretation 2004, pp 182–186
43. Huang S, Chu Y, Lai S, Novak CL (2009) Learning-based vertebra detection and iterative normalized-cut segmentation for spinal MRI. Med Imaging, IEEE Trans 28(10):1595–1605
44. Jackson A, Sheppard S, Johnson A, Annesley D, Laitt R, Kassner A (1999) Combined fat- and water-suppressed MR imaging of orbital tumors. AJNR Am J Neuroradiol 20 (10):1963–1969
45. Jäger F, Hornegger J, Schwab S, Janka R (2009) Computer-aided assessment of anomalies in the scoliotic spine in 3D MRI images. In: Proceedings of the 12th international conference on medical image computing and computer-assisted intervention: part II, MICCAI'09. Springer, Berlin, pp 819–826
46. Kadoury S, Labelle H, Paragios N (2011) Automatic inference of articulated spine models in CT images using high-order Markov random fields. Med Image Anal 15(4):426–437 (special section on IPMI 2009)
47. Kaminsky J, Klinge P, Rodt T, Bokemeyer M, Luedemann W, Samii M (2004) Specially adapted interactive tools for an improved 3D-segmentation of the spine. Comput Med Imaging Graph 28(3):119–127
48. Karangelis G, Zimeras S (2002) An accurate 3D segmentation method of the spinal canal applied to CT data. In: Meiler M, Saupe D, Kruggel F, Handels H, Lehmann T (eds) Bildverarbeitung fr die Medizin 2002. Springer Berlin, Informatik aktuell, pp 370–373
49. Kass M, Wittkin A, Terzopoulos D (1987) Snakes, active contour models. Int J Comput Vision 1:321–331
50. Kim Y, Kim D (2009) A fully automatic vertebra segmentation method using 3D deformable fences. Comput Med Imaging Graph 33(5):343–352
51. Klinder T, Wolz R, Lorenz C, Franz A, Ostermann J (2008) Spine segmentation using articulated shape models. In: Proceedings of the 11th international conference on medical image computing and computer-assisted intervention—part I, MICCAI'08. Springer, Berlin, pp 227–234
52. Klinder T, Ostermann J, Ehm M, Franz A, Kneser R, Lorenz C (2009) Automated model-based vertebra detection, identification, and segmentation in CT images. Med Image Anal 13 (3):471–482
53. Koh J, Kim T, Chaudhary V, Dhillon G (2010) Automatic segmentation of the spinal cord and the dural SAC in lumbar MR images using gradient vector flow field. In: Proceedings of the 32nd annual international conference of the IEEE Engineering in Medicine and Biology Society, EMBC, pp 2117–2120
54. Koh J, Scott PD, Chaudhary V, Dhillon G (2011) An automatic segmentation method of the spinal canal from clinical MR images based on an attention model and an active contour model. In: Proceedings on IEEE international symposium on biomedical imaging, ISBI, pp 1467–1471
55. Koh J, Chaudhary V, Dhillon G (2012) Disc herniation diagnosis in MRI using a CAD framework and a two-level classifier. Int J Comput Assist Radiol Surg 7(6):861–869
56. Koompairojn S, Hua KA, Bhadrakom C (2006) Automatic classification system for lumbar spine X-ray images. In: Proceedings of the 19th IEEE symposium on computer-based medical systems, CBMS'06. IEEE Computer Society, Washington, DC, pp 213–218

57. Krudy AG (1992) MR myelography using heavily T2-weighted fast spin-echo pulse sequences with fat presaturation. Am J Roentgenol 159(6):1315–1320
58. Law M, Tay K, Leung A, Garvin GJ, Li S (2013) Intervertebral disc segmentation in MR images using anisotropic oriented flux. Med Image Anal 17(1):43–61
59. Lecron F, Benjelloun M, Mahmoudi S (2012) Fully automatic vertebra detection in X-ray images based on multi-class SVM. In: Proceedings of SPIE, vol 8314, p 83142D
60. Long LR, Thoma GR (2000) Use of shape models to search digitized spine X-rays. In: Proceedings of 13th IEEE symposium on computer-based medical systems (CBMS), pp 255–260
61. Ma J, Lu L (2013) Hierarchical segmentation and identification of thoracic vertebra using learning-based edge detection and coarse-to-fine deformable model. Comput Vis Image Underst 117(9):1072–1083
62. Madan S, Deanery M (2003) Interobserver error in interpretation of the radiographs for degeneration of the lumbar spine. Iowa Orthop J 32:51–56
63. Masaki T, Lee Y, Tsai DY, Sekiya M, Kazama K (2006) Automatic determination of the imaging plane in lumbar MRI. In: Proceedings of SPIE of medical imaging, pp 1252–1259
64. Mastmeyer A, Engelke K, Fuchs C, Kalender WA (2006) A hierarchical 3D segmentation method and the definition of vertebral body coordinate systems for QCT of the lumbar spine. Med Image Anal 10(4):560–577 [special issue on Functional Imaging and Modelling of the Heart (FIMH 2005)]
65. McIntosh C, Hamarneh G (2006) Spinal crawlers: deformable organisms for spinal cord segmentation and analysis. In: Proceedings of the 9th international conference on medical image computing and computer-assisted intervention—volume part I, MICCAI'06. Springer, Berlin, pp 808–815
66. Michopoulou S, Boniatis I, Costaridou L, Cavouras D, Panagiotopoulos E, Panayiotakis G (2009) Computer assisted characterization of cervical intervertebral disc degeneration in MRI. J Instrum 4:287–293
67. Michopoulou S, Costaridou L, Panagiotopoulos E, Speller R, Panayiotakis G, Todd-Pokropek A (2009) Atlas-based segmentation of degenerated lumbar intervertebral discs from MR images of the spine. IEEE Trans Biomed Eng 56:2225–2231
68. Michopoulou SK, Costaridou L, Panagiotopoulos E, Speller R, Panayiotakis G, Todd-pokropek A (2009) Atlas-based segmentation of degenerated lumbar intervertebral discs from MR images of the spine. Biomed Eng, IEEE Trans 56:2225–2231
69. Mukherjee DP, Cheng I, Ray N, Mushahwar V, Lebel M, Basu A (2010) Automatic segmentation of spinal cord MRI using symmetric boundary tracing. Inf Technol Biomed, IEEE Trans 14(5):1275–1278
70. Mulconrey D, Knight R, Bramble J, Paknikar S, Harty P (2006) Interobserver reliability in the interpretation of diagnostic lumbar MRI and nuclear imaging. Spine J 6:177–184
71. Neubert A, Fripp J, Engstrom C, Schwarz R, Lauer L, Salvado O, Crozier S (2012) Automated detection, 3D segmentation and analysis of high resolution spine MR images using statistical shape models. Phys Med Biol 9:8357–8376
72. Neubert A, Fripp J, Engstrom C, Walker D, Schwarz R, Crozier S (2013) Automatic quantification of 3D morphology and appearance of intervertebral discs in high resolution MRI. In: Annual meeting and exhibition on International Society for Magnetic Resonance in Medicine (ISMRM). International Society for Magnetic Resonance in Medicine (ISMRM), Salt Lake City, p 1612
73. Nicholson P, Haddaway MJ, Davie M, Evans SF (1993) A computerized technique for vertebral morphometry. Physiol Meas 14(2):195
74. Nieniewski M, Serneels R (2002) Segmentation of spinal cord images by means of watershed and region merging together with inhomogeneity correction. MG&V 11(1):101–121
75. Nyúl LG, Kanyó J, Máté E, Makay G, Balogh E, Fidrich M, Kuba A (2005) Method for automatically segmenting the spinal cord and canal from 3D CT images. In: CAIP, pp 456–463

76. Oktay AB, Akgul YS (2011) Localization of the lumbar discs using machine learning and exact probabilistic inference. In: Medical image computing and computer-assisted intervention (MICCAI), vol 3
77. Pekar V, Bystrov D, Heese HS, Dries S, Schmidt S, Grewer R, den Harder CJ, Bergmans RC, Simonetti AW, van Muiswinkel AM (2007) Automated planning of scan geometries in spine MRI scans. In: Medical image computing and computer-assisted intervention—MICCAI 2007, vol 4791. Springer, Berlin, pp 601–608
78. Peng Z, Zhong J, Wee W, Lee J (2005) Automated vertebra detection and segmentation from the whole spine MR images. In: 27th annual international conference of the Engineering in Medicine and Biology Society, IEEE-EMBS 2005, pp 2527–2530
79. Perona P, Shiota T, Malik J (1994) Geometry-driven diffusion in computer vision. Springer, Berlin
80. Rasoulian A, Rohling R, Abolmaesumi P (2013) Lumbar spine segmentation using a statistical multi-vertebrae anatomical shape+ pose model. IEEE Trans Med Imaging 32 (10):1890–1900
81. Roberts AN, Gratin C, Whitehouse GH (1997) MRI analysis of lumbar intervertebral disc height in young and older populations. J Magn Reson Imaging 7(5):880–886
82. Schmidt S, Kappes H, Bergtholdt M, Pekar V, Dries S, Bystrov D, Schnörr C (2007) Spine detection and labeling using a parts-based graphical model. In: IPMI'07, vol 4584., Lecture notes in computer science. Springer, Berlin, pp 122–133
83. Seifert S, Wachter I, Schmelzle G, Dillmann R (2009) A knowledge-based approach to soft tissue reconstruction of the cervical spine. IEEE Trans Med Imaging 28(4):494–507
84. Shen H, Litvin A, Alvino C (2008) Localized priors for the precise segmentation of individual vertebras from CT volume data. In: The proceedings of medical imaging computing and computer assisted intervention (MICCAI'08), vol 5241., LNCS, pp 367–375
85. Shi J, Malik J (2000) Normalized cuts and image segmentation. IEEE Trans Pattern Anal Mach Intell 22:888–905
86. Shi R, Sun D, Qiu Z, Weiss KL (2007) An efficient method for segmentation of MRI spine images. In: IEEE/ICME international conference on complex medical engineering (CME 2007), pp 713–717
87. Smyth PP, Taylor CJ, Adams JE (1997) Automatic measurement of vertebral shape using active shape models. In: Duncan J, Gindi G (eds) Information processing in medical imaging, vol 1230., Lecture notes in computer scienceSpringer, Berlin, pp 441–446
88. Snell RS (2007) Clinical anatomy by regions, 8th edn. Lippincott Williams and Wilkins, Philadelphia
89. Stanley RJ, Antani S, Long R, Thoma G, Gupta K, Das M (2008) Size-invariant descriptors for detecting regions of abnormal growth in cervical vertebrae. Comput Med Imaging Graph 32(1):44–52
90. Tan S, Yao J, Ward M, Yao L, Summers R (2006) Computer aided evaluation of ankylosing spondylitis. In: 3rd IEEE international symposium on biomedical imaging: nano to macro, 2006, pp 339–342
91. Tan S, Yao J, Ward MM, Yao L, Summers RM (2007) 3D multi-scale level set segmentation of vertebrae. In: 4th IEEE international symposium on biomedical imaging: from nano to macro, ISBI 2007, pp 896–899
92. Tartaro A, Onofrj M, Delli C, Bonomo L, Thomas A, Fulgente T, Gambi D (1996) Long time echo stir sequence magnetic resonance imaging of optic nerves in optic neuritis. Ital J Neurol Sci 17(1):35–42
93. Tien RD (1992) Fat-suppression MR imaging in neuroradiology: techniques and clinical application. Am J Roentgenol 158(2):369–379
94. Tsai M, Jou S, Hsieh M (2002) A new method for lumbar herniated inter-vertebral disc diagnosis based on image analysis of transverse sections. Comput Med Imaging Graph 26 (6):369–380
95. University S (2009) A patient's guide to lumbar spine anatomy. Website:www.spineuniversity.com

96. Vrtovec T, Likar B, Pernus F (2005) Automated curved planar reformation of 3D spine images. Phys Med Biol 50(19):4527–4540
97. Vrtovec T, Ourselin S, Gomes L, Likar B, Pernus F (2007) Automated generation of curved planar reformations from MR images of the spine. Phys Med Biol 52(10):2865–2878
98. Štern D, Likar B, Pernuš F, Vrtovec T (2011) Parametric modelling and segmentation of vertebral bodies in 3D CT and MR spine images. Phys Med Biol 56(23):7505
99. Wachter I, Seifert S, Dillmann R (2005) Automatic segmentation of cervical soft tissue from MR images. In: J Troccaz, P Merloz (eds) Proceedings of Surgetica, Chambery
100. Weiss KL, Storrs JM, Banto RB (2006) Automated spine survey iterative scan technique. Radiology 239(1):255–262
101. Xu C, Prince JL (1998) Snakes, shapes, and gradient vector flow. Image Process, IEEE Trans 7(3):359–369
102. Yao J, O'Connor S, Summers R (2006) Automated spinal column extraction and partitioning. In: 3rd IEEE international symposium on biomedical imaging: nano to macro 2006, pp 390–393
103. Zewail R, Elsafi A, Durdle N (2009) Vertebral segmentation using contourlet-based salient point matching and localized multiscale shape prior. In: Pluim JPW, Dawant BM (eds) Medical imaging 2009: image processing, vol 7259. SPIE, Bellingham, p 72594Z

Automated Determination of the Spine-Based Coordinate System for an Efficient Cross-Sectional Visualization of 3D Spine Images

Tomaž Vrtovec

Abstract The most common and straightforward visualization of three-dimensional (3D) images of the spine and vertebrae, usually obtained by computed tomography (CT) or magnetic resonance (MR) imaging techniques, is based on multi-planar cross-sections that are positioned in the coordinate system of the 3D image. However, such multi-planar cross-sections may not provide sufficient or qualitative enough diagnostic information, because they cannot follow the curvature of the spine. As a result, not all of the important details can be shown simultaneously in any cross-section. To overcome this problem, cross-sections have to be generated in the coordinate system of the spine, which can be achieved by curved-planar 3D image reformation that reduces the structural complexity in favor of an improved feature perception of anatomical structures. The parameters for such cross-sectional image reformation are determined from the spine-based coordinate system, which is defined by the curve representing the vertebral column and by the rotation of vertebrae about the spine curve. This chapter is focused on the techniques for automated determination of the spine-based coordinate system for an efficient cross-sectional visualization of 3D spine images, acquired by the CT and MR imaging techniques. The reformatted cross-sections are diagnostically valuable, enable easier navigation, manipulation and orientation in 3D space, and are useful for initializing segmentation and other automated image analysis tasks.

1 Introduction

Medical images are of extreme importance for diagnosing and understanding of normal and pathological conditions of the human body. To some extent, the quality of image-assisted medical examinations depends on the acquisition of images, interpretation of the information present in images achieved through proper visualization,

T. Vrtovec (✉)
Faculty of Electrical Engineering, University of Ljubljana, Tržaška cesta 25, SI-1000 Ljubljana, Slovenia
e-mail: tomaz.vrtovec@fe.uni-lj.si

© Springer International Publishing Switzerland 2015

S. Li and J. Yao (eds.), *Spinal Imaging and Image Analysis*,
Lecture Notes in Computational Vision and Biomechanics 18,
DOI 10.1007/978-3-319-12508-4_8

and on the research activity and clinical environment that stimulate image formation and its application. Image formation can be defined as the process of mapping selected properties of the imaged object into the image space. The image space represents the basis for visualization of the object and its properties, and may be further used for quantitative evaluation of its structure or function, and interpretation of the information it contains. As quantitative evaluation and interpretation of images depend on the quality of the information of interest, the main purpose of image visualization is effective information extraction. In the field of medical image visualization, the extraction of clinically relevant information is therefore of significant importance for the development of accurate and non-invasive techniques for medical diagnosis and treatment.

Technological advances in medical imaging and computerized medical image processing led to the development of new three-dimensional (3D) image acquisition techniques that have become important clinical tools in modern diagnostic radiology and medical health care. Two-dimensional (2D) images, especially radiographs (X-ray images), are still widely present in clinical examination due to a relatively low acquisition price and wide area of application. However, the continuous increase in the number of acquired cross-sections, reduction in cross-sectional thickness and relatively short acquisition times led to the expansion of 3D imaging techniques [70]. Among the most important 3D techniques are computed tomography (CT) and magnetic resonance (MR) imaging, which provide qualitative data of the imaged structures. However, characteristic features of these techniques and variable positioning of the patient during image acquisition still represent a major source of variability that causes errors in the interpretation of image information. On the other hand, human capability of discovering and diagnosing diseases by proper interpretation of medical images is limited due to our non-systematic search patterns. Moreover, the presence of noise may conceal the natural anatomical background, such as actual geometrical relationships among anatomical structures, which may further hamper mental reconstruction of the 3D image information. Errors in interpretation may also be caused by similar characteristics of normal and pathological conditions, and by the natural biological variability of human anatomy. Image interpretation and quantitative evaluation therefore to a great extent depend on adequate visualization of the information about anatomical structures. As the information of interest is often associated with characteristic features of the selected structure or process, it is crucial to use specially designed image processing techniques for visualization and quantitative evaluation. As the number of acquired medical images is rapidly growing [7], computerized tools and devices may potentially help radiologists to reduce their workloads [73, 77] and direct clinicians towards an accurate interpretation and quantitative evaluation of the large amount of data within a reasonable time frame [75, 76]. Techniques for visualization and quantitative evaluation of medical images are therefore extremely valuable in the development of image-assisted diagnosis, planning of surgical interventions and assessment of medical treatment outcomes.

Both CT and MR are established image acquisition techniques for diagnosing and managing spinal and spine-related disorders [12, 81], as they provide

qualitative insight into the spinal anatomy, and have become indispensable investigative tools in related clinical decision making [44]. The CT imaging technique is appropriate for observing bones and other dense structures of the spine, especially relatively small and complex structures such as vertebral pedicles, facets and processes. On the other hand, the MR imaging technique allows examination of soft tissues, such as intervertebral discs, spinal cord and nerve roots. However, the identification, visualization and quantitative evaluation of many spinal disorders by routine examinations is difficult because the spine is a complex and articulated anatomical structure. The most common and straightforward visualization of 3D CT and MR images of the spine and vertebrae is based on multi-planar cross-sections that are defined in the coordinate system of the 3D image. However, such multi-planar cross-sections may not provide sufficient or qualitative enough diagnostic information, because they cannot follow the curvature of the spine. As a result, not all of the important details can be shown simultaneously in any cross-section. To overcome this problem, cross-sections have to be generated in the coordinate system of the spine, which can be achieved by curved-planar 3D image reformation. The parameters for such cross-sectional image reformation are determined from the spine-based coordinate system, which is defined by the curve representing the vertebral column and by the rotation of vertebrae around the spine curve.

This chapter is focused on the techniques for automated determination of the spine-based coordinate system for an efficient cross-sectional visualization of 3D spine images. The geometrical representation of the spine in 3D images (Sect. 2) is used to define the spine-based coordinate system, which provides means for quantitative evaluation of the spine as the observed anatomical structure, and represents the basis for cross-sectional reformation of 3D spine images (Sect. 3). Cross-sectional image reformation reduces the structural complexity in favor of an improved feature perception of the spinal anatomy, and results in valuable cross-sections that enable not only easier navigation, manipulation and orientation in 3D space, but can be also used for initializing spine image segmentation or other automated image analysis techniques. Further development of spine visualization and quantitative evaluation techniques may improve medical diagnosis and the design of more effective strategies for the treatment of spinal disorders. The fields of visualization and quantitative evaluation of spine images are closely related, as knowledge of spine parameters may provide a more effective spine visualization, and, on the other hand, proper spine visualization may allow a more effective measurement of spine parameters.

2 Geometrical Representation of the Spine in 3D Images

Geometrical representation of anatomical structures can be considered as a generalization of the anatomical information, as the most relevant properties of the observed anatomical structures are described by geometrical primitives such as points, lines or line segments, curves or curve segments, etc. Geometrical primitives

allow to quantify the location, size and shape of anatomical structures observed in medical images without making use of advanced image analysis techniques, such as segmentation. In the case of 3D spine images, the most relevant geometrical properties of the spine as the observed anatomical structure can be represented by the spine curve and axial rotation of vertebrae. To enable a proper geometrical representation of the spine in a 3D image, image-based and spine-based coordinate systems are introduced (Sect. 2.1), the spine-based coordinate system is formally defined (Sect. 2.2), and methods for automated determination of the spine-based coordinate system are presented (Sect. 2.3).

2.1 Coordinate Systems

Given a 3D image of an anatomical structure, the *image-based coordinate system* (Sect. 2.1.1) describes the acquisition of the image with respect to the observed anatomical structure. However, several 3D anatomical structures (e.g. spine, colon, arteries) have geometrical properties that enable the introduction of the *structure-based* or *anatomy-based* coordinate systems, which can be in the case of the spine as the observed anatomical structure termed *spine-based coordinate system* (Sect. 2.1.2).

2.1.1 Image-Based Coordinate System

The image-based coordinate system is the coordinate system of the acquired image. In the case of 3D images, it is defined by a 3-tuple $(x, y, z) \in \mathbb{R}^3_I$ of mutually orthogonal axes x, y and z, represented by unit vectors $\hat{\mathbf{e}}_{Ix} = [1, 0, 0]_I$, $\hat{\mathbf{e}}_{Iy} = [0, 1, 0]_I$ and $\hat{\mathbf{e}}_{Iz} = [0, 0, 1]_I$, respectively. In the case of 3D spine images, the image-based coordinate system usually corresponds to the coordinate system of the imaging device (i.e. CT or MR scanner), in which the body is longitudinally aligned. The image-based coordinate system is, in general, aligned with the coordinate system of the body, and therefore its axes represent the following anatomical directions:

- the *sinistro-dexter* axis x represents the direction $\hat{\mathbf{e}}_{Ix}$ from the left to the right part of the body,
- the *ventro-dorsal* axis y represents the direction $\hat{\mathbf{e}}_{Iy}$ from the anterior to the posterior part of the body,
- the *cranio-caudal* axis z represents the direction $\hat{\mathbf{e}}_{Iz}$ from the superior to the inferior part of the body (i.e. the longitudinal axis of the body).

As a result, the following three imaging planes can be defined:

- a *sagittal* or *lateral* plane (y, z) is any plane that passes from the anterior to the posterior, and from the superior to the inferior part of the body, therefore dividing the body into left and right sections (the sagittal plane that passes

through the midline of the body and, assuming bilateral symmetry, divides the body into left and right halves is the mid-sagittal plane),

- a *coronal* or *frontal* plane (x, z) is any plane that passes from the left to the right, and from the superior to the inferior part of the body, therefore dividing the body into anterior and posterior sections (the coronal plane that passes through the midline of the body and, assuming bilateral symmetry, divides the body into anterior and posterior halves is the mid-coronal plane),
- an *axial* or *transverse* plane (x, y) is any plane that passes from the left to the right, and from the anterior to the posterior part of the body, therefore dividing the body into superior and inferior sections.

The image-based coordinate system is therefore a right-handed Cartesian coordinate system with a standard orthonormal basis (Fig. 1).

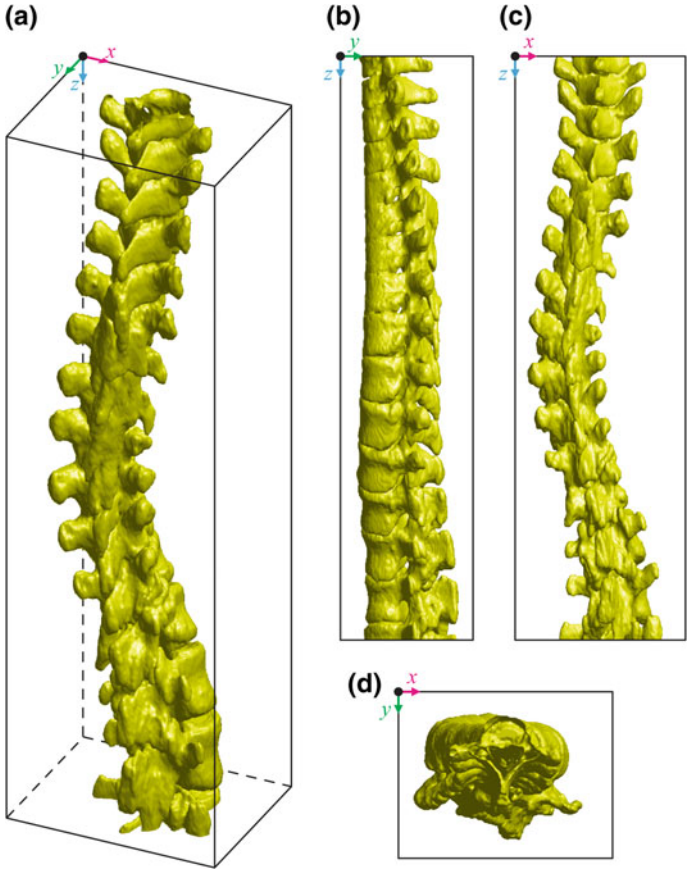


Fig. 1 The image-based coordinate system $\mathbb{R}^3_I \rightarrow (x, y, z)$ of a 3D image of a scoliotic spine, shown in **a** 3D view, **b** left sagittal view, **c** posterior coronal view and **d** superior axial view

2.1.2 Spine-Based Coordinate System

The spine-based coordinate system [84] is the coordinate system of the spine as the observed anatomical structure. In the case of 3D spine images, it is defined by a 3-tuple $(u, v, w) \in \mathbb{R}^3_S$ of mutually orthogonal axes u, v and w, represented by unit vectors $\hat{\mathbf{e}}_{Su} = [1, 0, 0]_S$, $\hat{\mathbf{e}}_{Sv} = [0, 1, 0]_S$ and $\hat{\mathbf{e}}_{Sw} = [0, 0, 1]_S$, respectively. The spine-based coordinate system is aligned with the spine, and its axes represent the following structural directions:

- the anatomical axis u represents the direction $\hat{\mathbf{e}}_{Su}$ from the left to the right part of the spine,
- the anatomical axis v represents the direction $\hat{\mathbf{e}}_{Sv}$ from the anterior to the posterior part of the spine,
- the anatomical axis w represents the direction $\hat{\mathbf{e}}_{Sw}$ from the superior to the inferior part of the spine (i.e. the longitudinal axis of the spine).

As a result, the following three anatomical (structural) planes can be defined:

- a *sagittal* or *lateral* plane (v, w) is any plane that passes from the anterior to the posterior, and from the superior to the inferior part of the spine, therefore dividing the spine into left and right sections,
- a *coronal* or *frontal* plane (u, w) is any plane that passes from the left to the right, and from the superior to the inferior part of the spine, therefore dividing the spine into anterior and posterior sections,
- an *axial* or *transverse* plane (u, v) is any plane that passes from the left to the right, and from the anterior to the posterior part of the spine, therefore dividing the spine into superior and inferior sections.

The spine-based coordinate system is therefore a right-handed Cartesian coordinate system with a standard orthonormal basis (Fig. 2).

2.2 Definition of the Spine-Based Coordinate System

To define the spine-based coordinate system, two geometrical properties of the spine have to be determined, namely the *spine curve* (Sect. 2.2.1) and *axial vertebral rotation* (Sect. 2.2.2), which serve to define the transformation from the image-based to the spine-based coordinate system (Sect. 2.2.3).

2.2.1 Spine Curve

The spine curve C is the curve that follows the curvature of the spine along its entire longitudinal length. If i is an independent parameter that denotes an arbitrary location on the spine, then $\mathbf{c}(i)$ is the parametrization of the spine curve C:

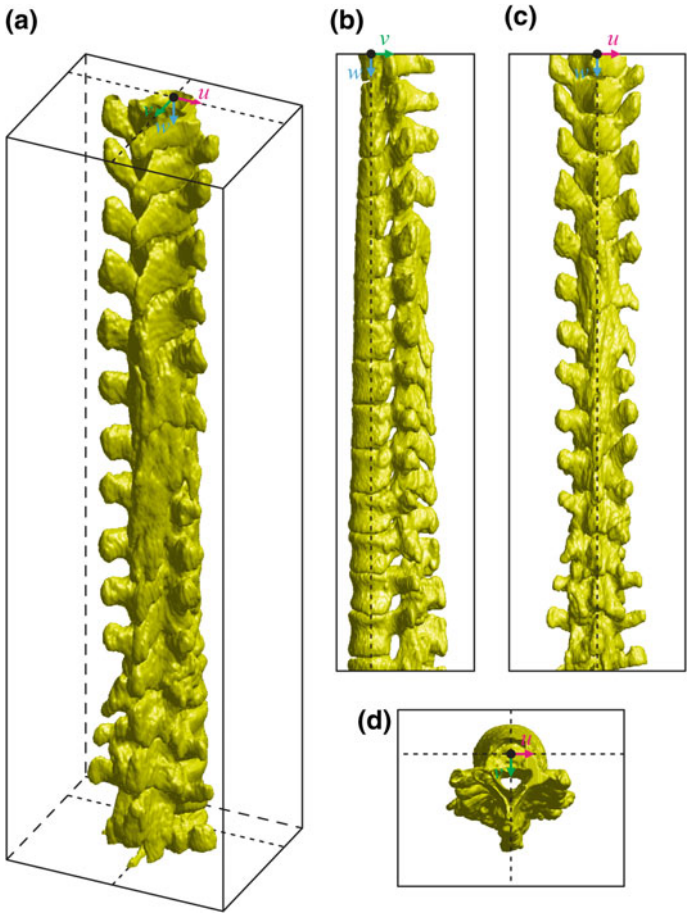


Fig. 2 The spine-based coordinate system $\mathbb{R}^3_S \rightarrow (u, v, w)$ of a 3D image of a scoliotic spine, shown in **a** 3D view, **b** left sagittal view, **c** posterior coronal view and **d** superior axial view (*Note* The spine corresponds to Fig. 1)

$$C: \quad \mathbf{c}(i) = (c_x(i), c_y(i), c_z(i)); \quad i \in [i_{\text{sp}}, i_{\text{ep}}], \tag{1}$$

where $i = i_{\text{sp}}$ and $i = i_{\text{ep}}$ represent the locations on the spine at its start and end point of observation, respectively, and $c_x(i)$, $c_y(i)$ and $c_z(i)$ represent the sagittal, coronal and axial coordinate, respectively, of the same anatomical reference point at any location i on the spine in the image-based coordinate system. Although arbitrary anatomical reference points can be chosen (e.g. the centers of the spinal canal), the most established anatomical reference points are the centers of vertebral bodies. For K observed consecutive vertebrae, let points $\{\mathbf{v}(k) = (v_x(k), v_y(k), v_z(k)); k = 1, 2, \ldots, K\}$ represent the corresponding centers of vertebral bodies. The spine

curve $\mathbf{c}(i)$ can be then obtained by continuous interpolation of $\mathbf{v}(k)$ between $\mathbf{c}(i_{sp}) = \mathbf{v}(1)$ and $\mathbf{c}(i_{ep}) = \mathbf{v}(K)$ (Fig. 3).

For a differentiable curve $\mathbf{c}(i)$, the geometric properties of the curve can be described in terms of differential geometry by the Frenet-Serret frame, which is defined as an orthonormal basis by the unit tangent, normal and binormal vectors to the curve. The unit tangent vector $\hat{\mathbf{t}}(i)$ represents the direction of the curve that corresponds to increasing values of parameter i:

$$\hat{\mathbf{t}}(i) = \frac{\mathbf{t}(i)}{\|\mathbf{t}(i)\|}; \quad \mathbf{t}(i) = \frac{\mathrm{d}\mathbf{c}(i)}{\mathrm{d}i}, \tag{2}$$

where $\mathbf{t}(i)$ is the tangent vector to the curve $\mathbf{c}(i)$, obtained as the first derivative of $\mathbf{c}(i)$ with respect to i, and $\|\cdot\|$ denotes the vector norm. The unit normal vector $\hat{\mathbf{n}}(i)$ represents the deviation of the curve from being a straight line:

$$\hat{\mathbf{n}}(i) = \frac{\mathbf{n}(i)}{\|\mathbf{n}(i)\|}; \quad \mathbf{n}(i) = \frac{\mathrm{d}\hat{\mathbf{t}}(i)}{\mathrm{d}i} = \frac{\mathrm{d}\mathbf{c}(i)}{\mathrm{d}i} \times \left(\frac{\mathrm{d}^2\mathbf{c}(i)}{\mathrm{d}i^2} \times \frac{\mathrm{d}\mathbf{c}(i)}{\mathrm{d}i}\right), \tag{3}$$

where $\mathbf{n}(i)$ is the normal vector to the curve $\mathbf{c}(i)$, obtained as the first derivative of $\hat{\mathbf{t}}(i)$ with respect to i, and $\times$ denotes the cross vector product. To satisfy the orthonormality of the basis, the unit binormal vector $\hat{\mathbf{b}}(i)$ is orthogonal to both the unit tangent vector and the unit normal vector:

$$\hat{\mathbf{b}}(i) = \hat{\mathbf{t}}(i) \times \hat{\mathbf{n}}(i) = \frac{\mathbf{b}(i)}{\|\mathbf{b}(i)\|}; \quad \mathbf{b}(i) = \frac{\mathrm{d}\mathbf{c}(i)}{\mathrm{d}i} \times \frac{\mathrm{d}^2\mathbf{c}(i)}{\mathrm{d}i^2}, \tag{4}$$

where $\mathbf{b}(i)$ is the binormal vector to the curve $\mathbf{c}(i)$, obtained as the cross vector product of the first and the second derivative of the curve $\mathbf{c}(i)$.

The unit tangent vector $\hat{\mathbf{t}}(i)$ and the unit normal vector $\hat{\mathbf{n}}(i)$ at location i on curve $\mathbf{c}(i)$ define the osculating plane at that location. The deviation of the curve from

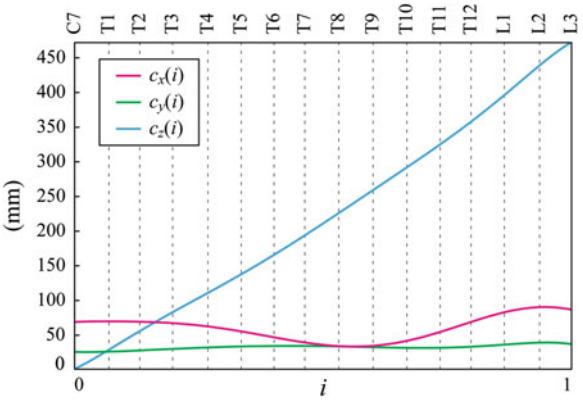


Fig. 3 The sagittal $c_x(i)$, coronal $c_y(i)$ and axial $c_z(i)$ component of the spine curve $\mathbf{c}(i)$ against the independent parameter $i \in [0, 1]$. Labels *C7*, *T1*, …, *L3* indicate vertebral segments (*Note* The spine corresponds to Fig. 1)

being a straight line relative to the osculating plane is measured by the geometrical curvature $\kappa(i)$ (Fig. 4):

$$\kappa(i) = \frac{1}{r_\kappa(i)} = \frac{\left\| \frac{\mathrm{d}\mathbf{c}(i)}{\mathrm{d}i} \times \frac{\mathrm{d}^2\mathbf{c}(i)}{\mathrm{d}i^2} \right\|}{\left\| \frac{\mathrm{d}\mathbf{c}(i)}{\mathrm{d}i} \right\|^3}, \tag{5}$$

where $r_\kappa(i)$ is the radius of curvature that represents the radius of the osculating circle in the osculating plane. On the other hand, the deviation of the curve from being a plane curve, represented by the rotation of the unit binormal vector $\hat{\mathbf{b}}(i)$ about the unit tangent vector $\hat{\mathbf{t}}(i)$, is measured by the geometrical torsion $\tau(i)$ (Fig. 5):

$$\tau(i) = \frac{1}{r_\sigma(i)} = \frac{\left(\frac{\mathrm{d}\mathbf{c}(i)}{\mathrm{d}i} \times \frac{\mathrm{d}^2\mathbf{c}(i)}{\mathrm{d}i^2} \right) \cdot \frac{\mathrm{d}^3\mathbf{c}(i)}{\mathrm{d}i^3}}{\left\| \frac{\mathrm{d}\mathbf{c}(i)}{\mathrm{d}i} \times \frac{\mathrm{d}^2\mathbf{c}(i)}{\mathrm{d}i^2} \right\|^2}, \tag{6}$$

where $\cdot$ denotes the dot vector product, and $r_\sigma(i)$ is the radius of torsion. By using the geometrical curvature and torsion, the resulting Frenet-Serret frame can be written in matrix form as:

$$\frac{\mathrm{d}}{\mathrm{d}i} \begin{bmatrix} \hat{\mathbf{t}}(i) \\ \hat{\mathbf{n}}(i) \\ \hat{\mathbf{b}}(i) \end{bmatrix} = \left\| \frac{\mathrm{d}\mathbf{c}(i)}{\mathrm{d}i} \right\| \begin{bmatrix} 0 & \kappa(i) & 0 \\ -\kappa(i) & 0 & \tau(i) \\ 0 & -\tau(i) & 0 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{t}}(i) \\ \hat{\mathbf{n}}(i) \\ \hat{\mathbf{b}}(i) \end{bmatrix}. \tag{7}$$

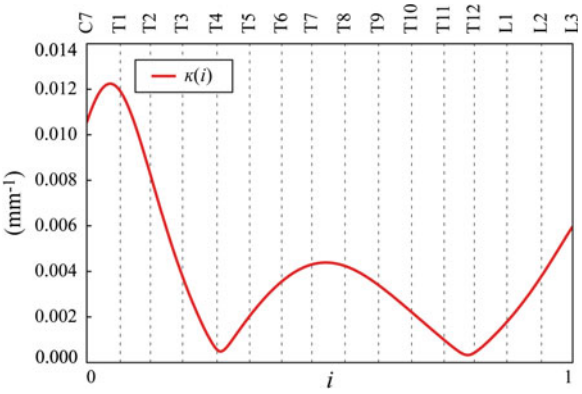


Fig. 4 The geometrical curvature $\kappa(i)$ of the spine curve $\mathbf{c}(i)$ against the independent parameter $i \in [0, 1]$. Labels *C7*, *T1*, …, *L3* indicate vertebral segments (*Note* The spine corresponds to Fig. 1)

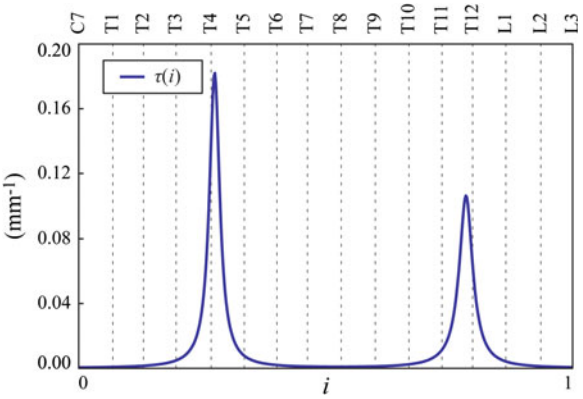


Fig. 5 The geometrical torsion $\tau(i)$ of the spine curve $\mathbf{c}(i)$ against the independent parameter $i \in [0, 1]$. Labels *C7*, *T1*, …, *L3* indicate vertebral segments (*Note* The spine corresponds to Fig. 1)

The form of Eqs. 2–7 corresponds to a regular parametrization of the curve by location i on the curve (Eq. 1). In the case the curve is reparameterized by its arc length s, the natural parametrization $\mathbf{c}(s)$ of $\mathbf{c}(i)$ is yielded:

$$\mathbf{c}(s) = \mathbf{c}(i(s)); \quad i(s) = s^{-1}(i); \quad s(i) = \int_{i_{sp}}^{i} \left\| \frac{\mathrm{d}\mathbf{c}(\lambda)}{\mathrm{d}\lambda} \right\| \mathrm{d}\lambda. \tag{8}$$

Considering the natural parametrization of the curve, the unit tangent vector $\hat{\mathbf{t}}(s)$, unit normal vector $\hat{\mathbf{n}}(s)$ and unit binormal vector $\hat{\mathbf{b}}(s)$ are computed as:

$$\hat{\mathbf{t}}(s) = \frac{\mathrm{d}\mathbf{c}(s)}{\mathrm{d}s}; \quad \hat{\mathbf{n}}(s) = \frac{\frac{\mathrm{d}\hat{\mathbf{t}}(s)}{\mathrm{d}s}}{\left\| \frac{\mathrm{d}\hat{\mathbf{t}}(s)}{\mathrm{d}s} \right\|}; \quad \hat{\mathbf{b}}(s) = \hat{\mathbf{t}}(s) \times \hat{\mathbf{n}}(s), \tag{9}$$

the corresponding geometrical curvature $\kappa(s)$ and torsion $\tau(s)$ are computed as:

$$\kappa(s) = \left\| \frac{\mathrm{d}\hat{\mathbf{t}}(s)}{\mathrm{d}s} \right\| = \left\| \frac{\mathrm{d}^2\mathbf{c}(s)}{\mathrm{d}s^2} \right\|; \quad \tau(s) = -\hat{\mathbf{n}}(\mathbf{s}) \cdot \frac{\mathrm{d}\hat{\mathbf{b}}(s)}{\mathrm{d}s}, \tag{10}$$

and the Frenet-Serret frame in the matrix form is:

$$\frac{\mathrm{d}}{\mathrm{d}s} \begin{bmatrix} \hat{\mathbf{t}}(s) \\ \hat{\mathbf{n}}(s) \\ \hat{\mathbf{b}}(s) \end{bmatrix} = \begin{bmatrix} 0 & \kappa(s) & 0 \\ -\kappa(s) & 0 & \tau(s) \\ 0 & -\tau(s) & 0 \end{bmatrix} \begin{bmatrix} \hat{\mathbf{t}}(s) \\ \hat{\mathbf{n}}(s) \\ \hat{\mathbf{b}}(s) \end{bmatrix}. \tag{11}$$

However, the natural parametrization is in the case of spine curves rare, as it is often based on the axial coordinate z at the start and end point of the spine curve (i.e. $i_{sp} = z_1$ and $i_{ep} = z_2$, where $\mathbf{p}_1 = (x_1, y_1, z_1)$ is the start point and $\mathbf{p}_2 = (x_2, y_2, z_2)$ is the end point), or on pre-defined constant values (e.g. $i_{sp} = 0$ and $i_{ep} = 1$).

2.2.2 Axial Vertebral Rotation

Axial vertebral rotation is the rotation Φ of vertebrae around their longitudinal axes when projected onto the transverse plane of the corresponding coordinate system. If i is an independent parameter that denotes an arbitrary location on the spine, then $\varphi(i)$ is the parametrization of the axial vertebral rotation Φ:

$$\Phi: \quad \varphi(i); \quad i \in [i_{\text{sp}}, i_{\text{ep}}], \tag{12}$$

where $i = i_{\text{sp}}$ and $i = i_{\text{ep}}$ represent the locations on the spine at its start and end point of observation, respectively. For K observed consecutive vertebrae, let angles $\{\delta(k); k = 1, 2, \ldots, K\}$ represent the corresponding axial vertebral rotation angles. The axial vertebral rotation $\varphi(i)$ can be then obtained by continuous interpolation of $\delta(k)$ between $\varphi(i_{\text{sp}}) = \delta(1)$ and $\varphi(i_{\text{ep}}) = \delta(K)$ (Fig. 6).

Axial vertebral rotation cannot be uniformly measured, as the vertebral anatomy observed in the transverse plane is not completely symmetrical due to normal developmental as well as pathological conditions affecting the spine. As a result, several methods were developed to determine axial vertebral rotation from 3D images of the spine [88] that measured the angle between the reference sagittal plane and a line connecting specific anatomical reference points in the transverse plane. For example, axial vertebral rotation was measured as the angle between the reference sagittal plane and the line connecting the posterior junction of the two laminae of the vertebral arch with the center of vertebral body [1], as the angle between the reference sagittal plane and the line bisecting the angle between the two lines connecting the junction of each lamina and the pedicle with the posterior junction of the two laminae [26], as the angle between the reference sagittal plane and the line connecting the tip of the spinous process with the center of vertebral body [39], or as the angle between the reference sagittal plane and the line connecting the most posterior points of the two pedicles [20].

Let points $\mathbf{r}(i) = (r_x(i), r_y(i), r_z(i))$ represent locations of the selected anatomical reference points in the corresponding transverse planes of measurement, and let

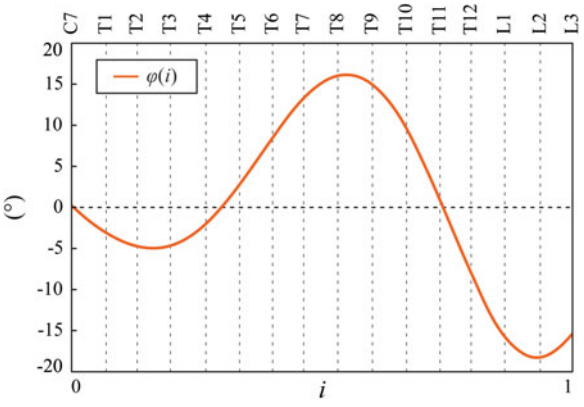


Fig. 6 The axial vertebral rotation $\varphi(i)$ against the independent parameter $i \in [0, 1]$. Labels *C7*, *T1*, …, *L3* indicate vertebral segments (*Note* The spine corresponds to Fig. 1)

$\mathbf{c}(i) = (c_x(i), c_y(i), c_z(i))$ represent the spine curve that passes through the centers of vertebral bodies. If the transverse planes of measurement are image-based, i.e. orthogonal to axis z of the image-based coordinate system, then the axial vertebral rotation $\varphi(i) = \varphi_z(i)$ can be determined by considering $r_z(i) = c_z(i)$ as:

$$\varphi_z(i) = \arctan\left(\frac{r_x(i) - c_x(i)}{r_y(i) - c_y(i)}\right), \tag{13}$$

which for every i corresponds to the angle between the line connecting the anatomical reference point with the center of the vertebral body, and the line representing the reference sagittal plane (i.e. the line in the direction of axis y of the image-based coordinate system). However, because vertebrae can be sagittally or coronally inclined against axis z, the centers of vertebral bodies and the corresponding reference anatomical points may not represent corresponding anatomical locations along the longitudinal vertebral axes. On the other hand, if the transverse planes of measurement are spine-based, i.e. orthogonal to axis w of the spine-based coordinate system and therefore orthogonal to $\mathbf{c}(i)$, then the axial vertebral rotation $\varphi(i) = \varphi_w(i)$ is measured at corresponding anatomical locations and can be determined as:

$$\varphi_w(i) = \arccos\left(\frac{(\mathbf{r}(i) - \mathbf{c}(i)) \cdot \tilde{\mathbf{e}}_{Iy}(i)}{\|\mathbf{r}(i) - \mathbf{c}(i)\|}\right), \quad \tilde{\mathbf{e}}_{Iy}(i) = \hat{\mathbf{t}}(i) \times \left(\hat{\mathbf{e}}_{Iy} \times \hat{\mathbf{t}}(i)\right), \tag{14}$$

where $\tilde{\mathbf{e}}_{Iy}(i)$ is the unit vector in the direction of the projection of $\hat{\mathbf{e}}_{Iy} = [0, 1, 0]_I$ to the plane orthogonal to the spine curve, defined by the unit tangent vector $\hat{\mathbf{t}}(i)$ as the normal of that plane.

2.2.3 Transformation from Image-Based to Spine-Based Coordinate System

The continuous transformation from the image-based coordinate system (Fig. 7) to the spine-based coordinate system (Fig. 8) is possible by having a continuous description of the spine curve C (Eq. 1) and axial vertebral rotation Φ (Eq. 12) that are parameterized by the same variable i representing the location on the spine:

$$\{C, \Phi\}: \quad \zeta(i) = (\mathbf{c}(i), \varphi(i)) = (c_x(i), c_y(i), c_z(i), \varphi(i)); \quad i = [i_{\text{sp}}, i_{\text{ep}}], \tag{15}$$

where $i = i_{\text{sp}}$ and $i = i_{\text{ep}}$ represent the locations of the spine curve and axial vertebral rotation at the start and end point of observation on the spine, respectively.

The Frenet-Serret frame (Eq. 7) describes the geometrical properties of the curve and can be therefore applied to the spine curve $\mathbf{c}(i)$. However, the spine-based coordinate system $(u, v, w) \in \mathbb{R}^3_S$ has to represent also the course of the axial vertebral rotation $\varphi(i)$. The unit tangent vector $\hat{\mathbf{t}}(i)$ defines the unit vector $\hat{\mathbf{e}}_{Sw} = [0, 0, 1]_S$

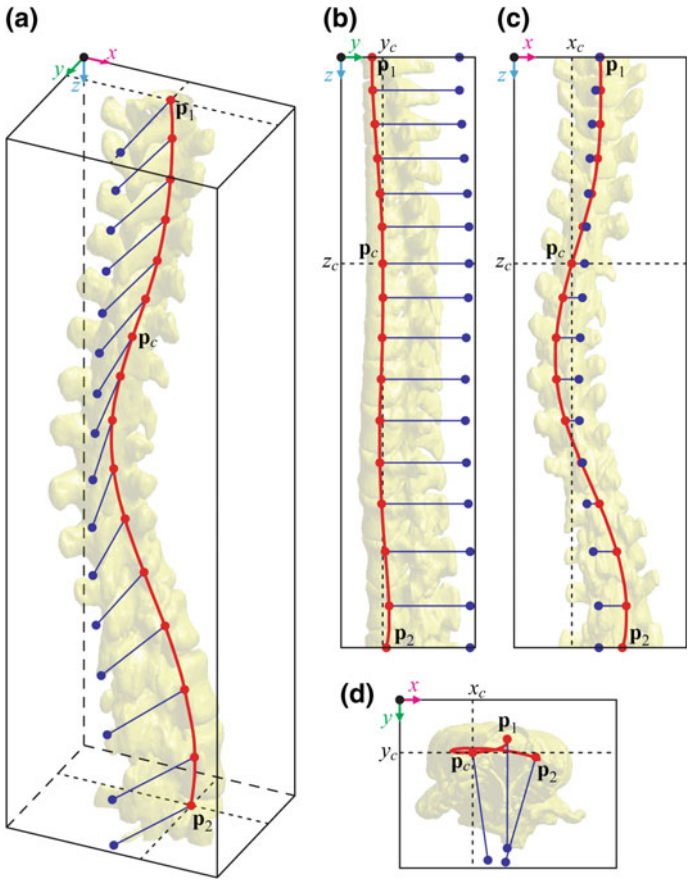


Fig. 7 The spine curve (*red line*) and axial vertebral rotation (*blue directions*) in the image-based coordinate system $\mathbb{R}^3_I$ of a 3D image of a scoliotic spine, shown in **a** 3D view, **b** left sagittal view, **c** posterior coronal view and **d** superior axial view. The spine curve is defined between the start point $\mathbf{p}_1$ and end point $\mathbf{p}_2$, while the coordinates are shown also for a selected point $\mathbf{p}_c = (x_c, y_c, z_c)$ on the spine curve

representing axis w at any point i on the spine curve. At the same time it defines, as the normal, the plane orthogonal to the spine curve at any point i, which is the plane that contains unit vectors $\hat{\mathbf{e}}_{Su} = [1, 0, 0]_S$ and $\hat{\mathbf{e}}_{Sv} = [0, 1, 0]_S$ representing axes u and v, respectively. However, $\hat{\mathbf{e}}_{Su}$ and $\hat{\mathbf{e}}_{Sv}$ do not correspond to the unit normal vector $\hat{\mathbf{n}}(i)$ or unit binormal vector $\hat{\mathbf{b}}(i)$, although they lie in the same plane, i.e. the plane orthogonal to the spine curve, because $\hat{\mathbf{n}}(i)$ and $\hat{\mathbf{b}}(i)$ rotate about $\hat{\mathbf{t}}(i)$ as described by the geometrical torsion $\tau(i)$. The directions of $\hat{\mathbf{e}}_{Su}$ and $\hat{\mathbf{e}}_{Sv}$ are namely defined by the axial vertebral rotation $\varphi(i)$. If $\varphi(i) = \varphi_z(i)$ (Eq. 13), meaning that the axial vertebral rotation is defined in transverse planes of measurement that are orthogonal to axis z of

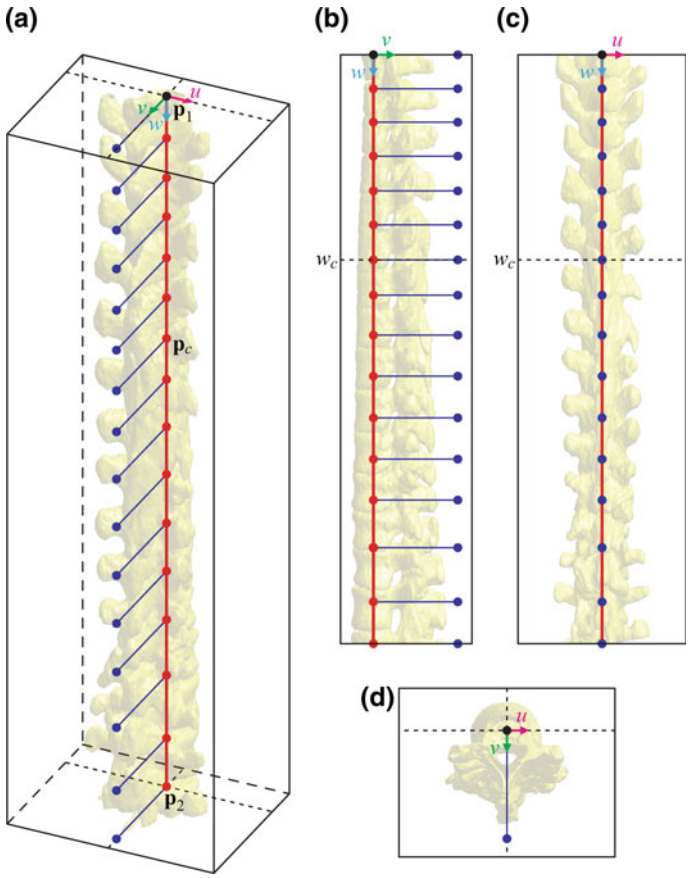


Fig. 8 The spine curve (*red line*) and axial vertebral rotation (*blue directions*) in the spine-based coordinate system $\mathbb{R}^3_S$ of a 3D image of a scoliotic spine, shown in **a** 3D view, **b** left sagittal view, **c** posterior coronal view and **d** superior axial view. The spine curve is defined between the start point $\mathbf{p}_1$ and end point $\mathbf{p}_2$, while the coordinates are shown also for a selected point $\mathbf{p}_c = (u_c, v_c, w_c)$ on the spine curve

the image-based coordinate system, then the modified unit normal vector $\hat{\mathbf{n}}_\varphi(i)$ equals:

$$\hat{\mathbf{n}}_\varphi(i) = \hat{\mathbf{t}}(i) \times (R_z(\varphi_z(i))\hat{\mathbf{e}}_{Iy} \times \hat{\mathbf{t}}(i)), \tag{16}$$

where matrix $R_z(\varphi_z(i))$ represents the rotation for angle $\varphi_z(i)$ about axis z of the image-based coordinate system:

$$R_z(\varphi_z(i)) = \begin{bmatrix} \cos\varphi_z(i) & -\sin\varphi_z(i) & 0 \\ \sin\varphi_z(i) & \cos\varphi_z(i) & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{17}$$

with the center of rotation at point $\mathbf{c}(i) = (c_x(i), c_y(i), c_z(i))$. On the other hand, if $\varphi(i) = \varphi_w(i)$ (Eq. 14), meaning that the axial vertebral rotation is defined in transverse planes of measurement that are orthogonal to the spine curve $\mathbf{c}(i)$, then the modified unit normal vector $\hat{\mathbf{n}}_\varphi(i)$ equals:

$$\hat{\mathbf{n}}_\varphi(i) = R_{\hat{\mathbf{t}}(i)}(\varphi_w(i))\tilde{\mathbf{e}}_{Iy}(i), \tag{18}$$

where $\tilde{\mathbf{e}}_{Iy}(i)$ is the unit vector in the direction of the projection of $\hat{\mathbf{e}}_{Iy} = [0,1,0]_I$ to the plane orthogonal to the spine curve (Eq. 14), and matrix $R_{\hat{\mathbf{t}}(i)}(\varphi_w(i))$ represents the rotation for angle $\varphi_w(i)$ about the axis defined by the unit tangent vector $\hat{\mathbf{t}}(i)$:

$$\begin{aligned} R_{\hat{\mathbf{t}}(i)}(\varphi_w(i)) = &\cos(\varphi_w(i))I_3 \\ &+ \sin(\varphi_w(i))[\hat{\mathbf{t}}(i)]_\times \\ &+ (1 - \cos(\varphi_w(i)))[\hat{\mathbf{t}}(i) \otimes \hat{\mathbf{t}}(i)], \end{aligned} \tag{19}$$

with the center of rotation at point $\mathbf{c}(i) = (c_x(i), c_y(i), c_z(i))$. In Eq. 19, I_3 denotes the identity matrix of size 3×3, and $[\hat{\mathbf{t}}(i)]_\times$ and $[\hat{\mathbf{t}}(i) \otimes \hat{\mathbf{t}}(i)]$ are, respectively, the cross and tensor product matrix of $\hat{\mathbf{t}}(i)$:

$$[\hat{\mathbf{t}}(i)]_\times = \begin{bmatrix} 0 & -\hat{t}_z(i) & \hat{t}_y(i) \\ \hat{t}_z(i) & 0 & -\hat{t}_x(i) \\ -\hat{t}_y(i) & \hat{t}_x(i) & 0 \end{bmatrix}, \tag{20}$$

$$[\hat{\mathbf{t}}(i) \otimes \hat{\mathbf{t}}(i)] = \begin{bmatrix} (\hat{t}_x(i))^2 & \hat{t}_x(i)\hat{t}_y(i) & \hat{t}_x(i)\hat{t}_z(i) \\ \hat{t}_x(i)\hat{t}_y(i) & (\hat{t}_y(i))^2 & \hat{t}_y(i)\hat{t}_z(i) \\ \hat{t}_x(i)\hat{t}_z(i) & \hat{t}_y(i)\hat{t}_z(i) & (\hat{t}_z(i))^2 \end{bmatrix}. \tag{21}$$

In both cases, the unit binormal vector also changes its direction to fit the orthonormal basis, and is therefore equal to $\hat{\mathbf{b}}_\varphi(i) = \hat{\mathbf{t}}(i) \times \hat{\mathbf{n}}_\varphi(i)$. The resulting axes u, v and w of the spine-based coordinate system are therefore represented by:

$$u: \quad \hat{\mathbf{e}}_{Su} = [1,0,0]_S \quad \leftrightarrow \quad \hat{\mathbf{b}}_\varphi(i), \tag{22}$$

$$v: \quad \hat{\mathbf{e}}_{Sv} = [0,1,0]_S \quad \leftrightarrow \quad \hat{\mathbf{n}}_\varphi(i), \tag{23}$$

$$w: \quad \hat{\mathbf{e}}_{Sw} = [0,0,1]_S \quad \leftrightarrow \quad \hat{\mathbf{t}}(i). \tag{24}$$

In contrast to the image-based coordinate system, which is defined in the Euclidean space, the spine-based coordinate system is defined in a non-Euclidean space, and therefore morphometric measurements based on Euclidean metrics can not be obtained directly from the spine-based coordinate system.

2.3 Automated Determination of the Spine-Based Coordinate System

The spine-based coordinate system can be manually determined by identifying distinctive anatomical points on each vertebra (e.g. the centers of vertebral bodies) and the corresponding rotation of vertebrae, and then interpolating through these points to obtain a continuous description of both the spine curve and axial vertebral rotation along the whole length of the spine. However, navigation through 3D spine images is time-consuming and subjective, moreover it is practically impossible to manually define the plane orthogonal to the spine curve, in which axial vertebral rotation is defined, basing only on the identified anatomical points on each vertebra. As a result, several automated and semi-automated methods based on image processing and analysis techniques were proposed to determine the spine curve and/or axial vertebral rotation in 3D spine images.

2.3.1 Automated Determination of the Spine Curve

In the past, the only possibility for measuring the geometrical properties of the spine curve was based on examining the antero-posterior and/or lateral radiographs. As a result, the spine curve in 3D was observed as its projection in 2D in the form of coronal and sagittal spinal curvatures. Moreover, these curvatures were usually evaluated by one-dimensional measures including angles of curvature (e.g. the Ferguson angle, the Cobb angle, the tangent line angle) and indices of curvature (e.g. the Greenspan index, the Ishihara index). A detailed review of methods for the determination of spinal curvature was performed by Vrtovec et al. [89].

With the development of 3D imaging techniques, methods that captured the 3D nature of the spine started to emerge, followed by application of computerized techniques that automatically or semi-automatically (i.e. with minimal observer interaction) determined the spine curve in 3D images. Due to the continuous course of the spinal curvature, a number of studies attempted to model the spine curve with a mathematical curve in stereoradiographic (i.e. in multiple radiographs acquired at different angles), CT or MR spine images. Different functions were used for modeling, such as harmonic functions (i.e. sines, cosines or Fourier series) [13, 15, 27, 58, 74], spline functions [6, 33, 55, 82] and standard polynomial functions [56, 83, 85, 86], as well as statistical interpolation techniques, such as kriging [59].

By computerized least-squares aligning of a parametric sine function to the stereographically reconstructed landmarks, Stokes et al. [74] measured the Cobb angle between the normals to the obtained curve at inflection points in the coronal and sagittal plane, and in the plane of maximal curvature. Drerup and Hierholzer [13–15] also considered the sine function appropriate, as it most resembled the appearance of curves in idiopathic scoliosis. On the other hand, Patwardhan et al. [55] justified the use of spline functions by stating that splines are used to describe geometries with continuously changing curvature, such as scoliotic spines. In their framework for spine segmentation from CT images, Kaminsky et al. [33] used spline functions because they proved appropriate to describe both the anatomical shape and scoliotic deformations of the spine. Berthonnaud and Dimnet [6] constructed the spine curve separately in coronal and sagittal projections by computing the average of two spline functions that connected the anatomical landmarks on vertebral body walls. On the other hand, Peng et al. [56] used polynomial functions to detect and segment vertebrae from MR images using vertebral disc templates. Polynomial functions were also used to model both normal and pathological spine curves in CT images by Vrtovec et al. [83]. The spine curve was automatically determined by aligning the polynomial function with the centers of vertebral bodies in 3D, obtained by maximizing the distance from the edges of vertebral bodies. The same authors also developed a method for MR images [86], where the center of vertebral body was first automatically detected in each axial cross-section by maximizing the entropy of image intensities inside a circular region, and the detected centers of vertebral bodies in 3D were then joined by a polynomial function using the robust least-trimmed-squares regression. The method was also used by Neubert et al. [49] for extracting the spine curve from MR images of high resolution, obtained by applying the sequence named sampling perfection with application optimized contrasts using different flip angle evolution (SPACE). The work was continued by Štern et al. [78], who proposed a modality-independent method for the determination of the spine curve that was extracted from locations where lines connecting opposite edge points on vertebral body walls in the direction of corresponding image intensity gradients most often intersected. Kadoury et al. [30, 31] determined the spine curve of a scoliotic spine in biplanar radiographs by first embedding cubic B-spline functions onto a non-linear manifold to predict an initial curve according to a given database of scoliotic curves, and then performing analytical regression to obtain a statistical model of the final curve.

To extract the spinal canal centerline from CT images, Yao et al. [96] applied a watershed algorithm followed by a graph search, while Hay et al. [24] applied the fast marching minimal path technique that was based on the distance transform of the spinal canal segmentation, obtained by morphological region growing. On the other hand, Klinder et al. [36] segmented the spinal canal by a progressive adaptation of small tubular segments, represented as triangulated surface meshes, and then determined the spinal canal centerline by calculating the centers of mass for all contours of the obtained tubular mesh. A similar approach was proposed by Forsberg et al. [17, 18], who extracted the spinal canal centerline in CT images by

first modeling the vertebral foramen in 2D axial cross-sections with circles, and then fitting a cubic B-spline function to the centers of the obtained circles.

Besides modeling the spine curve in 3D, different geometrical descriptors of spinal curvature were derived from mathematical functions. Poncet et al. [58] proposed the geometrical torsion as a measure for classifying spinal deformities, and Kadoury et al. [29, 32] further showed that it can be potentially used to discriminate among different types of thoracolumbar deformations of the spine. Vrtovec et al. [85] showed that clinically relevant features of the spine can be identified in 3D by observing the geometrical curvature as well as the curvature angle, which was defined as the angular magnitude of the geometrical curvature on an arbitrary spine section, and was as such independent of the size of the spine. Hay et al. [24] observed both the geometrical curvature and geometrical torsion that were scaled to a subject-independent coordinate system, and showed that they can be used to detect and quantify pathological spinal curvatures.

2.3.2 Automated Determination of the Axial Vertebral Rotation

Similarly as for the spinal curvature, measurements of axial vertebral rotation were in the past possible only by examining the location of pedicles and spinous processes in relation to corresponding vertebral bodies in antero-posterior radiographs. As a result, the axial vertebral rotation in 3D was observed as its projection in 2D, and several methods based on indices (e.g. the Cobb method, the Nash-Moe method, the Fait-Janovec method) or actual angles (e.g. the Bunnell method, the Drerup method, the Stokes et al. method) were proposed. With the introduction of 3D imaging techniques, cross-sectional imaging in the axial plane became possible and stimulated the development of methods that were based on manual identification of distinctive anatomical reference points (e.g. the tip of spinous process, the center of the vertebral body, etc.). A detailed review of methods for the determination of axial vertebral rotation was performed by Lam et al. [40] and Vrtovec et al. [88].

The measurement of axial vertebral rotation was also approached by computerized techniques based on image processing and analysis, although manual initialization was still required. Haughton et al. [23] and Rogers et al. [66, 67] proposed a method that required manual determination of the axial CT [66] or MR [23, 67] cross-section, the center of rotation and the circular area that encompassed the observed lumbar vertebra. After initialization, the method automatically measured the axial vertebral rotation relative to the cross-section of a different vertebra by searching for the maximal correlation of image intensities between the circular areas determined in both cross-sections. Oblique CT cross-sections were used by Adam and Askin [2], who determined the axial vertebral rotation from the line that bisected the thresholded image of the vertebral body according to the symmetry ratio, defined by the maximal correlation of image intensities in the bisected regions. Kouwenhoven et al. [37, 38] manually selected axial cross-sections through the centers of vertebral bodies in CT [38] and MR [37] images of normal

spines, and applied automated region growing segmentation to obtain reference points, such as the center of the vertebral canal, the center of the sternum at the T5 vertebra and the center of the anterior half of the vertebral body, which were used to define the axial vertebral rotation. Axial vertebral rotation was studied in both CT and MR images of whole spines also by Vrtovec et al. [83, 86]. In CT images [83], circular cross-sections that were orthogonal to the spine curve were first automatically extracted, and axial vertebral rotation was then defined from the line that bisected the cross-section and resulted in the maximal correlation of image intensities in the bisected regions. For MR images [86], the rotation was defined in an optimization procedure that searched for the orientation angle of the line of symmetry in each axial-cross section, and then smoothed with a polynomial function along the whole spine using the least-trimmed-squares regression technique. The same authors also combined both approaches into a method that was modality-independent, i.e. applicable to both CT and MR images [87]. Basing on the predefined location of the vertebral body center in 3D, they obtained the relation between the image-based and vertebra-based coordinate systems by matching image intensity gradients that defined the best available symmetry of vertebral anatomical structures. The method was thoroughly evaluated and compared to established manual methods when applied to CT [91] and MR [90] images of normal and scoliotic spines. To segment vertebral bodies in both CT and MR images, Štern et al. [79] proposed to use a parametric model based on superquadrics that, among several shape parameters, included also the axial rotation of the vertebral body, and which was later used to perform quantitative vertebral morphometry in CT images of normal and fractured vertebrae [80]. Axial vertebral rotation was determined from the symmetry of vertebral anatomical structures also in the study of Forsberg et al. [18], who for each vertebra in CT images extracted a cross-section that was orthogonal to the spine curve and passed through the center of the vertebral body, and then minimized the sum of absolute differences in image intensities over the line that bisected the cross-section at the evaluated rotation angle. The same group of authors also developed a method for segmentation of vertebrae by registering a spine model to CT spine images, and then measured axial vertebral rotation from landmarks that were placed at distinctive anatomical locations in the spine model and mapped to each CT image by using the obtained registration transformation fields [17].

2.3.3 Examples of Automated Determination of the Spine Curve and Axial Vertebral Rotation

Among automated methods for the determination of the spine curve $\mathbf{c}(i)$ and/or axial vertebral rotation $\varphi(i)$, the following approaches are presented in detail:

- *automated determination of the spine curve and axial vertebral rotation in CT images* [83] (section Automated Determination of the Spine Curve and Axial Vertebral Rotation in CT Images),

- *automated determination of the spine curve and axial vertebral rotation in MR images* [86] (section Automated Determination of the Spine Curve and Axial Vertebral Rotation in MR Images),
- *automated modality-independent determination of the spine curve and axial vertebral rotation in 3D images* [78, 83, 86] (section Automated Modality-Independent Determination of the Spine Curve and Axial Vertebral Rotation in 3D Images).

In all of the presented approaches [78, 83, 86], the spine-based coordinate system is determined from 3D spine images of normal and scoliotic subjects by parameterizing the spine curve $\mathbf{c}(i) = (c_x(i), c_y(i), c_z(i))$ and axial vertebral rotation $\varphi(i)$ with polynomial functions:

$$\mathbf{c}(i) = \left(\sum_{k=0}^{K_x} \frac{b_{x,k}}{\hat{b}_k} i^k, \sum_{k=0}^{K_y} \frac{b_{y,k}}{\hat{b}_k} i^k, \sum_{k=0}^{K_z} \frac{b_{z,k}}{\hat{b}_k} i^k \right), \tag{25}$$

$$\varphi(i) = \sum_{k=0}^{K_\varphi} \frac{b_{\varphi,k}}{\hat{b}_k} i^k, \tag{26}$$

where $b_x = \{b_{x,k}; k = 0, 1, \ldots, K_x\}$, $b_y = \{b_{y,k}; k = 0, 1, \ldots, K_y\}$ and $b_z = \{b_{z,k}; k = 0, 1, \ldots, K_z\}$ are the parameters of polynomial functions $c_x(i)$, $c_y(i)$ and $c_z(i)$ of degrees K_x, K_y and K_z, respectively, corresponding to the spine curve $\mathbf{c}(i)$, and $b_\varphi = \{b_{\varphi,k}; k = 0, 1, \ldots, K_\varphi\}$ are the parameters of the polynomial function $\varphi(i)$ of the degree K_φ corresponding to the axial vertebral rotation $\varphi(i)$. The normalization coefficients $\hat{b}_k$:

$$\hat{b}_k = \int_{i_{sp}}^{i_{ep}} |i^k| \mathrm{d}i, \tag{27}$$

where $i = i_{sp}$ and $i = i_{ep}$ represent the locations on the spine at its start and end point of observation, respectively, regularize the impact of each polynomial parameter to the absolute variation of the corresponding term. With such parametrization, the goal is to automatically determine polynomial parameters $b_\mathbf{c} = b_x \cup b_y \cup b_z$ that describe the spine curve $\mathbf{c}(i)$, and polynomial parameters b_φ that define the axial vertebral rotation $\varphi(i)$ in the given 3D spine image.

Moreover, it is assumed that the 3D image is cropped to a volume of interest according to the start and end point of observation along axis z of the image-based coordinate system, so that the resulting cropped 3D image contains only axial cross-sections that display the observed anatomy of the spine. Although the presented examples may be therefore labeled as semi-automated, manual determination of the volume of interest does not represent a very demanding or time-consuming task. An advantage of such an assumption is that the parametrization of the spine curve and

axial vertebral rotation can be based on axial pixel coordinates z at the start and end point of observation, resulting in $i_{sp} = 1$ and $i_{ep} = Z$, respectively, with the corresponding number of samples equal to $N = i_{ep} - i_{sp} + 1 = Z$, where Z is the number of axial cross-sections in the 3D image. Such parametrization is, considering the usual in-plane resolution and slice thickness of CT and MR spine images, in general sufficient for a smooth and continuous description of the spine curve $\mathbf{c}(i)$ and axial vertebral rotation $\varphi(i)$.

Automated Determination of the Spine Curve and Axial Vertebral Rotation in CT Images

Vrtovec et al. [83] proposed a method for automated determination of the spine curve and axial vertebral rotation in CT images. If the spine curve $\mathbf{c}(i)$ is represented by a curve that passes through the centers of vertebral bodies, then its determination can be based on the anatomical property that vertebral bodies are locally the largest bone structures of the spine, and on the geometrical property that the center of each vertebral body is represented by the point that is most distant from corresponding edges of the vertebral body. To obtain a quantitative representation of these properties, a distance transform function based on Euclidean metrics is applied twice to image $I = I(\mathbf{p})$, resulting in distance map $D_I = D_I(\mathbf{p})$:

$$D_I(\mathbf{p}) = \begin{cases} +d_<(\mathbf{p}, \tilde{\mathbf{p}}); & I(\mathbf{p}) \geq T, \quad I(\tilde{\mathbf{p}}) < T, \\ -d_\geq(\mathbf{p}, \tilde{\mathbf{p}}); & I(\mathbf{p}) < T, \quad I(\tilde{\mathbf{p}}) \geq T, \end{cases} \tag{28}$$

where $d_<(\mathbf{p}, \tilde{\mathbf{p}})$ and $d_\geq(\mathbf{p}, \tilde{\mathbf{p}})$ are the Euclidean distances between the observed point $\mathbf{p} = (x, y, z)$ and point $\tilde{\mathbf{p}} = (\tilde{x}, \tilde{y}, \tilde{z})$, which represents the closest point to $\mathbf{p}$ with $I(\tilde{\mathbf{p}}) < T$ and $I(\tilde{\mathbf{p}}) \geq T$, respectively. The image intensity threshold T, which separates the bone structures from the background, can be in the case of CT images determined from the corresponding Hounsfield values. Each value at point $\mathbf{p}$ in the resulting distance map D_I, which is of the same size as image I, represents the Euclidean distance from $\mathbf{p}$ to the edges of the bone structures, and this distance is positive when $\mathbf{p}$ is located inside and negative when $\mathbf{p}$ is located outside the bone structures (Fig. 9). As vertebral bodies are locally the largest bone structures of the spine, distance map values are expected to be the highest in geometrical centers of vertebral bodies and smoothly decrease by moving away from the centers. The optimal polynomial parameters $b_{\mathbf{c}}^*$ that define the spine curve $\mathbf{c}(i)$ (Eq. 25) are finally obtained in an optimization procedure that searches for the combination of parameters that corresponds to the maximal sum of distance map values along the spine curve:

$$b_{\mathbf{c}}^* = \arg\max_{b_{\mathbf{c}}} \left(\sum_{i=1}^{N} D_I(\mathbf{c}(i)|b_{\mathbf{c}}) \right). \tag{29}$$

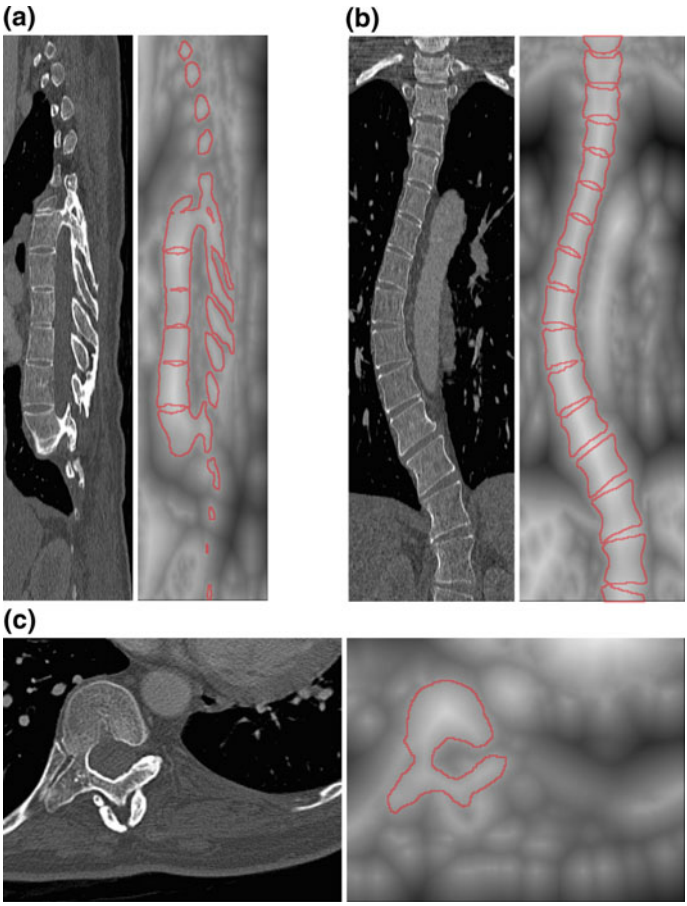


Fig. 9 Original image I (*left*) and the corresponding distance transform D_I with superimposed spine boundaries (*right*), displayed for a selected **a** sagittal, **b** coronal and **c** axial cross-section of a 3D CT image of a scoliotic spine. In the distance transform D_I, *brighter elements* represent positive distances, while *darker elements* represent negative distances

In order to increase the optimization robustness, two additional mechanisms can be applied. First, the amount of information taken into account during optimization can be increased by considering distance map values within a radius from the spine curve, which can be defined from the quantitative morphometrical vertebral analysis [52–54]. Second, optimization can be designed hierarchically and performed on multiple levels. On the first level, the polynomial functions are initialized with degrees of $K_x = K_y = K_z = 1$, i.e. as straight lines. When the optimization reaches the termination criterion, polynomial degrees are increased by 1 and the optimization is restarted on the next level, using the polynomial parameters from the previous level for initialization (optionally, the degree of the polynomial function $c_z(i)$ can be fixed to $K_z = 1$, therefore always representing a straight line that

describes the longitudinal axis of the spine). The maximal polynomial degree can be determined from the flexion points in normal and scoliotic spinal curvatures, as the number of flexion points of a polynomial function is equal to its degree decreased by one (e.g. a straight line is of the first degree and has zero flexion points). For example, when observing the thoracolumbar section of the spine, three distinctive flexion points exist in a normal spinal curvature, i.e. the maximal thoracic kyphosis, the thoracolumbar junction, and the maximal lumbar lordosis, and therefore polynomial functions of the fourth degree represent an adequate choice. Accordingly, a scoliotic spinal curvature can be adequately described by polynomial functions of the fifth degree.

As the axial vertebral rotation is defined as the rotation of vertebrae about the spine curve, it can be determined automatically in planes that are orthogonal to the spine curve $\mathbf{c}(i)$, defined by polynomial parameters $b^*_{\mathbf{c}}$ (Eq. 29). In each such ith plane, a circular domain of radius d and centered at point $\mathbf{c}(i) = (c_x(i), c_y(i), c_z(i))$ is defined (Fig. 10a). The circular domain is then bisected by a line that passes through the center of the domain and is inclined for angle φ against the projection $\tilde{\mathbf{e}}_{Iy}(i)$ of the unit vector $\hat{\mathbf{e}}_{Iy} = [0, 1, 0]_I$ to the plane (Eq. 14). In the obtained halves A and B of the ith circular domain, s_A and s_B represent image intensities at mirror pixels according to the line of bisection:

$$\begin{aligned} s_A(i, j, \varphi) &= I(R_z(\varphi)[-u, v, c_z(i)]) \\ &= I(R_{\hat{\mathbf{t}}(i)}(\varphi)[-x, y, 0] + [c_x(i), c_y(i), c_z(i)]), \end{aligned} \tag{30}$$

$$\begin{aligned} s_B(i, j, \varphi) &= I(R_z(\varphi)[+u, v, c_z(i)]) \\ &= I(R_{\hat{\mathbf{t}}(i)}(\varphi)[+x, y, 0] + [c_x(i), c_y(i), c_z(i)]), \end{aligned} \tag{31}$$

where j is the index of the mirror point pair (a total of J mirror point pairs exist), consecutively assigned on the basis of each (u, v) with $u > 0$ and $u^2 + v^2 \leq d^2$, or each (x, y) with $x > 0$ and $x^2 + y^2 \leq d^2$ (Fig. 10b). Matrices R_z (Eq. 17) and $R_{\hat{\mathbf{t}}(i)}$ (Eq. 19) represent, respectively, the rotation[1] about axis w of the spine-based coordinate system and the rotation about the unit tangent vector $\hat{\mathbf{t}}(i)$ to the spine curve $\mathbf{c}(i)$ at point i in the image-based coordinate system.

Radius d is defined so that the anatomy of the whole vertebra (and not only of the vertebral body) is captured within the circular domain, and can be defined from the quantitative morphometrical vertebral analysis [52–54]. From the obtained mirror image intensity pairs, the in-plane similarity between the two mirror halves of the ith circular domain at inclination φ can be quantitatively evaluated by the correlation coefficient $R_{AB}(i, \varphi)$:

[1] It is assumed that point $\mathbf{p} = [x, y, z]$ is a column vector. If $\mathbf{p}$ is a row vector, vector transpose operation is required, therefore the equation $\mathbf{p}' = (x', y', z') = R[x, y, z] = R\mathbf{p}$ turns into $\mathbf{p}' = (x', y', z') = \left(R[x, y, z]^T\right)^T = (R\mathbf{p}^T)^T$.

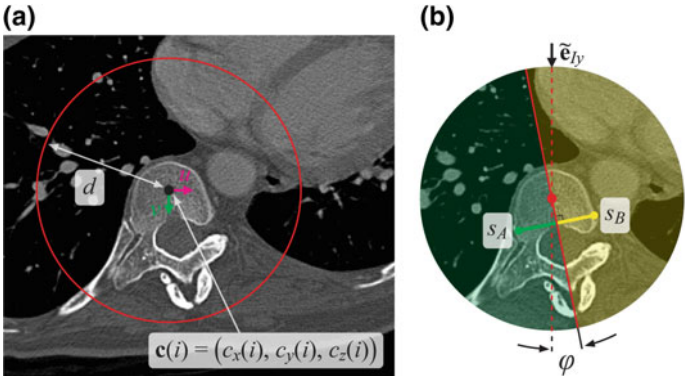


Fig. 10 **a** A circular domain of radius d and centered at the point on the spine curve $\mathbf{c}(i)$, defined in the plane orthogonal to the spine curve $\mathbf{c}(i)$. **b** For a bisecting line that is inclined for an angle φ against the direction of $\tilde{\mathbf{e}}_{Iy}$, image intensities s_A and s_B at mirror point pairs are used to evaluate the similarity between the two halves of the circular domain

$$R_{AB}(i,\varphi) = \frac{\sum_{j=1}^{J}(s_A(i,j,\varphi) - \bar{s}_A(i,\varphi))(s_B(i,j,\varphi) - \bar{s}_B(i,\varphi))}{\sqrt{\sum_{j=1}^{J}(s_A(i,j,\varphi) - \bar{s}_A(i,\varphi))^2 \sum_{j=1}^{J}(s_B(i,j,\varphi) - \bar{s}_B(i,\varphi))^2}}, \tag{32}$$

where $\bar{s}_A(i,\varphi) = \frac{1}{J}\sum_{j=1}^{J} s_A(i,j,\varphi)$ and $\bar{s}_B(i,\varphi) = \frac{1}{J}\sum_{j=1}^{J} s_B(i,j,\varphi)$ are the mean image intensities in parts A and B, respectively, of the circular domain. It is important to note that features other than image intensities can be extracted at mirror points (e.g. image intensity gradients), and similarity measures other than the correlation coefficient can be computed (e.g. mutual information) for the corresponding circular domain halves. Nevertheless, the domain must be circular so that irrespectively of inclination φ, the same points are taken into account for the computation of the in-plane similarity. The optimal polynomial parameters b_φ^* that define the axial vertebral rotation $\varphi(i)$ (Eq. 26) are finally obtained in an optimization procedure that searches for the combination of parameters that corresponds to the maximal sum of correlation coefficients along the spine curve:

$$b_\varphi^* = \arg\max_{b_\varphi}\left(\sum_{i=1}^{N} R_{AB}(i,\varphi)|b_\varphi\right). \tag{33}$$

Similarly as in the case of the spine curve, optimization can be designed hierarchically and performed on multiple levels. However, in the case of the axial vertebral rotation, on the first level the optimization starts with a zero degree polynomial, i.e. $K_\varphi = 0$, and with polynomial parameter $b_{\varphi,0} = 0$, meaning that the initial axial vertebral rotation is constant and equal to zero along the whole length of

the spine. When the optimization reaches the termination criterion, the polynomial degree is increased by 1 and the optimization is restarted on the next level, using the polynomial parameters from the previous level for initialization. The maximal polynomial degree of 5 is adequate for modeling the axial vertebral rotation in both normal and scoliotic spines.

The method was evaluated [83, 85] on 30 normal and one scoliotic CT image of the thoracolumbar spine, and the reported mean difference between the obtained spine curve in 3D and manually defined ground truth points was 2.1 ± 1.4 mm. The performance of the determination of the axial vertebral rotation was, using the sum of absolute differences instead of the correlation coefficient (Eq. 32), evaluated [18] on 68 vertebrae extracted from CT images, and the reported mean difference against reference values was around $-0.6°$ with the corresponding 95% confidence interval of around $4.8°$.

Automated Determination of the Spine Curve and Axial Vertebral Rotation in MR Images

Vrtovec et al. [86] proposed a method for automated determination of the spine curve and axial vertebral rotation in MR images. The method is based on anatomical properties that vertebral bodies and intervertebral discs are nearly circular in shape and that vertebrae are nearly symmetrical over the lines that pass through the centers of vertebral bodies. In each axial cross-section $z = z_i$ of the MR image (Fig. 11a), an arbitrary in-plane line $y_i = y_i(x)$ that divides the observed cross-section into two image parts, i.e. part A and part B, can be defined as:

$$y_i(x) = (x - \lambda_i) \tan\left(\frac{\pi}{2} - \gamma_i\right), \tag{34}$$

where $\tan\left(\frac{\pi}{2} - \gamma_i\right)$ is the slope and λ_i is the intersection of the line with axis x of the image-based coordinate system (Fig. 11b). The line that splits the observed cross-section into two symmetrical parts can be obtained by maximizing the similarity between image parts A and B:

$$\{\gamma_i^*, \lambda_i^*\} = \arg\max_{\{\gamma_i, \lambda_i\}} (S(I(\mathbf{p}_A, z_i), I(\mathbf{p}_B, z_i))), \tag{35}$$

where $\mathbf{p}_A = (x_A, y_A); \forall \mathbf{p} = (x, y) > y_i(x)$ and $\mathbf{p}_B = (x_B, y_B); \forall \mathbf{p}_B = (x, y) < y_i(x)$ represent all points $\mathbf{p} = (x, y)$ that lie in image parts A and B, respectively, and S is the similarity measure that can be computed as the standard mutual information between image parts A and B:

$$S(I(\mathbf{p}_A, z_i), I(\mathbf{p}_B, z_i)) = \sum_{\mathbf{p}_A \in A} \sum_{\mathbf{p}_B \in B} p_Q(\mathbf{p}_A, \mathbf{p}_B) \log\left(\frac{p_Q(\mathbf{p}_A, \mathbf{p}_B)}{p_Q(\mathbf{p}_A) p_Q(\mathbf{p}_B)}\right), \tag{36}$$

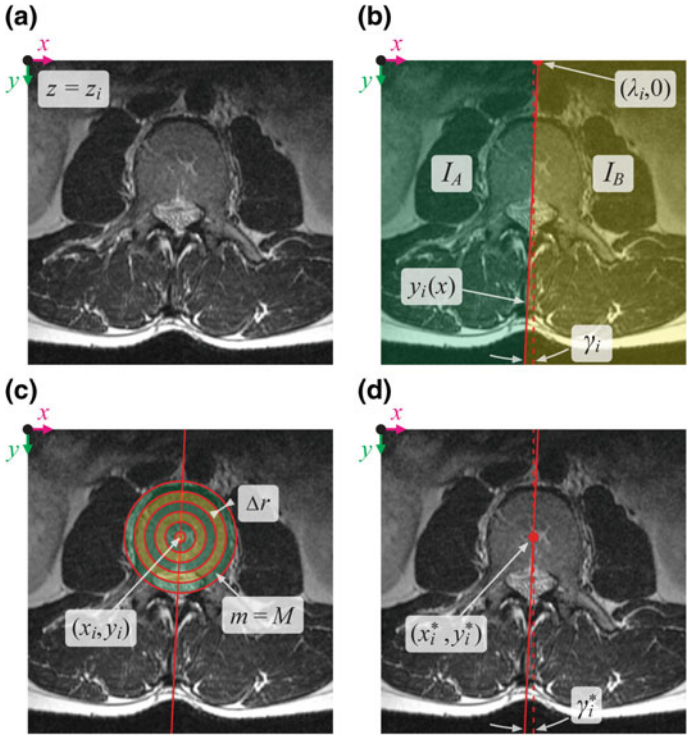


Fig. 11 **a** The search for the line of in-plane symmetry is performed in each *i*th axial orthogonal multi-planar cross-section of the MR spine image. **b** The in-plane line $y_i(x)$ is defined by parameters γ_i and λ_i. **c** The operator Γ is centered in (x_i, y_i) and consists of M *concentric rings*; $m = 1, 2, \ldots, M$, each with radial width of Δr. **d** The resulting center of the vertebral body (x_i^*, y_i^*) and in-plane rotation γ_i^*

where $p_Q(\mathbf{p}_A)$ and $p_Q(\mathbf{p}_B)$ are the probability distributions of image intensities in image parts A and B, respectively, $p_Q(\mathbf{p}_A, \mathbf{p}_B)$ is the joint probability distribution of image intensities, and Q is the number of bins used for probability estimation. The resulting parameters γ_i^* and λ_i^* (Eq. 34) define the in-plane line of symmetry, which passes through the center of the vertebral body in the observed *i*th axial cross-section.

To determine the exact location of the center of the vertebral body, the shape properties of the vertebral anatomy are combined with the appearance of the vertebral body in MR images. When observed in axial cross-sections, the shape of the vertebral body is relatively circular. For a circular structure displayed in a MR image, a certain amount of variation in image intensities is always present along any radial direction, however, in the tangential direction, the variation in image intensities is relatively small. To obtain a quantitative estimation of these properties, the entropy-based operator Γ is introduced that is centered at an arbitrary point $\mathbf{p} = (x, y)$ and consists

of M concentric rings of radii $\{r_m; m = 1, 2, \ldots, M\}$ with $\forall m: r_m < r_{m+1}$ and radial width of each ring equal to $\Delta r = r_{m+1} - r_m$ (Fig. 11c):

$$\Gamma = \frac{\sum_{m=1}^{M} w_m H_m}{H \sum_{m=1}^{M} w_m}; \quad w_m = \exp\left(-\frac{1}{2}\left(\frac{m}{M}L\right)\right), \tag{37}$$

where $H_m = -\sum_{q=1}^{Q} p_{q,m} \log p_{q,m}$ is the entropy defined by the probability distribution $p_{q,m}$ of image intensities in the mth ring, $H = -\sum_{q=1}^{Q} p_q \log p_q$ is the entropy defined by the probability distribution p_q of image intensities within the entire operator (i.e. within all rings), and Q is the number of bins used for probability estimation. The ring weights w_m are chosen to be within L standard deviations of the Gaussian distribution, so that the inner rings have a relatively stronger impact to the operator response in comparison to the outer rings. The number of rings can be automatically adjusted from $M = 15$ rings in the cervical region, to $M = 20$ rings in the thoracic region, and to $M = 30$ rings in the lumbar region of the spine. The radial width of each ring can be set to $\Delta r = 1$ mm, the ring weights to be within $L = 2$ standard deviations of the Gaussian distribution, and the probability distributions can be computed using $Q = 16$ bins. The variation in image intensities in the tangential direction is estimated by the sum of entropies H_m in individual concentric rings, while the variation in image intensities in the radial direction is estimated by the entropy H within the entire operator, which also serves to penalize the regions that are homogeneous in image intensity. The center of the vertebral body (x_i^*, y_i^*) is found by minimizing the response of the entropy-based operator Γ along the in-plane line of symmetry $y_i(x)$:

$$x_i^* = \arg\min_x (I(x, y_i(x), z_i) | \Gamma(x, y_i(x))), \tag{38}$$

$$y_i^* = y_i(x_i^*) = \left(x_i^* - \lambda_i^*\right) \tan\left(\frac{\pi}{2} - \gamma_i^*\right). \tag{39}$$

The resulting points $\{\mathbf{c}_i = (x_i^*, y_i^*, z_i); i = 1, 2, \ldots, N\}$ represent the detected centers of vertebral bodies in each ith axial cross-section of the MR spine image (Fig. 11d). A continuous representation of the spine curve $c(i)$ is obtained by fitting polynomial functions to points $\{\mathbf{c}_i\}$ to determine the optimal polynomial parameters $b_{\mathbf{c}}^*$ (Eq. 25). However, as each point in $\{\mathbf{c}_i\}$ is obtained independently and therefore outliers may be present, it is recommended to apply a robust regression method, for example, the non-linear least trimmed squares (LTS) regression [69]:

$$b_{\mathbf{c}}^* = \arg\min_{b_{\mathbf{c}}} \left(\sum_{i=1}^{h_{\mathbf{c}}} r_{\mathbf{c},[i]}^2 | b_{\mathbf{c}} \right), \tag{40}$$

where $r_{\mathbf{c},[i]}^2 = (\mathbf{c}_i - \mathbf{c}(i))^2$ represent the ordered squared residuals in increasing order; $r_{\mathbf{c},[1]}^2 \leq r_{\mathbf{c},[2]}^2 \leq \cdots \leq r_{\mathbf{c},[N]}^2$, and $h_{\mathbf{c}}$ is the trimming constant that satisfies the

condition $0.5 < \frac{h_c}{N} \leq 1$ and determines the number of ordered residuals used in the computation.

Basing on the obtained spine curve $\mathbf{c}(i)$, planes that are orthogonal to the spine curve can be extracted from the MR image for each location i on the spine. In these planes, axial vertebral rotations $\{\varphi_i; i = 1, 2, \ldots, N\}$ are computed by finding the in-plane lines of symmetry (Eqs. 34–36) and using angles $\{\gamma_i^*; i = 1, 2, \ldots, N\}$ as initialization values. A continuous representation of the axial vertebral rotation $\varphi(i)$ is then obtained by fitting a polynomial function to the resulting angles $\{\varphi_i\}$. To determine the optimal polynomial parameters b_φ^* (Eq. 26), the non-linear LTS regression can be again applied:

$$b_\varphi^* = \arg\min_{b_\varphi} \left(\sum_{i=1}^{h_\varphi} r_{\varphi,[i]}^2 | b_\varphi \right), \tag{41}$$

where $r_{\varphi,[i]}^2 = (\varphi_i - \varphi(i))^2$ represent the ordered squared residuals in increasing order; $r_{\varphi,[1]}^2 \leq r_{\varphi,[2]}^2 \leq \ldots \leq r_{\varphi,[N]}^2$, and h_φ is the trimming constant that satisfies the condition $0.5 < \frac{h_\varphi}{N} \leq 1$ and determines the number of ordered residuals used in the computation. Two thirds of ordered residuals can be used to determine the optimal polynomial parameters for the spine curve (i.e. $h_\mathbf{c} = \frac{2}{3}N$ in Eq. 40) and axial vertebral rotation (i.e. $h_\varphi = \frac{2}{3}N$ in Eq. 41).

The method was evaluated [86] on 21 MR images of the thoracic and lumbar region of the spine, and the reported mean difference between the obtained spine curve in 3D and manually defined ground truth points was 2.5 ± 1.1 mm, while the reported mean difference between the obtained axial vertebral rotation and manually defined ground truth angles was $1.7 \pm 0.9°$.

Automated Modality-Independent Determination of the Spine Curve and Axial Vertebral Rotation in 3D Images

Štern et al. [78] proposed a method for automated determination of the spine curve $\mathbf{c}(i)$ that is applicable to both CT and MR images of the spine, and can be therefore regarded as a modality-independent method. The method is based on the anatomical property that the walls of each vertebral body usually form a cylindrically shaped structure, which can be represented by a closed surface that contains the edges of the vertebral body in 3D, and on the geometrical property that any line orthogonal to vertebral body walls intersects with these edges in two points located on the opposite sides of vertebral body walls. The spine curve, defined as a curve in 3D that passes through the center of each vertebral body, is therefore located in the middle of any pair of opposite edge points on vertebral body walls.

The directions of lines that define the pairs of opposite edge points can be represented by image intensity gradient vectors in 3D, which are orthogonal to the

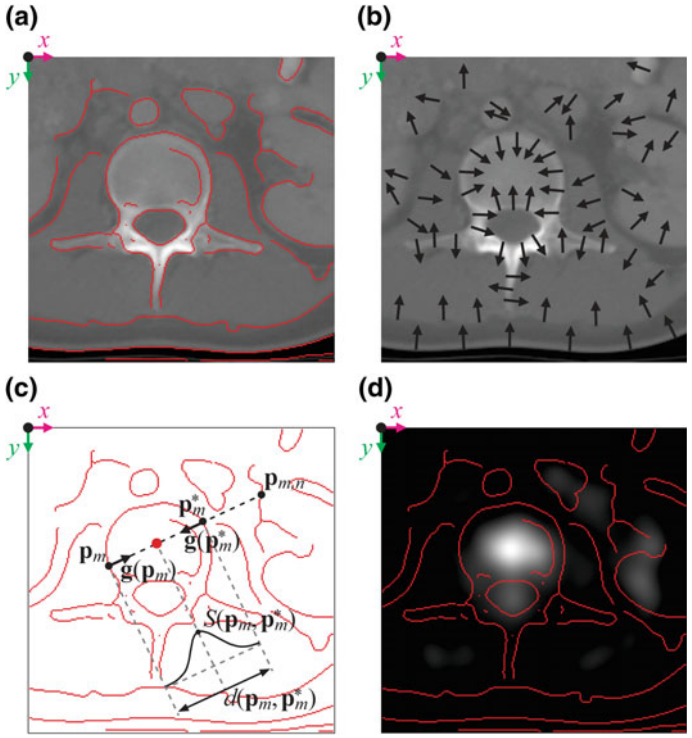


Fig. 12 An axial cross-section of a CT spine image I. **a** Edge points $\{\mathbf{p}_m\}$, extracted by the 3D Canny edge detector. **b** Image intensity gradient vectors $\{\mathbf{g}(\mathbf{p}_m)\}$, extracted by the 3D Sobel gradient operator (the number of gradient vectors was considerably reduced for visualization purposes). **c** An illustration of the determination of a pair of opposite edge points $(\mathbf{p}_m, \mathbf{p}^*_m)$ in the direction of gradient vector $+\mathbf{g}(\mathbf{p}_m)$. **d** The resulting accumulator A_I with superimposed edge points

extracted edges in the 3D image, while their magnitude is proportional to the strength of the extracted edges in the 3D image. For a given 3D image, the edge points $\{\mathbf{p}_m = (x_m, y_m, z_m); m = 1, 2, \ldots, M\}$ of anatomical structures (Fig. 12a) can be extracted by the 3D Canny edge detector (using e.g. high threshold t_{high} that captures 50 % of image pixels and low threshold $t_{\text{low}} = 0.4\,t_{\text{high}}$ for hysteresis thresholding), while the corresponding image intensity gradient vectors $\{\mathbf{g}(\mathbf{p}_m) = [g_x(x_m), g_y(y_m), g_z(z_m)]; m = 1, 2, \ldots, M\}$ (Fig. 12b) can be extracted by the 3D Sobel gradient operator (using e.g. kernel size of $3 \times 3 \times 3$ mm^3) and normalized so that $\forall \mathbf{g}(\mathbf{p}_m) \Rightarrow 0 \leq \|\mathbf{g}(\mathbf{p}_m)\| \leq 1$. Before computing the edges and gradient vectors, the images can be smoothed with a 3D Gaussian filter (with e.g. standard deviation of $\sigma = 1.5$ mm).

For each edge point $\mathbf{p}_m$, a search for the opposite edge point $\mathbf{p}^*_m$ is therefore performed in the direction of the normalized gradient vector $\mathbf{g}(\mathbf{p}_m)$, however,

multiple candidate points $\{\mathbf{p}_{m,n}; n = 1, 2, \ldots, N\}$ for the opposite edge point $\mathbf{p}_m^*$ exist along the search direction (Fig. 12c). By computing the absolute dot vector product between the corresponding normalized gradient vectors $\mathbf{g}(\mathbf{p}_m)$ and $\mathbf{g}(\mathbf{p}_{m,n})$:

$$S(\mathbf{p}_m, \mathbf{p}_{m,n}) = \sum_{q,r=0}^{Q} \left| \mathbf{g}(\mathbf{p}_m)|_q . \mathbf{g}(\mathbf{p}_{m,n})|_r \right|, \tag{42}$$

the congruence $S(\mathbf{p}_m, \mathbf{p}_{m,n})$ between the edge point $\mathbf{p}_m$ and its candidate opposite point $\mathbf{p}_{m,n}$ is evaluated, i.e. the more parallel the normalized gradient vectors $\mathbf{g}(\mathbf{p}_m)$ and $\mathbf{g}(\mathbf{p}_{m,n})$, the larger the congruence between the corresponding points, as edge points on the opposite sides of the vertebral body walls have gradient vectors of similar magnitudes and approximately opposite directions. However, the determination of such edge points is obstructed by edges that do not represent vertebral body walls, and an opposite edge point may not always exist for each edge point. To increase the robustness of the computation, gradient vectors in neighbourhoods of edge points, represented with $Q + 1$ points on planes orthogonal to the gradient vectors, are taken into account. As a result, the summation in Eq. 42 is performed over every qth point in the neighborhood of edge point $\mathbf{p}_m$ and the corresponding rth mirror point in the neighborhood of edge point $\mathbf{p}_{m,n}$:

$$\begin{aligned} \mathbf{g}(\mathbf{p}_m)|_q . \mathbf{g}(\mathbf{p}_{m,n})|_r = \; & g_x(x_m + q) g_x(x_{m,n} + r) \\ & + g_y(y_m + q) g_y(y_{m,n} + r) \\ & + g_y(y_m + q) g_y(y_{m,n} + r). \end{aligned} \tag{43}$$

As the size of vertebral bodies varies due to the biological variability of human anatomy (e.g. age, gender, pathology), among all candidate opposite edge points $\{\mathbf{p}_{m,n}; n = 1, 2, \ldots, N\}$ for edge point $\mathbf{p}_m$, only points that are less than $d_{\text{vb,min}} = 7$ mm and up to $d_{\text{vb,max}} = 60$ mm distant from $\mathbf{p}_m$ are taken into account (values are determined according to the average size of the human vertebral body [47, 52, 53]):

$$\mathbf{p'}_{m,n} \mapsto \mathbf{p}_{m,n}; \quad \forall p_{m,n}: \; d_{vb,\text{min}} \leq d(\mathbf{p}_m, \mathbf{p}_{m,n}) \leq d_{vb,\text{max}}; \quad n = 1, 2, \ldots, N, \tag{44}$$

where $d(\mathbf{p}_m, \mathbf{p}_n)$ is the Euclidean distance between points $\mathbf{p}_m$ and $\mathbf{p}_{m,n}$. Among the resulting candidate points $\{\mathbf{p'}_{m,n}; n = 1, 2, \ldots, N'\}$; $N' \leq N$, the point that results in the largest congruence $S(\mathbf{p}_m, \mathbf{p'}_{m,n})$ is selected as the opposite edge point $\mathbf{p}_m^*$ to edge point $\mathbf{p}_m$:

$$\mathbf{p}_m^* = \arg\max_{\mathbf{p'}_{m,n}} (S(\mathbf{p}_m, \mathbf{p'}_{m,n})); \quad n = 1, 2, \ldots, N'. \tag{45}$$

The search for opposite edge points is performed in both the positive direction $+\mathbf{g}(\mathbf{p}_m)$ and the negative direction $-\mathbf{g}(\mathbf{p}_m)$ of the gradient vector, as it may point inwards or outwards vertebral body walls due to variations in the distribution of

image intensities in different image modalities (i.e. CT, T_1-weighted MR and T_2-weighted MR images). As a result, two opposite edge points $\mathbf{p}_m^{*+}$ and $\mathbf{p}_m^{*-}$ are respectively determined, and by repeating the procedure for all edge points $\{\mathbf{p}_m; m = 1, 2, \ldots, M\}$ extracted from the 3D spine image, two sets of pairs of opposite edge points $\{(\mathbf{p}_m, \mathbf{p}_m^{*+}); m = 1, 2, \ldots, M^+; M^+ \leq M\}$ and $\{(\mathbf{p}_m, \mathbf{p}_m^{*-}); m = 1, 2, \ldots, M^-; M^- \leq M\}$ are respectively formed. The final set of pairs of opposite edge points $\{(\mathbf{p}_m, \mathbf{p}_m^*) = (\mathbf{p}_m, \mathbf{p}_m^{*+}) \cup (\mathbf{p}_m, \mathbf{p}_m^{*-}); m = 1, 2, \ldots, M'; M' \leq 2M\}$ is obtained by joining the results of the search in the positive and negative gradient vector directions.

As the spine curve passes through the center of each vertebral body, it is located where the lines connecting pairs $\{(\mathbf{p}_m, \mathbf{p}_m^*); m = 1, 2, \ldots, M'\}$ of opposite edge points most often intersect. For this purpose, a 3D accumulator A_I, which is of the same size as the observed 3D spine image I, is generated in the image-based coordinate system and initialized with zero values, i.e. $A_I(\mathbf{p}) = 0; \forall \mathbf{p} = (x, y, z) \in I$. Each line connecting a pair $(\mathbf{p}_m, \mathbf{p}_m^*)$ of opposite edge points is assigned a weighting function, normally distributed according to the distance $d(\mathbf{p}_m, \mathbf{p}_m^*)$ and scaled according to the congruence $S(\mathbf{p}_m, \mathbf{p}_m^*)$ between opposite edge points $\mathbf{p}_m$ and $\mathbf{p}_m^*$. The 3D accumulator value $A_I(\mathbf{p}_l)$ at point $\mathbf{p}_l$ is then increased by the value of the weighting function at each point $\{\mathbf{p}_l = (x_l, y_l, z_l); l = 1, 2, \ldots, L\}$ along the connecting line:

$$A_I(\mathbf{p}_l) = A_I(\mathbf{p}_l) + S(\mathbf{p}_m, \mathbf{p}_m^*) \exp\left(-\frac{(d(\mathbf{p}_m, \mathbf{p}_l) - d(\mathbf{p}_m, \mathbf{p}_m^*)/2)^2}{2(d(\mathbf{p}_m, \mathbf{p}_m^*)/6)^2}\right), \tag{46}$$

where $d(\mathbf{p}_m, \mathbf{p}_l)$ is the Euclidean distance between edge point $\mathbf{p}_m$ and each point $\mathbf{p}_l$ on the line, and $d(\mathbf{p}_m, \mathbf{p}_m^*)/6$ represents the standard deviation of the normally distributed weighting values. By accumulating the lines connecting all pairs $\{(\mathbf{p}_m, \mathbf{p}_m^*); m = 1, 2, \ldots, M'\}$ of opposite edge points in the 3D spine image, the values in the 3D accumulator A_I increase most along the longitudinal axes of vertebral bodies, as vertebral body walls contribute to most pairs of opposite edge points. The resulting normalized accumulator values, $\forall \mathbf{p} = (x, y, z) \in I \Rightarrow 0 \leq |A_I(\mathbf{p})| \leq 1$, therefore represent the probability that the spine curve passes through the corresponding locations (Fig. 12d).

Although maximal accumulator values point to the location of the spine curve, they may not always represent its exact location, as the generation of the accumulator is obstructed by edges that do not represent vertebral body walls. To determine the exact location of the spine curve, the coordinates of $H = 5$ largest maxima are extracted from each axial cross-section of the 3D accumulator A_I and connected into line segments. The line segments that are shorter than one half of the average size of the human vertebral body $d_{\text{vb}} = 30$ mm [47, 52, 53] are discarded (i.e. segments shorter than $d_{\text{vb}}/2 = 15$ mm). The remaining line segments, i.e. the set of points $\{\mathbf{p}_j; j = 1, 2, \ldots, J\}$, represent the candidate locations for the spine curve. A robust estimation of the polynomial parameters $b_{\mathbf{c}}$ of the spine curve $\mathbf{c}(i)$ (Eq. 25) can be obtained by applying the random sample consensus (RANSAC)

technique [16]. By randomly selecting $K+1=4$ points $\{\mathbf{p}_k; k=1,2,\ldots,K+1\}$ from set $\{\mathbf{p}_j\}$, i.e. $\{\mathbf{p}_k\} \subseteq \{\mathbf{p}_j\}$, a curve in the form of a polynomial function $\mathbf{c}'(i)$ of degree $K=3$ can be generated, resulting in a combination of polynomial parameters $b_\mathbf{c}$. To evaluate the agreement of curve $\mathbf{c}'(i)$ against points in $\{\mathbf{p}_j; j=1,2,\ldots,J\}$, the criterion $C(b_\mathbf{c})$ is computed as the number of points in $\{\mathbf{p}_j\}$ that are less than $r=3$ mm distant from $\mathbf{c}'(i)$ in each axial cross-section:

$$C(b_\mathbf{c}) = \text{count}(d(\mathbf{c}'(i)|_{b_\mathbf{c}}, \mathbf{p}_j) < r); \quad j=1,2,\ldots,J, \tag{47}$$

where $d(\mathbf{c}'(i)|_{b_\mathbf{c}}, \mathbf{p}_j)$ denotes the Euclidean distance between curve $\mathbf{c}'(i)$ and point $\mathbf{p}_j$. Among 1000 generated polynomial functions $\mathbf{c}'(i)$ (the number of iterations can be even larger), the optimal polynomial parameters $b_\mathbf{c}^*$ of the spine curve $\mathbf{c}(i)$ correspond to the maximal criterion $C(b_\mathbf{c})$:

$$b_\mathbf{c}^* = \arg\max_{b_\mathbf{c}}(C(b_\mathbf{c})). \tag{48}$$

The obtained spine curve $\mathbf{c}(i)$ therefore represents the best fit to the maxima of the 3D accumulator, which are located in the middle of vertebral body walls (Fig. 13).

The method was evaluated [78] on 42 3D images of the lumbar spine (29 CT images and 13 MR images), and the reported mean difference between the obtained spine curves in 3D and manually defined ground truth points was 1.8 ± 1.1 mm (1.7 ± 1.0 mm for CT images and 2.3 ± 1.5 mm for MR images).

Once the spine curve $\mathbf{c}(i)$ is determined, the axial vertebral rotation $\varphi(i)$ can be obtained by extracting planes that are orthogonal to the spine curve. By intersecting each of these planes with a line passing through $\mathbf{c}(i)$ and inclined for an angle that corresponds to the polynomial function defined by parameters b_φ, the optimal polynomial parameters b_φ^* representing the axial vertebral rotation $\varphi(i)$ (Eq. 26) can be computed by maximizing the in-plane similarity of the resulting image parts, e.g. by finding the maximal correlation of image intensities between mirror image halves (Eqs. 30–33) or the maximal mutual information of image intensities between plane parts (Eqs. 34–36), or by using a different similarity measure.

3 Cross-Sectional Reformation of 3D Spine Images

Volumetric image visualization can be defined as the transformation of image information from a 3D image space onto a 2D display device. The most straightforward 2D visualization of 3D images is based on original cross-sections that display the primarily reconstructed images, composed of original pixels in image reconstruction planes. A 2D cross-section of a 3D image is defined as the

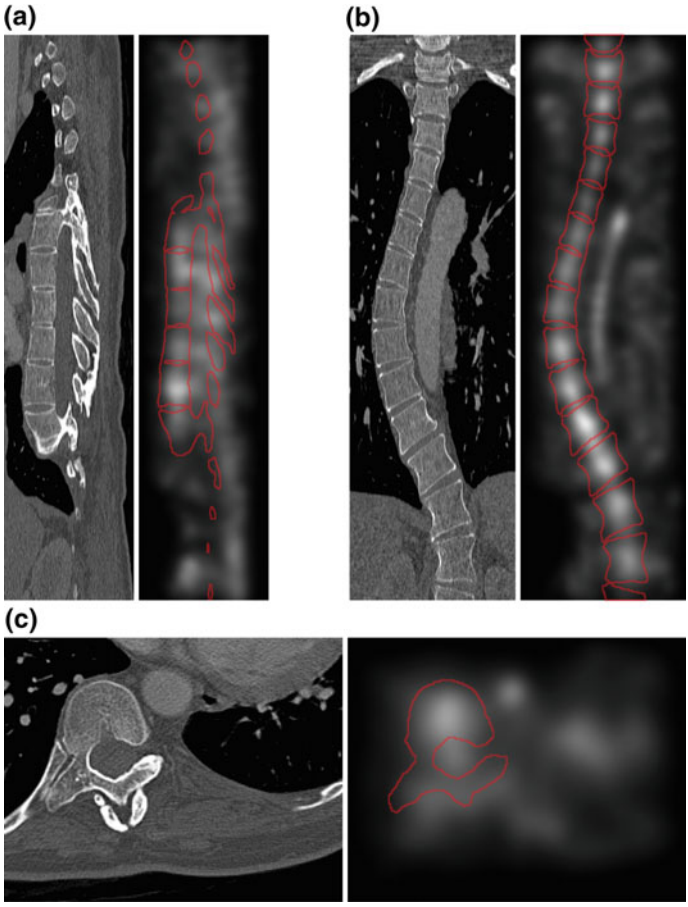


Fig. 13 Original image I (*left*) and the corresponding accumulator A_I with superimposed spine boundaries (*right*), displayed for a selected **a** sagittal, **b** coronal and **c** axial cross-section of a 3D CT image of a scoliotic spine. In the accumulator A_I, *brighter elements* correspond to a larger number of intersections of lines connecting opposite edge points

intersection of the 3D image with the reconstruction plane, on which image intensities are sampled. In CT imaging, reconstruction planes are usually transversely oriented, while in MR imaging they are usually oriented parallel to the excited slab. In both cases, the image reconstruction planes and the corresponding original cross-sections define the image-based coordinate system (Sect. 2.1.1).

The established techniques for 2D visualization of anatomical structures are therefore based on multi-planar reformation (MPR) that results in a series of sagittal, coronal and axial multi-planar cross-sections. However, multi-planar cross-sections do not always follow curved or tubular anatomical structures (e.g. spine, arteries, colon). As all of the important parts of the structure are not simultaneously visible in a single multi-planar cross-section, the visualization of such structures is

often unsatisfying, which may seriously affect the quality of the diagnostic information of the observed curved structures. When visualizing 3D spine images with multi-planar cross-sections, the spine may intersect with sagittal and coronal planes, while the axial plane may not always be located at the same level of vertebral bodies or intervertebral discs. The important structural parts of the spine may therefore not be displayed simultaneously in any single multi-planar cross-section, which may therefore not provide sufficient or qualitative enough diagnostic information, because they cannot follow the curvature of the spine and the rotation of vertebrae. This is already the case when visualizing a normal spine due to its natural "S"-shaped curvature, and is even more emphasized in pathological spinal curvatures, for example in the case of scoliosis or increased kyphosis/lordosis. Curved planar reformation (CPR) is a 2D image visualization technique that displays the originally reconstructed pixels along any user-defined curved surface that is flattened in order to appear as a plane. The use of the CPR visualization technique, which generates cross-sections that are orthogonal or tangent to the curve along the structure, represents a solution to the above mentioned problem. The standard coordinate system, which is determined by the 3D image, is transformed into a coordinate system that is determined by the observed 3D anatomical structure, for example, into the spine-based coordinate system in the case of 3D spine images (Sect. 2.1.2).

As a visualization technique, CPR is used in the field of angiography to display and evaluate blood vessels [25, 34, 35, 46, 51, 61, 63, 64, 72], in the field of pancreatography to display and evaluate pancreatic diseases [21, 60], for brain visualization [43], in the field of bronchoscopy [42, 57] and in the field of colonoscopy [19, 71, 92]. In all of the CPR visualization approaches, the determination of the curve that represents the central course of the visualized tubular structure is of utmost importance [3, 5, 9, 41, 97]. Dedicated commercial software or software provided by CT and MR scanner manufacturers already allows generation of curved cross-sections, however, this requires manual determination of the curve that follows the anatomical structure. Although MR scanners allow arbitrary orientation of the imaging plane and can therefore simulate the generation of oblique cross-sections, such visualization is greatly influenced by the scanner operator that determines the orientation of the imaging plane and by the position of the subject in the scanner. Curved cross-sections can be also acquired directly from the MR scanner [10, 11, 28], however, the quality of the obtained images in not adequate due to low spatial resolution of images, presence of intensity modulation artefacts and the fact that images can be curved only in one dimension. New concepts in curved-slice imaging allow to maintain a close-to-rectangular voxel size [93] and constant cross-sectional thickness [94]. On the other hand, by applying a combination of linear and/or non-linear spatial encoding magnetic fields for excitation and geometrically matched local encoding of curved-slice imaging, it is possible to achieve an almost rectangular voxel size [93] and constant cross-sectional thickness [94] of curved cross-sections.

Many approaches that aim to improve quantitative and qualitative evaluation of spinal deformities by an effective visualization of CT spine images have already

been proposed. By generating oblique sagittal images, Rabassa et al. [62] showed that visualization of vertebral facet joints improved, while oblique axial images allowed views that were parallel to intervertebral discs. Although the visualization was limited to oblique cross-sections, the authors concluded that in certain clinical situations, such as for the evaluation of neural foraminal stenosis or localization of spinal lesions, reformatted images could supplement the original 3D images. Oblique cross-sections that were orthogonal to the long axis of both left and right neural foraminae of the cervical spine region were also generated by Roberts et al. [65], who showed that by oblique MPR, consistency in the interpretation of neural foraminal stenosis between observers was improved, and suggested that such an approach should be considered in routine evaluation. Rothman et al. [68] demonstrated that curved cross-sections, obtained by connecting manually selected points into a continuous curve, were useful for the evaluation of anatomical relationships in the coronal spine region. After reformation, structures such as nerve roots, vertebral facet joints and spinal cord could be observed in a single 2D cross-section. Congenital spinal abnormalities were examined by Newton et al. [50], who manually outlined the boundaries of the spine in multi-planar cross-sections and created curved cross-sections that improved the identification and interpretation of abnormalities. The benefit of curved cross-sections was, in comparison with multi-planar or oblique cross-sections, most valuable in the case of significant sagittal or coronal spinal curvature, as they may help spine surgeons to achieve a more complete understanding and evaluation of spinal deformities. Menten et al. [48] presented a curved planospheric reformation method that was based on the reconstruction from a cylindrical plane, defined around the approximate boundary of the spinal canal within an axial cross-section. As a result, the anterior and posterior anatomical structures of the spine were displayed simultaneously in the same plane, which improved the evaluation of congenital spinal deformities. Manual determination of points or curves that determined the curved cross-sections was required in all of the above mentioned studies. A semi-automated method was presented by Kaminsky et al. [33], who segmented the spine on reformatted 3D images in order to overcome the problems of orientation in the standard multi-planar configuration. The transformation axis was determined by a 3D spline, obtained either manually by delineating centerlines in sagittal and coronal cross-sections, or automatically by dropping spheres of maximum possible radius through vertebral bodies or the spinal canal. Vrtovec et al. [83, 85] extracted curved cross-section from 3D spine images by representing the spine curve and the rotation of vertebrae as polynomial functions in 3D that formed the transformation axes for the reformation procedure, while Klinder et al. [36] reformatted 3D images by stacking curved cross-sections in order to reduce the region of interest and make the subsequent detection of vertebrae independent of the spinal curvature. Hanaoka et al. [22] extracted curved cross-sections by simultaneously aligning one elliptical column to the vertebral bodies and intervertebral discs, and a second elliptical column to the spinal canal, which allowed virtual straightening of the 3D image.

Image reformation was also identified as a valuable visualization technique in MR imaging of the spine. Apicella and Mirowitz [4] reported that multi-planar

cross-sections could compensate for the apparent asymmetry of 3D anatomical structures, caused by improper patient positioning or patient motion during image acquisition, and that reformatting can be applied to different anatomical structures. In the case of spine images, reformatted images can be used to improve the visualization of the spinal canal and intervertebral foraminae. In order to avoid measurement errors, Birchall et al. [8] and Adam and Askin [2] computed the rotation of vertebrae from the position of landmarks that were manually placed in each oblique axial cross-section, defined in sagittal and coronal MR cross-sections through the superior and inferior vertebral endplates, or parallel to the endplates through the centers of each vertebral body. Liljenqvist et al. [45] focused their study on vertebral morphology related to pedicle screw placement for the treatment of scoliosis. The pedicle width, length and angle were measured in manually determined oblique MR cross-sections that were orthogonal to vertebral bodies. In a study of automated survey of MR spine images [95], it was reported that automated reformation of 3D spine images along the true sagittal, coronal or axial vertebral body axes may potentially facilitate image interpretation. Vrtovec et al. [86] generated curved cross-section from MR spine images by extracting the 3D spine curve and axial vertebral rotation, and representing them as polynomial functions that guided the reformation procedure. The same method for the extraction of the spine curve and axial vertebral rotation was used by Neubert et al. [49] to initialize statistical shape models for the purpose of segmentation and analysis of high-resolution MR spine images.

Although 3D image reformation is often used for observing and analysing a variety of anatomical structures and related pathologies, it can be concluded that 3D spine images can be in general reformatted according to the following two principles:

- *multi-planar reformation (MPR)* is defined and performed in the image-based coordinate system (Sect. 3.1),
- *curved-planar reformation (CPR)* is defined in the spine-based coordinate system and performed in the image-based coordinate system (Sect. 3.2).

In both MPR and CPR, the plane of reformation is defined in the corresponding coordinate system, which is then represented as a plane or a curved surface in the image-based coordinate system, where image intensities are sampled. By flattening the extracted cross-sections onto a plane, visualization of the spine in both the image-based and the spine-based coordinate system is enabled (Sect. 3.3).

3.1 Multi-planar Reformation

Multi-planar reformation is the most straightforward 3D image reformation. The volume of the 3D image is cut by a plane, and image intensities are sampled on that plane. According to the orientation of the sampling plane in the image-based

coordinate system, the following types of MPR can be applied to 3D images of the spine:

- *orthogonal MPR*, where the sampling plane is orthogonal to one of the axes of the image-based coordinate system (Sect. 3.1.1),
- *oblique MPR*, where the orthogonal sampling plane, defined in the image-based coordinate system, is rotated about the axes of the image-based coordinate system (Sect. 3.1.2).

The common characteristic of all types of MPR is that sampling planes are defined on the basis of the image-based coordinate system $\mathbb{R}^3_I$.

3.1.1 Orthogonal Multi-planar Reformation

The most common MPR is orthogonal, meaning that the sampling plane is orthogonal to one of the axes of the image-based coordinate system. By applying orthogonal MPR to a 3D spine image, the following orthogonal multi-planar cross-sections can be obtained:

- *sagittal orthogonal multi-planar cross-sections* are obtained by sampling the 3D image on selected sagittal planes defined in the-based coordinate system (section Sagittal Orthogonal Multi-planar Cross-Sections),
- *coronal orthogonal multi-planar cross-sections* are obtained by sampling the 3D image on selected coronal planes defined in the image-based coordinate system (section Coronal Orthogonal Multi-planar Cross-Sections),
- *axial orthogonal multi-planar cross-sections* are obtained by sampling the 3D image on selected axial planes defined in the image-based coordinate system (section Axial Orthogonal Multi-planar Cross-Sections).

By selecting a point $\mathbf{p}_c = (x_c, y_c, z_c)$ in the image-based coordinate system $\mathbb{R}^3_I$, exactly one sagittal ($x = x_c$), one coronal ($y = y_c$) and one axial ($z = z_c$) orthogonal multi-planar cross-section can be defined through $\mathbf{p}_c$. If $\mathbf{p}_c$ is located on the spine curve $\mathbf{c}(i)$, e.g. at point $i = i_p$ so that $\mathbf{p}_c = \big(c_x(i_p), c_y(i_p), c_z(i_p)\big)$, the obtained cross-sections show, depending on the shape of the spine, parts of the spinal anatomy.

Sagittal Orthogonal Multi-planar Cross-Sections

Sagittal orthogonal multi-planar cross-sections are obtained by sampling the 3D image on selected sagittal planes that are orthogonal to axis x of the image-based coordinate system. The sagittal orthogonal multi-planar cross-section $M_{x=x_c}$ is therefore obtained by selecting a fixed coordinate $x = x_c$, and sampling the 3D image I along coordinates y and z in the image-based coordinate system:

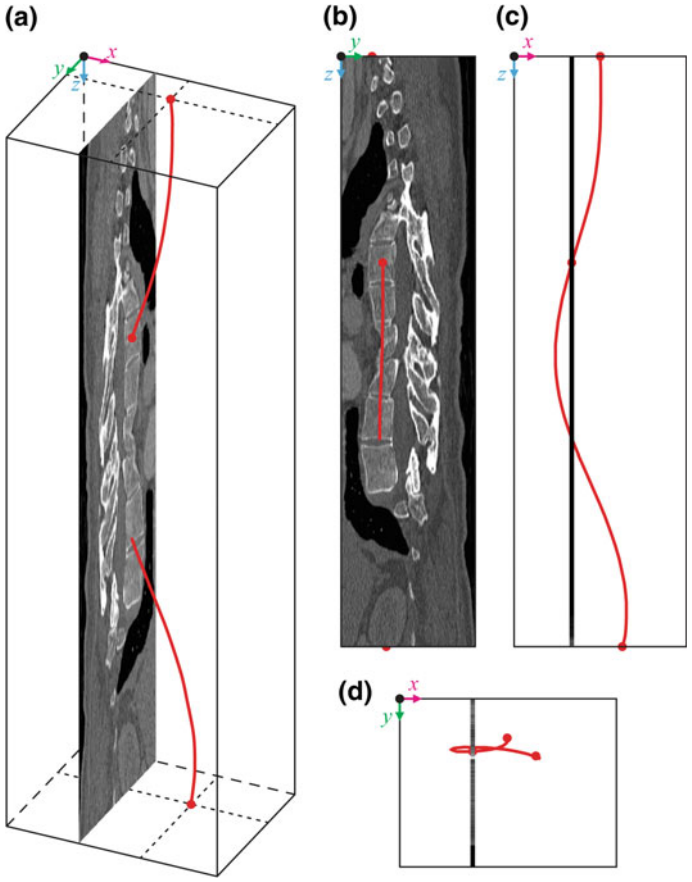


Fig. 14 A sagittal orthogonal multi-planar cross-section $M_{x=x_c}$ of a 3D CT image of a scoliotic spine, shown in **a** 3D view, **b** left sagittal view, **c** posterior coronal view and **d** superior axial view of the image-based coordinate system (*Note* The image-based coordinate system and the spine curve correspond to Figs. 1 and 7)

$$M_{x=x_c}(y,z) = I(x_c,y,z). \tag{49}$$

In the case of normal spines, the anatomy of all vertebrae can be usually observed simultaneously in selected sagittal orthogonal multi-planar cross-sections. However, in the case scoliotic spines, vertebrae come in and out of the sampling plane, and therefore the anatomy of all vertebrae cannot be simultaneously observed in any selected sagittal orthogonal multi-planar cross-section, as shown in Fig. 14.

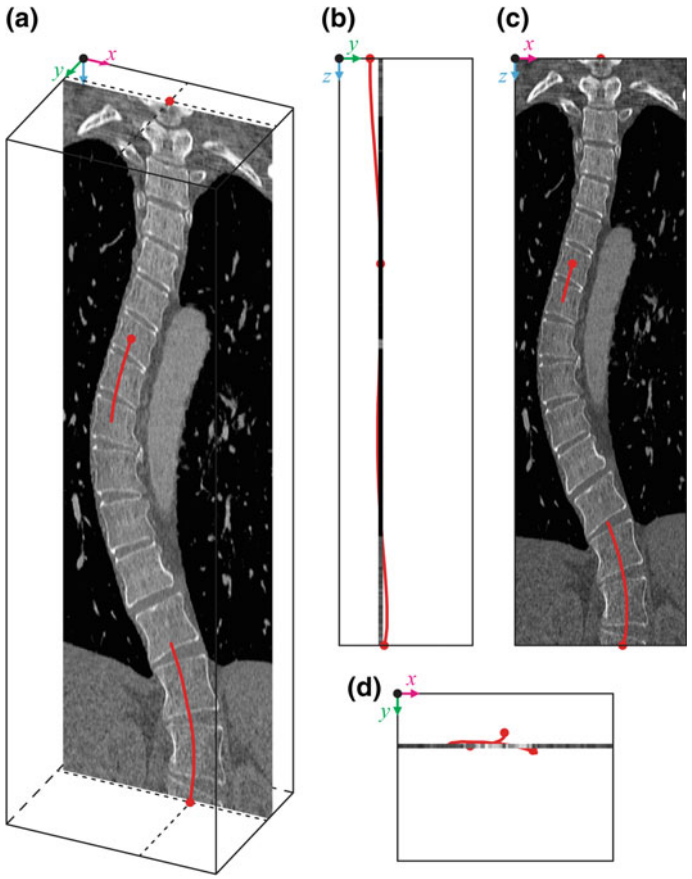


Fig. 15 A coronal orthogonal multi-planar cross-section $M_{y=y_c}$ of a 3D CT image of a scoliotic spine, shown in **a** 3D view, **b** left sagittal view, **c** posterior coronal view and **d** superior axial view of the image-based coordinate system (*Note* The image-based coordinate system and the spine curve correspond to Figs. 1 and 7)

Coronal Orthogonal Multi-planar Cross-Sections

Coronal orthogonal multi-planar cross-sections are obtained by sampling the 3D image on selected coronal planes that are orthogonal to axis y of the image-based coordinate system. The coronal orthogonal multi-planar cross-section $M_{y=y_c}$ is therefore obtained by selecting a fixed coordinate $y = y_c$, and sampling the 3D image I along coordinates x and z in the image-based coordinate system:

$$M_{y=y_c}(x, z) = I(x, y_c, z). \tag{50}$$

In the case of normal spines, the anatomy of all vertebrae cannot be simultaneously observed in any selected coronal orthogonal multi-planar cross-section, as vertebrae come in and out of the sampling plane. On the other hand, in the case of scoliotic spines, the anatomy of all vertebrae can be usually observed simultaneously in selected coronal orthogonal multi-planar cross-sections, as shown in Fig. 15.

Axial Orthogonal Multi-planar Cross-Sections

Axial orthogonal multi-planar cross-sections are obtained by sampling the 3D image on selected axial planes that are orthogonal to axis z of the image-based

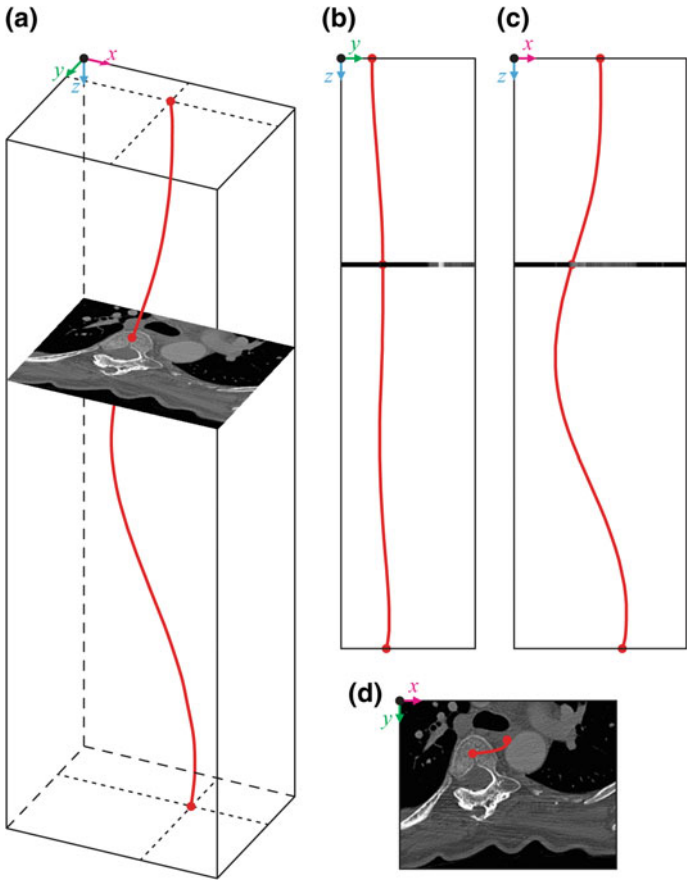


Fig. 16 An axial orthogonal multi-planar cross-section $M_{z=z_c}$ of a 3D CT image of a scoliotic spine, shown in **a** 3D view, **b** left sagittal view, **c** posterior coronal view and **d** superior axial view of the image-based coordinate system (*Note* The image-based coordinate system and the spine curve correspond to Figs. 1 and 7)

coordinate system. The axial orthogonal multi-planar cross-section $M_{z=z_c}$ is therefore obtained by selecting a fixed coordinate $z = z_c$, and sampling the 3D image I along coordinates x and y in the image-based coordinate system:

$$M_{z=z_c}(x, y) = I(x, y, z_c). \tag{51}$$

In the case of normal spines, vertebrae are usually sagittally inclined, while in the case of scoliotic spines, vertebrae are usually coronally inclined against axis z. As a result, axial orthogonal multi-planar cross-sections (Fig. 16) in general do not show a completely geometrically correct shape of the vertebral anatomy, because sampling planes cut through vertebrae at different anatomical locations. For example, similarly as an ellipse can be obtained by intersecting a circular cone with an inclined plane, the shape of the vertebral body is observed as a more elliptical structure than it may actually be.

3.1.2 Oblique Multi-planar Reformation

An established type of MPR is oblique (slanted), meaning that the orthogonal sampling plane is rotated (inclined) for selected angles about the axes of the image-based coordinate system. By applying oblique MPR to a 3D spine image, the following oblique multi-planar cross-sections can be obtained:

- *sagittal oblique multi-planar cross-sections* are obtained by sampling the 3D image on selected rotated sagittal orthogonal planes, defined in the image-based coordinate system (section Sagittal Oblique Multi-planar Cross-Sections),
- *coronal oblique multi-planar cross-sections* are obtained by sampling the 3D image on selected rotated coronal orthogonal planes, defined in the image-based coordinate system (section Coronal Oblique Multi-planar Cross-Sections),
- *axial oblique multi-planar cross-sections* are obtained by sampling the 3D image on selected rotated axial orthogonal planes, defined in the image-based coordinate system (section Axial Oblique Multi-planar Cross-Sections),
- *generalized oblique multi-planar cross-sections* are obtained by sampling the 3D image on planes that are arbitrarily defined in the image-based coordinate system (section Generalized Oblique Multi-planar Cross-Sections).

The rotation for angles α, β and γ about axes x, y and z, respectively, of the image-based coordinate system can be represented by rotation matrices $R_x(\alpha)$, $R_y(\beta)$ and $R_z(\gamma)$, respectively:

$$R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix}, \tag{52}$$

$$R_y(\beta) = \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix}, \tag{53}$$

$$R_z(\gamma) = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{54}$$

and the composition of extrinsic rotations about axes x (first), y (second) and z (last) can be represented by the rotation matrix $R(\alpha, \beta, \gamma)$:

$$R(\alpha, \beta, \gamma) = R_z(\gamma) R_y(\beta) R_x(\alpha). \tag{55}$$

For an arbitrary point $\mathbf{p} = (x, y, z)$ in the image-based coordinate system $\mathbb{R}^3_I$, its location $\mathbf{p}' = (x', y', z')$ after rotation is obtained by:

$$\begin{aligned} \mathbf{p}' = (x', y', z') &= R(\alpha, \beta, \gamma)[x, y, z] \\ &= R(\alpha, \beta, \gamma)\mathbf{p}, \end{aligned} \tag{56}$$

with the center of rotation at the origin $\mathbf{p}_0 = (0, 0, 0)$ of the image-based coordinate system.[2] If the center of rotation is at point $\mathbf{p}_c = (x_c, y_c, z_c)$, then the location of $\mathbf{p}$ after rotation is:

$$\begin{aligned} \mathbf{p}' = (x', y', z') &= (R(\alpha, \beta, \gamma)[x - x_c, y - y_c, z - z_c]) + [x_c, y_c, z_c] \\ &= (R(\alpha, \beta, \gamma)(\mathbf{p} - \mathbf{p}_c)) + \mathbf{p}_c, \end{aligned} \tag{57}$$

By selecting a point $\mathbf{p}_c = (x_c, y_c, z_c)$ in the image-based coordinate system $\mathbb{R}^3_I$ and rotation angles $\alpha = \alpha_p$, $\beta = \beta_p$, and $\gamma = \gamma_p$, three different sagittal $\big(x = x_c$ and $\beta = \beta_p,\ \gamma = \gamma_p$ or both$\big)$, coronal $\big(y = y_c$ and $\alpha = \alpha_p,\ \gamma = \gamma_p$ or both$\big)$ and axial ($z = z_c$ and $\alpha = \alpha_p$, $\beta = \beta_p$ or both) oblique multi-planar cross-sections can be defined through $\mathbf{p}_c$, which also represents the center of rotation. If $\mathbf{p}_c$ is located on the spine curve $\mathbf{c}(i)$, e.g. at point $i = i_p$ so that $\mathbf{p}_c = \big(c_x(i_p), c_y(i_p), c_z(i_p)\big)$, the obtained cross-sections show, depending on the shape of the spine, parts of the spinal anatomy.

Sagittal Oblique Multi-planar Cross-Sections

The sagittal oblique multi-planar cross-section $M_{x=x_c, \beta=\beta_p, \gamma=\gamma_p}$ is obtained by sampling the 3D image I on the sagittal orthogonal plane at the selected fixed

[2] It is assumed that $\mathbf{p}$ is a column vector. If $\mathbf{p}$ is a row vector, vector transpose operation is required, therefore Eq. 56 turns into $\mathbf{p}' = (x', y', z') = \big(R(\alpha, \beta, \gamma)[x, y, z]^T\big)^T = (R(\alpha, \beta, \gamma)\mathbf{p}^T)^T$.

coordinate $x = x_c$ that is additionally rotated for angle $\beta = \beta_p$ about axis y and/or for angle $\gamma = \gamma_p$ about axis z of the image-based coordinate system:

$$M_{x=x_c,\beta=\beta_p,\gamma=\gamma_p}(y,z) = I(R(0,\beta_p,\gamma_p)[x_c,y,z]), \tag{58}$$

with the center of rotation at point $\mathbf{p}_c = (x_c, y_c, z_c)$ (Eq. 57). The rotation angles must be on the closed interval $\{\beta_p, \gamma_p\} \in [-\frac{\pi}{4}, +\frac{\pi}{4}]$, otherwise the cross-section turns into an axial or coronal oblique multi-planar cross-section. However, in practice only one rotation is usually applied:

$$M_{x=x_c,\beta=\beta_p}(y,z) = I(R(0,\beta_p,0)[x_c,y,z]) = I(R_y(\beta_p)[x_c,y,z]), \tag{59}$$

$$M_{x=x_c,\gamma=\gamma_p}(y,z) = I(R(0,0,\gamma_p)[x_c,y,z]) = I(R_z(\gamma_p)[x_c,y,z]). \tag{60}$$

Figure 17 displays the sagittal oblique multi-planar cross-section $M_{x=x_c,\gamma=\gamma_p}$ (Eq. 60) at $x = x_c$ and $\gamma = \gamma_p = 25°$.

Coronal Oblique Multi-planar Cross-Sections

The coronal oblique multi-planar cross-section $M_{y=y_c,\alpha=\alpha_p,\gamma=\gamma_p}$ is obtained by sampling the 3D image I on the coronal orthogonal plane at the selected fixed coordinate $y = y_c$ that is additionally rotated for angle $\alpha = \alpha_p$ about axis x and/or for angle $\gamma = \gamma_p$ about axis z of the image-based coordinate system:

$$M_{y=y_c,\alpha=\alpha_p,\gamma=\gamma_p}(x,z) = I(R(\alpha_p,0,\gamma_p)[x,y_c,z]), \tag{61}$$

with the center of rotation at point $\mathbf{p}_c = (x_c, y_c, z_c)$ (Eq. 57). The rotation angles must be on the closed interval $\{\alpha_p, \gamma_p\} \in [-\frac{\pi}{4}, +\frac{\pi}{4}]$, otherwise the cross-section turns into an axial or sagittal oblique multi-planar cross-section. However, in practice only one rotation is usually applied:

$$M_{y=y_c,\alpha=\alpha_p}(x,z) = I(R(\alpha_p,0,0)[x,y_c,z]) = I(R_x(\alpha_p)[x,y_c,z]), \tag{62}$$

$$M_{y=y_c,\gamma=\gamma_p}(x,z) = I(R(0,0,\gamma_p)[x,y_c,z]) = I(R_z(\gamma_p)[x,y_c,z]). \tag{63}$$

Figure 18 displays the coronal oblique multi-planar cross-section $M_{y=y_c,\gamma=\gamma_p}$ (Eq. 63) at $y = y_c$ and $\gamma = \gamma_p = 25°$.

Axial Oblique Multi-planar Cross-Sections

The axial oblique multi-planar cross-section $M_{z=z_c,\alpha=\alpha_p,\beta=\beta_p}(x,y)$ is obtained by sampling the 3D image I on the axial orthogonal plane at the selected fixed

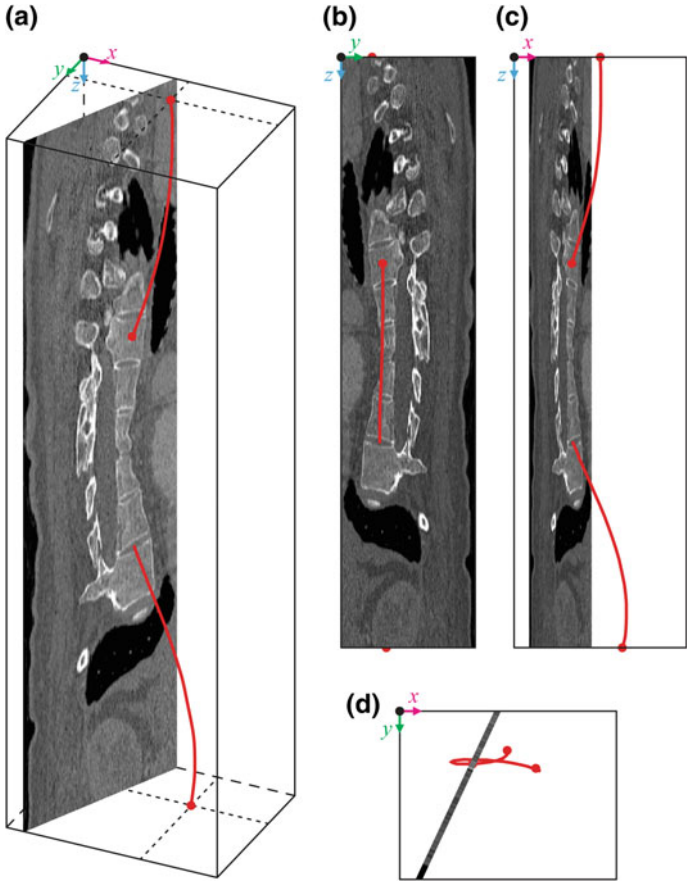


Fig. 17 A sagittal oblique multi-planar cross-section $M_{x=x_c,y=y_p}$ of a 3D CT image of a scoliotic spine, shown in **a** 3D view, **b** left sagittal view, **c** posterior coronal view and **d** superior axial view of the image-based coordinate system (*Note* The image-based coordinate system and the spine curve correspond to Figs. 1 and 7)

coordinate $z = z_c$ that is additionally rotated for angle $\alpha = \alpha_p$ about axis x and/or for angle $\beta = \beta_p$ about axis y of the image-based coordinate system:

$$M_{z=z_c,\alpha=\alpha_p,\beta=\beta_p}(x,y) = I(R(\alpha_p,\beta_p,0)[x,y,z_c]), \tag{64}$$

with the center of rotation at point $\mathbf{p}_c = (x_c, y_c, z_c)$ (Eq. 57). The rotation angles must be on the closed interval $\{\alpha_p, \beta_p\} \in [-\frac{\pi}{4}, +\frac{\pi}{4}]$, otherwise the cross-section turns into a coronal or sagittal oblique multi-planar cross-section. However, in practice only one rotation is usually applied:

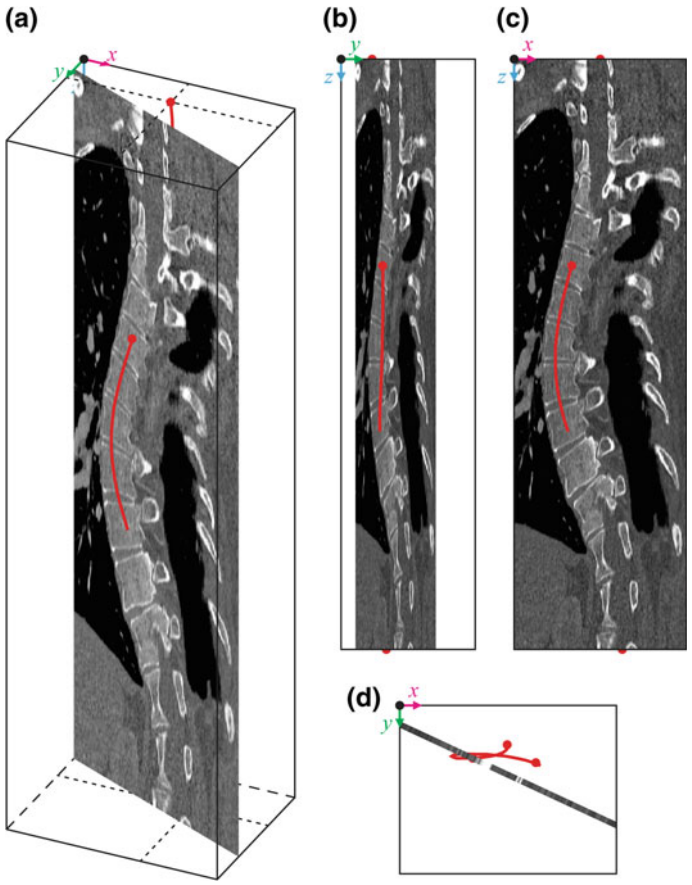


Fig. 18 A coronal oblique multi-planar cross-section $M_{y=y_c,\gamma=\gamma_p}$ of a 3D CT image of a scoliotic spine, shown in **a** 3D view, **b** left sagittal view, **c** posterior coronal view and **d** superior axial view of the image-based coordinate system (*Note* The image-based coordinate system and the spine curve correspond to Figs. 1 and 7)

$$M_{z=z_c,\alpha=\alpha_p}(x,y) = I(R(\alpha_p,0,0)[x,y,z_c]) = I(R_x(\alpha_p)[x,y,z_c]), \tag{65}$$

$$M_{z=z_c,\beta=\beta_p}(x,y) = I(R(0,\beta_p,0)[x,y,z_c]) = I(R_y(\beta_p)[x,y,z_c]). \tag{66}$$

Figure 19 displays the axial oblique multi-planar cross-section $M_{z=z_c,\alpha=\alpha_p}$ (Eq. 65) at $z = z_c$ and $\alpha = \alpha_p = 25°$.

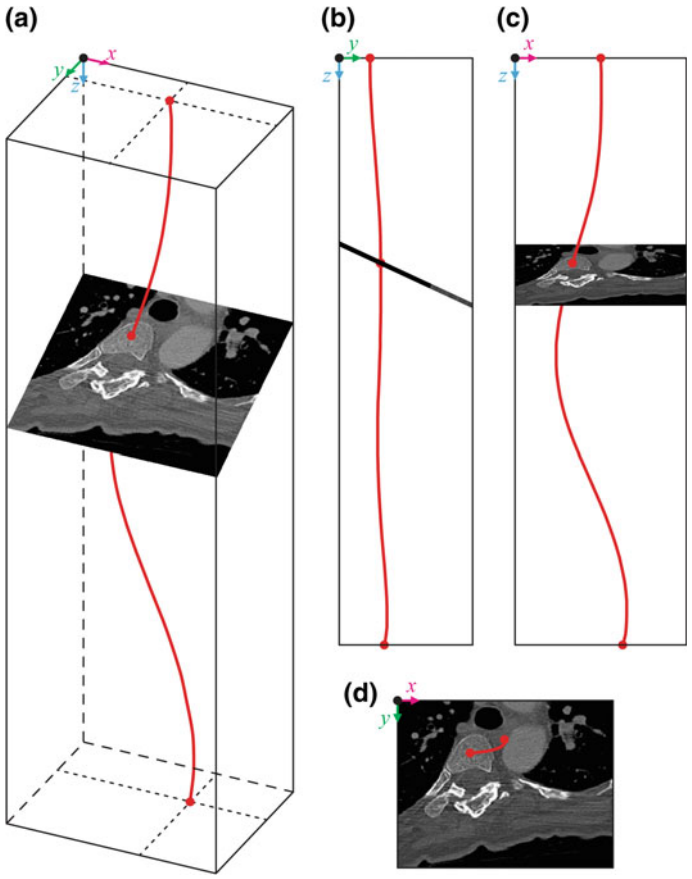


Fig. 19 An axial oblique multi-planar cross-section $M_{z=z_c,\alpha=\alpha_p}$ of a 3D CT image of a scoliotic spine, shown in **a** 3D view, **b** left sagittal view, **c** posterior coronal view and **d** superior axial view of the image-based coordinate system (*Note* The image-based coordinate system and the spine curve correspond to Figs. 1 and 7)

Generalized Oblique Multi-planar Cross-Sections

Generalized oblique multi-planar cross-sections are obtained by sampling the 3D image I on a sagittal ($x = x_c$), coronal ($y = y_c$) or axial ($z = z_c$) orthogonal plane that is additionally rotated for angles $\alpha = \alpha_p$, $\beta = \beta_p$ and/or $\gamma = \gamma_p$ about axes x, y and z, respectively, of the image-based coordinate system:

$$M_{x=x_c,\alpha=\alpha_p,\beta=\beta_p,\gamma=\gamma_p}(y,z) = I(R(\alpha_p,\beta_p,\gamma_p)[x_c,y,z]), \tag{67}$$

$$M_{y=y_c,\alpha=\alpha_p,\beta=\beta_p,\gamma=\gamma_p}(x,z) = I(R(\alpha_p,\beta_p,\gamma_p)[x,y_c,z]), \tag{68}$$

$$M_{z=z_c,\alpha=\alpha_p,\beta=\beta_p,\gamma=\gamma_p}(x,y) = I(R(\alpha_p,\beta_p,\gamma_p)[x,y,z_c]), \tag{69}$$

with the center of rotation at point $\mathbf{p}_c = (x_c, y_c, z_c)$ (Eq. 57). However, in the case of 3D spine images, the value of such generalized oblique multi-planar cross-sections is questionable since the arbitrarily defined sampling plane may not include any spinal anatomy.

A more valuable result can be achieved by defining the sampling plane according to the spine as the observed anatomical structure. By selecting three non-collinear points $\mathbf{p}_1 = (x_1, y_1, z_1)$, $\mathbf{p}_2 = (x_2, y_2, z_2)$ and $\mathbf{p}_3 = (x_3, y_3, z_3)$ on the spine, which may be located on the spine curve $\mathbf{c}(i)$, e.g. at points $i = i_1$, $i = i_2$ and $i = i_3$ so that $\mathbf{p}_1 = (c_x(i_1), c_y(i_1), c_z(i_1))$, $\mathbf{p}_2 = (c_x(i_2), c_y(i_2), c_z(i_2))$ and $\mathbf{p}_3 = (c_x(i_3), c_y(i_3), c_z(i_3))$, respectively, a sampling plane P can be uniquely defined in the image-based coordinate system and used to redefine the rotation matrix R (Eq. 55). The unit normal vector $\hat{\mathbf{n}}_P$ of plane P is:

$$\hat{\mathbf{n}}_P = [\hat{n}_{Px}, \hat{n}_{Py}, \hat{n}_{Pz}] = \frac{\mathbf{n}_P}{\|\mathbf{n}_P\|}; \quad \mathbf{n}_P = (\mathbf{p}_1 - \mathbf{p}_3) \times (\mathbf{p}_2 - \mathbf{p}_3). \tag{70}$$

To redefine the rotation matrix R, two unit vectors $\hat{\mathbf{e}}_1$ and $\hat{\mathbf{e}}_2$ have to be additionally defined that are, including $\hat{\mathbf{n}}_P$, mutually orthogonal. We have one degree of freedom for the selection of $\hat{\mathbf{e}}_1$, e.g.:

$$\begin{aligned} &\hat{\mathbf{n}}_P \cdot \hat{\mathbf{e}}_1 = 0 \\ &\hat{\mathbf{e}}_1 = [\hat{e}_{1x}, \hat{e}_{1y}, \hat{e}_{1z}] = \frac{\mathbf{e}_1}{\|\mathbf{e}_1\|}; \quad \mathbf{e}_1 = [-\hat{n}_{Pz}, -\hat{n}_{Pz}, \hat{n}_{Px} + \hat{n}_{Py}], \end{aligned} \tag{71}$$

which is then used to determine $\hat{\mathbf{e}}_2$:

$$\hat{\mathbf{e}}_2 = [\hat{e}_{2x}, \hat{e}_{2y}, \hat{e}_{2z}] = \hat{\mathbf{e}}_1 \times \hat{\mathbf{n}}_P. \tag{72}$$

In 3D spine images, normal spines are usually aligned with sagittal orthogonal planes, while scoliotic spines are usually aligned with coronal orthogonal planes. As a result, $\hat{\mathbf{n}}_P$ (Eq. 70) represents, in the coordinate system of plane P, the unit vector $\hat{\mathbf{e}}_{Px} = [1,0,0]_P$ in the case of normal spines, and the unit vector $\hat{\mathbf{e}}_{Py} = [0,1,0]_P$ in the case of scoliotic spines. On the other hand, $\hat{\mathbf{e}}_1$ (Eq. 71) is selected so that it always represents the unit vector $\hat{\mathbf{e}}_{Pz} = [0,0,1]_P$, while $\hat{\mathbf{e}}_2$ (Eq. 72) represents the remaining unit vector in the coordinate system of plane P. As a result, the rotation matrix is in the case of normal spines redefined as R^n and used to obtain the generalized oblique multi-planar cross-section $M^n_{\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3}$:

$$M^n_{\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3}(y,z) = I(R^n[x_j,y,z]); \quad R^n = \begin{bmatrix} \hat{n}_{Px} & \hat{e}_{2x} & \hat{e}_{1x} \\ \hat{n}_{Py} & \hat{e}_{2y} & \hat{e}_{1y} \\ \hat{n}_{Pz} & \hat{e}_{2z} & \hat{e}_{1z} \end{bmatrix}, \tag{73}$$

with the center of rotation at point $\mathbf{p}_j$ (Eq. 57), arbitrarily chosen among points $\mathbf{p}_1$, $\mathbf{p}_2$ and $\mathbf{p}_3$. On the other hand, in the case of scoliotic spines, the rotation matrix is redefined as R^s and used to obtain the generalized oblique multi-planar cross-section $M^s_{\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3}$:

$$M^s_{\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3}(x,z) = I(R^s[x,y_j,z]); \quad R^s = \begin{bmatrix} \hat{e}_{2x} & \hat{n}_{Px} & \hat{e}_{1x} \\ \hat{e}_{2y} & \hat{n}_{Py} & \hat{e}_{1y} \\ \hat{e}_{2z} & \hat{n}_{Pz} & \hat{e}_{1z} \end{bmatrix}, \tag{74}$$

with the center of rotation again at point $\mathbf{p}_j$ (Eq. 57), arbitrarily chosen among points $\mathbf{p}_1$, $\mathbf{p}_2$ and $\mathbf{p}_3$. Figure 20 displays the generalized oblique multi-planar cross-section

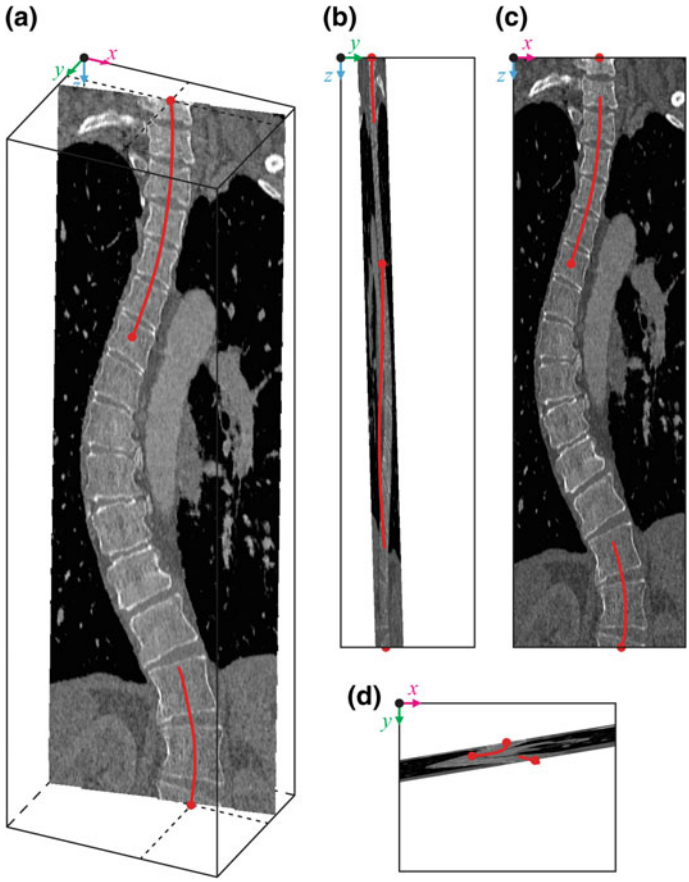


Fig. 20 A generalized oblique multi-planar cross-section $M^s_{\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3}$ of a 3D CT image of a scoliotic spine, shown in **a** 3D view, **b** left sagittal view, **c** posterior coronal view and **d** superior axial view of the image-based coordinate system (*Note* The image-based coordinate system and the spine curve correspond to Figs. 1 and 7)

$M^s_{\mathbf{p}_1,\mathbf{p}_2,\mathbf{p}_3}$ (Eq. 74) with $\mathbf{p}_1$ as the spine curve start point, $\mathbf{p}_2$ as the spine curve end point, and $\mathbf{p}_3 = \mathbf{p}_c = \mathbf{c}(i_p)$ as a point on the spine curve at $i = i_p$.

3.2 Curved-Planar Reformation

Curved-planar reformation is an efficient technique for cross-sectional visualization of curved 3D anatomical structures, where the goal is to visualize the structure along its entire length within a single cross-section. The volume of the 3D structure is cut by a plane, and image intensities are sampled on that plane. According to the orientation of the sampling plane in the spine-based coordinate system, the following types of CPR can be applied to 3D images of the spine:

- *orthogonal CPR*, where the sampling plane is orthogonal to one of the axes of the spine-based coordinate system (Sect. 3.2.1),
- *oblique CPR*, where the orthogonal sampling plane, defined in the spine-based coordinate system, is rotated about the axes of the spine-based coordinate system (Sect. 3.2.2).

The common characteristic of all types of CPR is that sampling planes are defined on the basis of the spine-based coordinate system $\mathbb{R}^3_S$. However, image intensities can be accessed only in the image-based coordinate system, therefore a transformation from the spine-based to the image-based coordinate system is required (Sect. 2.2.3) and achieved through a continuous representation of the spine curve $\mathbf{c}(i)$ (Sect. 2.3.1) and axial vertebral rotation $\varphi(i)$ (Sect. 2.3.2). As the axial vertebral rotation $\varphi(i)$ represents axes u and v in the spine-based coordinate system, it must be defined in planes orthogonal to axis w. The axial vertebral rotation can be therefore represented by matrix $R_w(\varphi(i))$ of rotation about axis w of the spine-based coordinate system, which has the same form as R_z (Eq. 54), i.e. $R_w(\varphi(i)) = R_z(\varphi(i))$. However, the rotation in the image-based coordinate system has to be performed about the axis defined by the unit tangent vector $\hat{\mathbf{t}}(i) = (\hat{t}_x(i), \hat{t}_y(i), \hat{t}_z(i))$ to the spine curve $\mathbf{c}(i)$, which represents axis w in the spine-based coordinate system. The rotation in the form of axis-angle representation can be achieved by matrix $R_{\mathbf{t}(i)}(\varphi(i))$ (Eq. 19).

3.2.1 Orthogonal Curved-Planar Reformation

The most straightforward approach to CPR is orthogonal, meaning that the sampling plane is orthogonal to one of the axes of the spine-based coordinate system. By applying orthogonal CPR to a 3D spine image, the following orthogonal curved-planar cross-sections can be obtained:

- *sagittal orthogonal curved-planar cross-sections* are obtained by sampling the 3D image on selected sagittal planes, defined in the spine-based coordinate system (section Sagittal Orthogonal Curved-Planar Cross-Sections),
- *coronal orthogonal curved-planar cross-sections* are obtained by sampling the 3D image on selected coronal planes, defined in the spine-based coordinate system (section Coronal Orthogonal Curved-Planar Cross-Sections),
- *axial orthogonal curved-planar cross-sections* are obtained by sampling the 3D image on selected axial planes, defined in the spine-based coordinate system (section Axial Orthogonal Curved-Planar Cross-Sections).

By selecting a point $\mathbf{p}_c = (u_c, v_c, w_c)$ in the spine-based coordinate system $\mathbb{R}^3_S$, exactly one sagittal ($u = u_c$), one coronal ($v = v_c$) and one axial ($w = w_c$) orthogonal curved-planar cross-section can be defined through $\mathbf{p}_c$. If $\mathbf{p}_c$ is located on the spine curve $\mathbf{c}(i)$, e.g. at point $i = i_p$ so that $\mathbf{p}_c = (c_x(i_p), c_y(i_p), c_z(i_p))$, the obtained sagittal and coronal cross-sections show, irrespectively of the shape of the spine, the spinal anatomy at its midline along its entire length, while the obtained axial cross-section shows the spinal anatomy in the direction orthogonal to its midline.

Sagittal Orthogonal Curved-Planar Cross-Sections

Sagittal orthogonal curved-planar cross-sections are obtained by sampling the 3D image on selected sagittal planes that are orthogonal to axis u of the spine-based coordinate system. The sagittal orthogonal curved-planar cross-section $C_{u=u_c}$ is therefore obtained by selecting a fixed coordinate $u = u_c$, and sampling the 3D image I along coordinates v and w in the spine-based coordinate system:

$$C_{u=u_c}(v, w) = I(u_c, v, w). \tag{75}$$

In the image-based coordinate system, the sampling plane is represented by a curved surface that is parallel to the spine curve $\mathbf{c}(i) = (c_x(i), c_y(i), c_z(i))$ and follows the axial vertebral rotation $\varphi(i)$. If the selected fixed coordinate u_c is represented as $u_c \leftrightarrow c_x(i) + \Delta x$, where Δx is a fixed offset in the left or right direction from the spine curve $\mathbf{c}(i)$ that corresponds to the sagittal offset of point $\mathbf{p}_c = (u_c, v_c, w_c)$ from the origin of the spine-based coordinate system, then the sagittal orthogonal curved-planar cross-section $C_{u=u_c}$ can be obtained as:

$$C_{u=u_c}(y, c_z(i)) = I(R_{\hat{\mathbf{t}}(i)}(\varphi(i))R_x(\alpha(i))[c_x(i) + \Delta x, y, c_z(i)]), \tag{76}$$

where matrix $R_{\hat{\mathbf{t}}(i)}(\varphi(i))$ (Eq. 19) represents the axial vertebral rotation for angle $\varphi(i)$ about axis defined by $\hat{\mathbf{t}}(i)$ (i.e. $\varphi(i) = \varphi_w(i)$, Eq. 14), and matrix $R_x(\alpha(i))$ (Eq. 52) represents the rotation for angle $\alpha(i) = \arctan(\hat{t}_y(i)/\hat{t}_z(i))$ about axis x of the image-based coordinate system, considering that $\hat{\mathbf{t}}(i) = [\hat{t}_x(i), \hat{t}_y(i), \hat{t}_z(i)]$ is the unit tangent vector to the spine curve, and $\mathbf{c}(i) = (c_x(i), c_y(i), c_z(i))$ is the center of

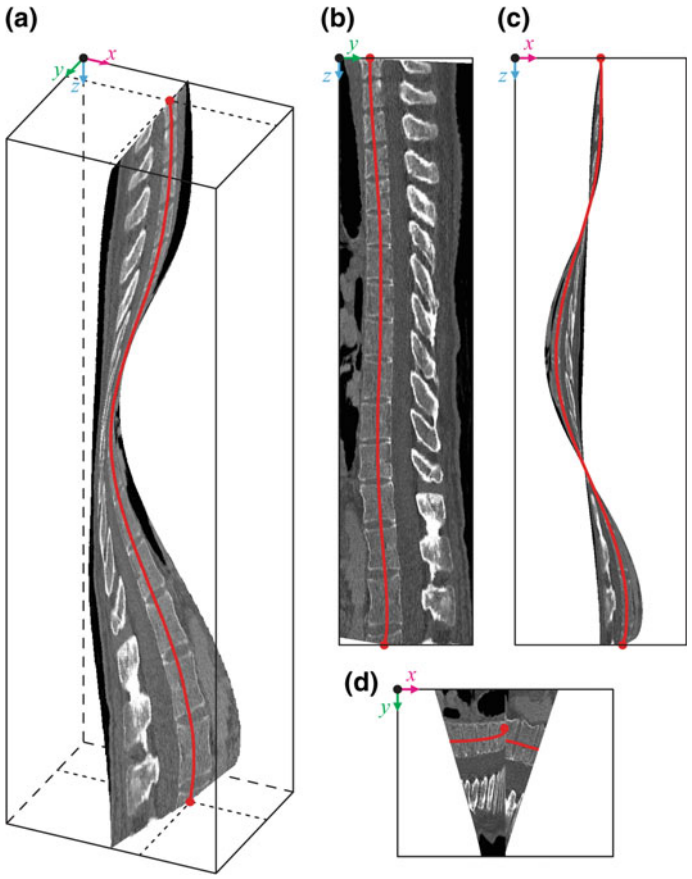


Fig. 21 A sagittal orthogonal curved-planar cross-section $C_{u=u_c}$ of a 3D CT image of a scoliotic spine, shown in **a** 3D view, **b** left sagittal view, **c** posterior coronal view and **d** superior axial view of the image-based coordinate system (*Note* The image-based coordinate system and the spine curve correspond to Figs. 1 and 7)

rotation (Eq. 57) for every point i on the spine curve $\mathbf{c}(i)$. Exactly one sagittal orthogonal curved-planar cross-section passes through the spine curve (i.e. $\Delta x = 0$) and therefore displays the spinal anatomy along its midline, which is represented by a straight line (Fig. 21).

However, in the resulting cross-sections, anatomical deformations are present that result from the intersections of sagittal profiles due to the rotation for angle $\alpha(i)$ about axis x of the image-based coordinate system. To avoid anatomical deformations, the sagittal orthogonal curved-planar cross-section $C_{u=u_c}$ can be obtained as:

$$C_{u=u_c}(y, c_z(i)) = I(R_z(\varphi(i))[c_x(i) + \Delta x, y, c_z(i)]), \tag{77}$$

where matrix $R_z(\varphi(i))$ (Eq. 54) represents the axial vertebral rotation for angle $\varphi(i)$ about axis z of the image-based coordinate system. In this case, the axial vertebral rotation $\varphi(i)$ has to be defined in transverse planes that are orthogonal to axis z of the image-based coordinate system (i.e. $\varphi(i) = \varphi_z(i)$, Eq. 13). As a result, the resulting cross-sections are no longer defined on the basis of the spine-based coordinate system, and therefore the spine curve is no longer represented by a straight line.

Coronal Orthogonal Curved-Planar Cross-Sections

Coronal orthogonal curved-planar cross-sections are obtained by sampling the 3D image on selected coronal planes that are orthogonal to axis v of the spine-based coordinate system. The coronal orthogonal curved-planar cross-section $C_{v=v_c}$ is therefore obtained by selecting a fixed coordinate $v = v_c$, and sampling the 3D image I along coordinates u and w in the spine-based coordinate system:

$$C_{v=v_c}(u, w) = I(u, v_c, w). \tag{78}$$

In the image-based coordinate system, the sampling plane is represented by a curved surface that is parallel to the spine curve $\mathbf{c}(i) = (c_x(i), c_y(i), c_z(i))$ and follows the axial vertebral rotation $\varphi(i)$. If the selected fixed coordinate v_c is represented as $v_c \leftrightarrow c_y(i) + \Delta y$, where Δy is a fixed offset in the anterior or posterior direction from the spine curve $\mathbf{c}(i)$ that corresponds to the coronal offset of point $\mathbf{p}_c = (u_c, v_c, w_c)$ from the origin of the spine-based coordinate system, then the coronal orthogonal curved cross-section $C_{v=v_c}$ can be obtained as:

$$C_{v=v_c}(x, c_z(i)) = I(R_{\hat{\mathbf{t}}(i)}(\varphi(i))R_y(\beta(i))[x, c_y(i) + \Delta y, c_z(i)]), \tag{79}$$

where matrix $R_{\hat{\mathbf{t}}(i)}(\varphi(i))$ (Eq. 19) represents the axial vertebral rotation for angle $\varphi(i)$ about axis defined by $\hat{\mathbf{t}}(i)$ (i.e. $\varphi(i) = \varphi_w(i)$, Eq. 14), and matrix $R_y(\beta(i))$ (Eq. 53) represents the rotation for angle $\beta(i) = \arctan(\hat{t}_x(i)/\hat{t}_z(i))$ about axis y of the image-based coordinate system, considering that $\hat{\mathbf{t}}(i) = [\hat{t}_x(i), \hat{t}_y(i), \hat{t}_z(i)]$ is the unit tangent vector to the spine curve, and $\mathbf{c}(i) = (c_x(i), c_y(i), c_z(i))$ is the center of rotation (Eq. 57) for every point i on the spine curve $\mathbf{c}(i)$. Exactly one coronal orthogonal curved-planar cross-section passes through the spine curve (i.e. $\Delta y = 0$) and therefore displays the spinal anatomy along its midline, which is represented by a straight line (Fig. 22).

However, in the resulting cross-sections, anatomical deformations are present that result from the intersections of coronal profiles due to the rotation for angle $\beta(i)$ about axis y of the image-based coordinate system. To avoid anatomical deformations, the coronal orthogonal curved-planar cross-section $C_{v=v_c}$ can be obtained as:

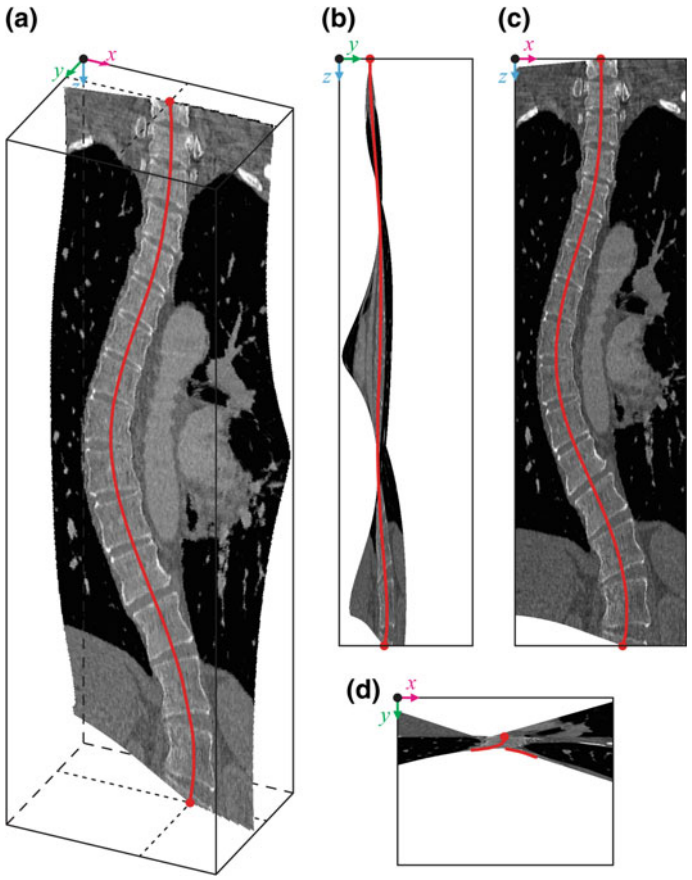


Fig. 22 A coronal orthogonal curved-planar cross-section $C_{v=v_c}$ of a 3D CT image of a scoliotic spine, shown in **a** 3D view, **b** left sagittal view, **c** posterior coronal view and **d** superior axial view of the image-based coordinate system (*Note* The image-based coordinate system and the spine curve correspond to Figs. 1 and 7)

$$C_{v=v_c}(x, c_z(i)) = I(R_z(\varphi(i))[x, c_y(i) + \Delta y, c_z(i)]), \tag{80}$$

where matrix $R_z(\varphi(i))$ (Eq. 54) represents the axial vertebral rotation for angle $\varphi(i)$ about axis z of the image-based coordinate system. In this case, the axial vertebral rotation $\varphi(i)$ has to be defined in transverse planes that are orthogonal to axis z of the image-based coordinate system (i.e. $\varphi(i) = \varphi_z(i)$, Eq. 13). As a result, the resulting cross-sections are no longer defined on the basis of the spine-based coordinate system, and therefore the spine curve is no longer represented by a straight line.

Axial Orthogonal Curved-Planar Cross-Sections

Axial orthogonal curved-planar cross-sections are obtained by sampling the 3D image on selected axial planes that are orthogonal to axis w of the spine-based coordinate system. The axial orthogonal curved-planar cross-section $C_{w=w_c}$ is therefore obtained by selecting a fixed coordinate $w = w_c$, and sampling the 3D image I along coordinates u and v in the spine-based coordinate system:

$$C_{w=w_c}(u, v) = I(u, v, w_c). \tag{81}$$

If the selected fixed coordinate w_c is represented as $w_c \leftrightarrow c_z(i_p)$, where $i = i_p$ defines the point $\mathbf{c}(i_p) = (c_x(i_p), c_y(i_p), c_z(i_p))$ on the spine curve $\mathbf{c}(i)$, then the sampling plane is, in the image-based coordinate system, orthogonal to the spine curve at point $\mathbf{c}(i_p)$ and rotationally aligned with the axial vertebral rotation $\varphi(i_p)$. The axial orthogonal curved-planar cross-section $C_{w=w_c}$ can be therefore obtained as:

$$C_{w=w_c}(x, y) = I(R_{\hat{\mathbf{t}}(i)}(\varphi(i_p))R_y(\beta(i_p))R_x(\alpha(i_p))[x, y, c_z(i_p)]), \tag{82}$$

where matrix $R_{\hat{\mathbf{t}}(i_p)}(\varphi(i_p))$ (Eq. 19) represents the axial vertebral rotation for angle $\varphi(i_p)$ about axis defined by $\hat{\mathbf{t}}(i_p)$ (i.e. $\varphi(i_p) = \varphi_w(i_p)$, Eq. 14), matrix $R_x(\alpha(i_p))$ (Eq. 52) represents the rotation for angle $\alpha(i_p) = \arctan(\hat{t}_y(i_p)/\hat{t}_z(i_p))$ about axis x of the image-based coordinate system, and matrix $R_y(\beta(i_p))$ (Eq. 53) represents the rotation for angle $\beta(i_p) = \arctan(\hat{t}_x(i_p)/\hat{t}_z(i_p))$ about axis y of the image-based coordinate system, considering that $\hat{\mathbf{t}}(i) = [\hat{t}_x(i), \hat{t}_y(i), \hat{t}_z(i)]$ is the unit tangent vector to the spine curve and $\mathbf{c}(i_p) = (c_x(i_p), c_y(i_p), c_z(i_p))$ is the center of rotation (Eq. 57) at the selected point $i = i_p$ on the spine curve $\mathbf{c}(i)$. Axial orthogonal curved-planar cross-sections in general show a geometrically correct shape of the vertebral anatomy, because sampling planes cut through vertebrae at the same anatomical locations (Fig. 23).

In the case the axial vertebral rotation $\varphi(i)$ is defined in transverse planes that are orthogonal to axis z of the image-based coordinate system (i.e. $\varphi(i) = \varphi_z(i)$, Eq. 13), then the axial orthogonal curved cross-section $C_{w=w_c}$ can be obtained as:

$$\begin{aligned} C_{w=w_c}(x, y) &= I(R_z(\varphi(i_p))R_y(\beta(i_p))R_x(\alpha(i_p))[x, y, c_z(i_p)]) \\ &= I(R(\alpha(i_p), \beta(i_p), \varphi(i_p))[x, y, c_z(i_p)]), \end{aligned} \tag{83}$$

where matrix $R_z(\varphi(i_p))$ (Eq. 54) represents the rotation for angle $\varphi(i_p)$ about axis z, and matrix $R(\alpha(i_p), \beta(i_p), \varphi(i_p))$ (Eq. 55) represents the composition of extrinsic rotations for angles $\alpha(i_p)$, β, (i_p) and $\varphi(i_p)$ about axes x, y and z, respectively, of the image-based coordinate system.

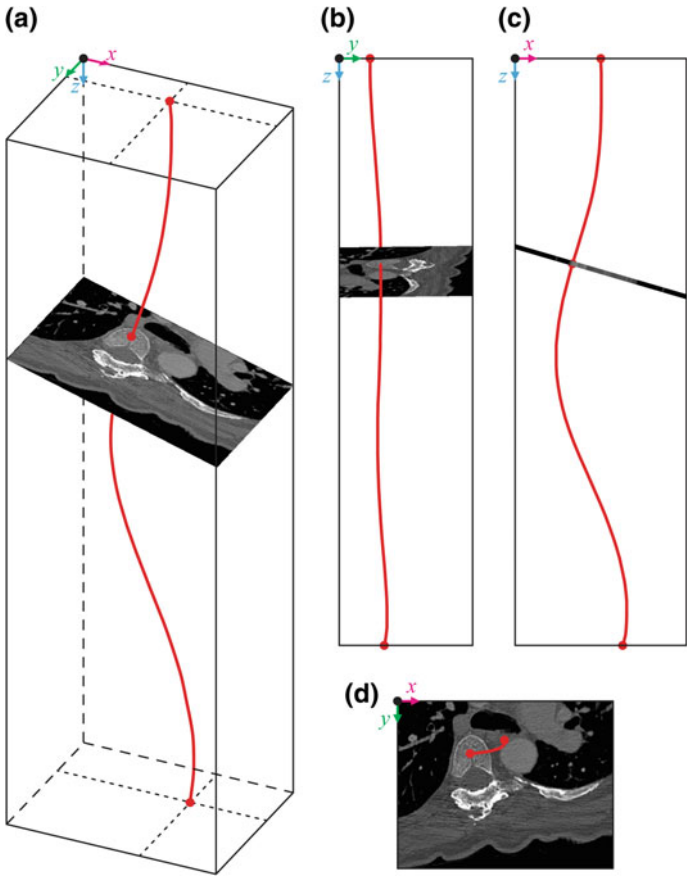


Fig. 23 An axial orthogonal curved-planar cross-section $C_{w=w_c}$ of a 3D CT image of a scoliotic spine, shown in **a** 3D view, **b** left sagittal view, **c** posterior coronal view and **d** superior axial view of the image-based coordinate system (*Note* The image-based coordinate system and the spine curve correspond to Figs. 1 and 7)

3.2.2 Oblique Curved-Planar Reformation

A useful type of CPR is oblique (slanted), meaning that the orthogonal sampling plane is rotated (inclined) for selected angles about the axes of the spine-based coordinate system. The rotation for angles ϑ, ψ and ϕ about axes u, v and w, respectively, of the spine-based coordinate system can be represented by rotation matrices $R_u(\vartheta) = R_x(\vartheta)$ (Eq. 52), $R_v(\psi) = R_y(\psi)$ (Eq. 53) and $R_w(\phi) = R_z(\phi)$ (Eq. 54), respectively. By applying oblique CPR to a 3D spine image, the following oblique curved-planar cross-sections can be obtained:

- *sagittal oblique curved-planar cross-sections* are obtained by sampling the 3D image on selected rotated sagittal orthogonal planes, defined in the spine-based coordinate system (section Sagittal Oblique Curved-Planar Cross-Sections),
- *coronal oblique curved-planar cross-sections* are obtained by sampling the 3D image on selected rotated coronal orthogonal planes, defined in the spine-based coordinate system (section Coronal Oblique Curved-Planar Cross-Sections),
- *axial oblique curved-planar cross-sections* are obtained by sampling the 3D image on selected rotated axial orthogonal planes, defined in the spine-based coordinate system (section Axial Oblique Curved-Planar Cross-Sections).

The rotation principle is therefore in general the same as in the case of oblique MPR (Sect. 3.1.2), however, in the case of oblique CPR, only rotation about one axis is usually applied for sagittal and coronal cross-sections so that they remain aligned with the observed spinal anatomy. As a result, by selecting a point $\mathbf{p}_c = (u_c, v_c, w_c)$ in the spine-based coordinate system $\mathbb{R}^3_S$ and rotation angles $\vartheta = \vartheta_p$, $\psi = \psi_p$, and $\phi = \phi_p$, exactly one sagittal ($u = u_c$ and $\phi = \phi_p$), one coronal ($v = v_c$ and $\phi = \phi_p$) and three axial ($w = w_c$ and $\vartheta = \vartheta_p$, $\psi = \psi_p$ or both) oblique curved-planar cross-sections can be defined through $\mathbf{p}_c$. If $\mathbf{p}_c$ is located on the spine curve $\mathbf{c}(i)$, e.g. at point $i = i_p$ so that $\mathbf{p}_c = (c_x(i_p), c_y(i_p), c_z(i_p))$, the obtained sagittal and coronal cross-sections show, irrespectively of the shape of the spine, the spinal anatomy at its midline along its entire length, while the obtained axial cross-section is arbitrarily inclined against the midline of the spinal anatomy.

Sagittal Oblique Curved-Planar Cross-Sections

The sagittal oblique curved-planar cross-section $C_{u=u_c,\phi=\phi_p}$ is obtained by sampling the 3D image I on the sagittal orthogonal plane at the selected fixed coordinate $u = u_c$ that is additionally rotated for angle $\phi = \phi_p$ about axis w of the spine-based coordinate system:

$$C_{u=u_c,\phi=\phi_p}(v, w) = I(R_w(\phi_p)[u_c, v, w]), \tag{84}$$

with the center of rotation at point $\mathbf{p}_c = (u_c, v_c, w_c)$ (Eq. 57). The rotation angle must be on the closed interval $\phi_p \in [-\frac{\pi}{4}, +\frac{\pi}{4}]$, otherwise the cross-section turns into a coronal oblique curved-planar cross-section. In the image-based coordinate system, the sampling plane is represented by a curved surface that is parallel to the spine curve $\mathbf{c}(i) = (c_x(i), c_y(i), c_z(i))$ and follows the axial vertebral rotation $\varphi(i)$ with an offset of $\phi = \phi_p$. If the selected fixed coordinate u_c is represented as $u_c \leftrightarrow c_x(i) + \Delta x$, where Δx is a fixed offset in the left or right direction from the spine curve $\mathbf{c}(i)$ that corresponds to the sagittal offset of point $\mathbf{p}_c = (u_c, v_c, w_c)$ from the origin of the spine-based coordinate system, then the sagittal oblique curved-planar cross-section $C_{u=u_c,\phi=\phi_p}$ can be obtained as:

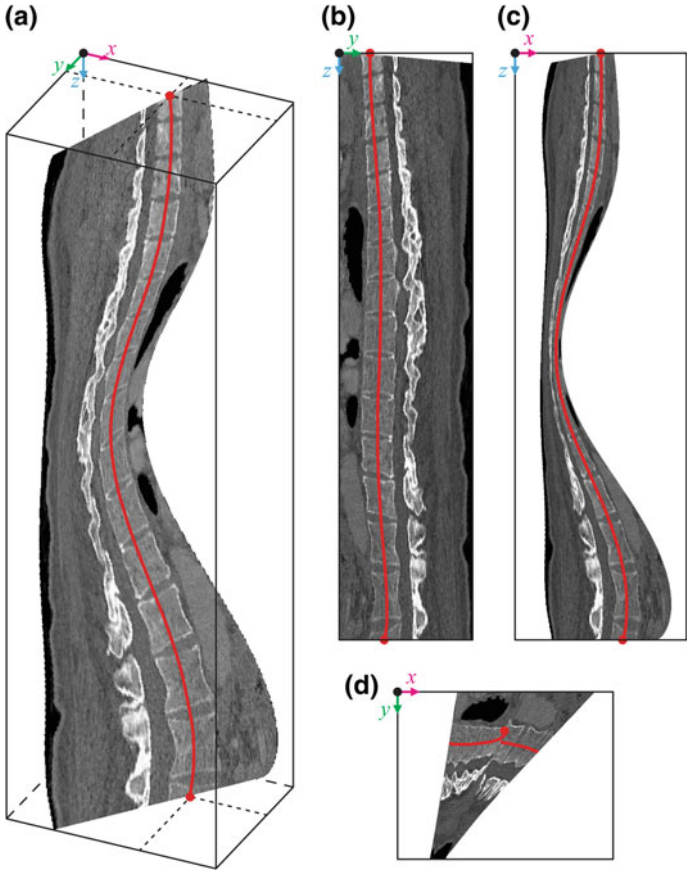


Fig. 24 A sagittal oblique curved-planar cross-section $C_{u=u_c,\phi=\phi_p}$ of a 3D CT image of a scoliotic spine, shown in **a** 3D view, **b** left sagittal view, **c** posterior coronal view and **d** superior axial view of the image-based coordinate system (*Note* The image-based coordinate system and the spine curve correspond to Figs. 1 and 7)

$$C_{u=u_c,\phi=\phi_p}(y, c_z(i)) = I(R_{\hat{t}(i)}(\varphi(i) + \phi_p)R_x(\alpha(i))[c_x(i) + \Delta x, y, c_z(i)]), \qquad (85)$$

where matrix $R_{\hat{\mathbf{t}}(i)}(\varphi(i) + \phi_p)$ (Eq. 19) represents the axial vertebral rotation for angle $\varphi(i) + \phi_p$ about axis defined by $\hat{\mathbf{t}}(i)$ (i.e. $\varphi(i) = \varphi_w(i)$, Eq. 14), and matrix $R_x(\alpha(i))$ (Eq. 52) represents the rotation for angle $\alpha(i) = \arctan(\hat{t}_y(i)/\hat{t}_z(i))$ about axis x of the image-based coordinate system, considering that $\hat{\mathbf{t}}(i) = [\hat{t}_x(i), \hat{t}_y(i), \hat{t}_z(i)]$ is the unit tangent vector to the spine curve, and $\mathbf{c}(i) = (c_x(i), c_y(i), c_z(i))$ is the center of rotation for every point i on the spine curve $\mathbf{c}(i)$. Figure 24 displays the sagittal oblique curved-planar cross-section that passes through the spine curve (i.e. $\Delta x = 0$) and is rotated for $\phi = \phi_p = 25°$.

However, in the resulting cross-sections, anatomical deformations are present that result from the intersections of sagittal profiles due to the rotation for angle $\alpha(i)$ about axis x of the image-based coordinate system. To avoid anatomical deformations, the sagittal oblique curved-planar cross-section $C_{u=u_c,\phi=\phi_p}$ can be obtained as:

$$C_{u=u_c,\phi=\phi_p}(y, c_z(i)) = I(R_z(\varphi(i) + \phi_p)[c_x(i) + \Delta x, y, c_z(i)]), \tag{86}$$

where matrix $R_z\big(\varphi(i) + \phi_p\big)$ (Eq. 54) represents the axial vertebral rotation for angle $\varphi(i) + \phi_p$ about axis z of the image-based coordinate system. In this case, the axial vertebral rotation $\varphi(i)$ has to be defined in transverse planes that are orthogonal to axis z of the image-based coordinate system (i.e. $\varphi(i) = \varphi_z(i)$, Eq. 13). As a result, the resulting cross-sections are no longer defined on the basis of the spine-based coordinate system, and therefore the spine curve is no longer represented by a straight line.

Coronal Oblique Curved-Planar Cross-Sections

The coronal oblique curved-planar cross-section $C_{v=v_c,\phi=\phi_p}$ is obtained by sampling the 3D image I on the coronal orthogonal plane at the selected fixed coordinate $v = v_c$ that is additionally rotated for angle $\phi = \phi_p$ about axis w of the spine-based coordinate system:

$$C_{v=v_c,\phi=\phi_p}(u, w) = I(R_w(\phi_p)[u, v_c, w]), \tag{87}$$

with the center of rotation at point $\mathbf{p}_c = (u_c, v_c, w_c)$. The rotation angle must be on the closed interval $\phi_p \in [-\frac{\pi}{4}, +\frac{\pi}{4}]$, otherwise the cross-section turns into a sagittal oblique curved-planar cross-section. In the image-based coordinate system, the sampling plane is represented by a curved surface that is parallel to the spine curve $\mathbf{c}(i) = \big(c_x(i), c_y(i), c_z(i)\big)$ and follows the axial vertebral rotation $\varphi(i)$ with an offset of $\phi = \phi_p$. If the selected fixed coordinate v_c is represented as $v_c \leftrightarrow c_y(i) + \Delta y$, where Δy is a fixed offset in the anterior or posterior direction from the spine curve $\mathbf{c}(i)$ that corresponds to the sagittal offset of point $\mathbf{p}_c = (u_c, v_c, w_c)$ from the origin of the spine-based coordinate system, then the coronal oblique curved-planar cross-section $C_{v=v_c,\phi=\phi_p}$ can be obtained as:

$$C_{v=v_c,\phi=\phi_p}(x, c_z(i)) = I(R_{\hat{\mathbf{t}}(i)}(\varphi(i) + \phi_p)R_y(\beta(i))[x, c_y(i) + \Delta y, c_z(i)]), \tag{88}$$

where matrix $R_{\hat{\mathbf{t}}(i)}(\varphi(i) + \phi_p)$ (Eq. 19) represents the axial vertebral rotation for angle $\varphi(i) + \phi_p$ about axis defined by $\hat{\mathbf{t}}(i)$ (i.e. $\varphi(i) = \varphi_w(i)$, Eq. 14), and matrix $R_y(\beta(i))$ (Eq. 53) represents the rotation for angle $\beta(i) = \arctan(\hat{t}_x(i)/\hat{t}_z(i))$ about axis y of the image-based coordinate system, considering that $\hat{\mathbf{t}}(i) = [\hat{t}_x(i), \hat{t}_y(i), \hat{t}_z(i)]$ is the unit tangent vector to the spine curve, and $\mathbf{c}(i) = (c_x(i), c_y(i), c_z(i))$ is the center of

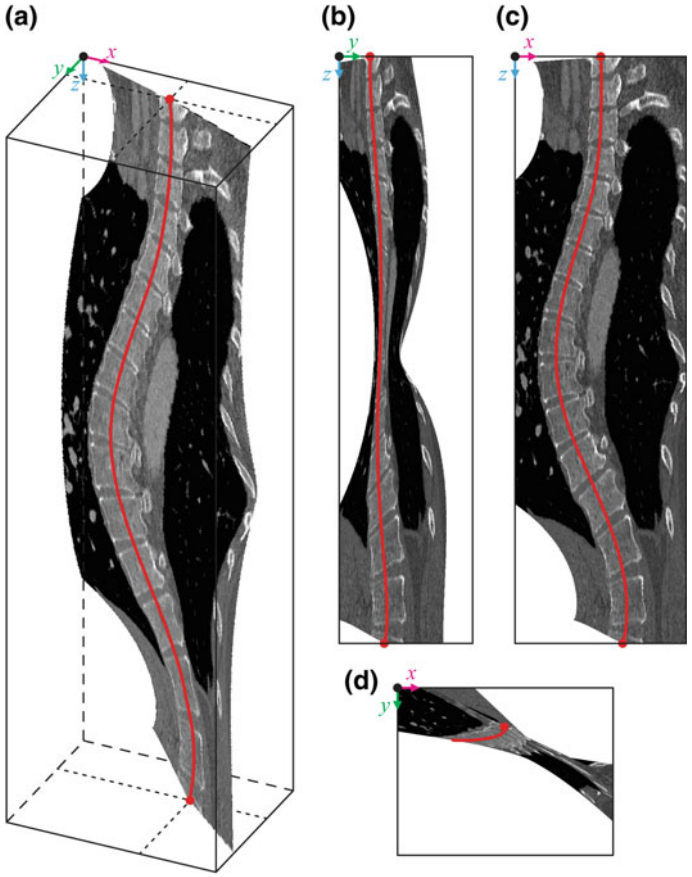


Fig. 25 A coronal oblique curved-planar cross-section $C_{v=v_c,\phi=\phi_p}$ of a 3D CT image of a scoliotic spine, shown in **a** 3D view, **b** left sagittal view, **c** posterior coronal view and **d** superior axial view of the image-based coordinate system (*Notes* The image-based coordinate system and the spine curve correspond to Figs. 1 and 7)

rotation (Eq. 57) for every point i on the spine curve $\mathbf{c}(i)$. Figure 25 displays the coronal oblique curved-planar cross-section that passes through the spine curve (i.e. $\Delta y = 0$) and is rotated for $\phi = \phi_p = 25°$.

However, in the resulting cross-sections, anatomical deformations are present that result from the intersections of coronal profiles due to the rotation for angle $\beta(i)$ about axis y of the image-based coordinate system. To avoid anatomical deformations, the coronal oblique curved-planar cross-section $C_{v=v_c,\phi=\phi_p}$ can be obtained as:

$$C_{v=v_c,\phi=\phi_p}(x, c_z(i)) = I(R_z(\varphi(i) + \phi_p)[x, c_y(i) + \Delta y, c_z(i)]), \tag{89}$$

where matrix $R_z(\varphi(i) + \phi_p)$ (Eq. 54) represents the axial vertebral rotation for angle $\varphi(i) + \phi_p$ about axis z of the image-based coordinate system. In this case, the axial vertebral rotation $\varphi(i)$ has to be defined in transverse planes that are orthogonal to axis z of the image-based coordinate system (i.e. $\varphi(i) = \varphi_z(i)$, Eq. 13). As a result, the resulting cross-sections are no longer defined on the basis of the spine-based coordinate system, and therefore the spine curve is no longer represented by a straight line.

Axial Oblique Curved-Planar Cross-Sections

Axial oblique curved-planar cross-sections are obtained by rotating the axial orthogonal curved-planar cross-section for angles $\vartheta = \vartheta_p$ about axis u and/or for angle $\psi = \psi_p$ about axis v of the spine-based coordinate system. However, the resulting cross-sections are similar to axial oblique multi-planar cross-sections (section Axial Oblique Multi-planar Cross-Sections), with the difference that they are centered at the selected point $\mathbf{p}_c = (u_c, v_c, w_c)$ in the spine-based coordinate system, and that rotation angles ϑ_p and ψ_p are defined against the axes of the spine-based and not against the axes of the image-based coordinate system.

3.3 Cross-Sectional Visualization

The efficiency of the spine-based coordinate system for cross-sectional visualization of 3D spine images can be observed in cross-sections that result from flattening different types of reformations (i.e. MPR and CPR) onto a 2D plane. The examples are presented for a 3D MR image of a normal spine (Figs. 26, 28 and 30) and for a 3D CT image of a scoliotic spine (Figs. 27, 29 and 31).

By applying MPR, sagittal orthogonal and oblique multi-planar cross-sections of the normal spine (Fig. 26a, b) simultaneously display the anatomy of all vertebrae along the whole length of the spine, although in the sagittal oblique multi-planar cross-section, cervical vertebrae go out of the sampling plane because the center of rotation was at a selected point on the thoracic spine curve. On the other hand, in the case of the scoliotic spine (Fig. 27a, b), the anatomy of all vertebrae cannot be simultaneously displayed in sagittal orthogonal or oblique multi-planar cross-sections, because the vertebrae come in and out of the sampling plane due to the curvature of the spine in the coronal plane. The situation is reversed in the case of coronal orthogonal and oblique multi-planar cross-sections. In the case of the normal spine (Fig. 28a, b), vertebrae go out of the sampling plane due to the curvature of the spine in the sagittal plane, while in the case of the scoliotic spine (Fig. 29a, b), all vertebrae can be simultaneously observed along the whole length of the spine in coronal orthogonal or oblique multi-planar cross-sections. Axial orthogonal multi-planar cross-sections (Figs. 30a and 31a) in general do not display the

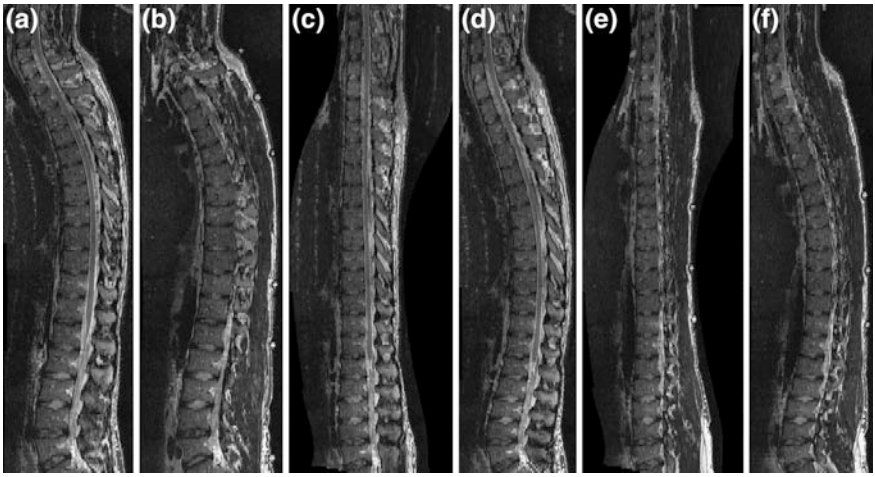


Fig. 26 A sagittal cross-section of a 3D MR image of a normal spine, obtained by **a** orthogonal MPR, **b** oblique MPR, **c** orthogonal CPR (with anatomical deformations), **d** orthogonal CPR (without anatomical deformations), **e** oblique CPR (with anatomical deformations) and **f** oblique CPR (without anatomical deformations)

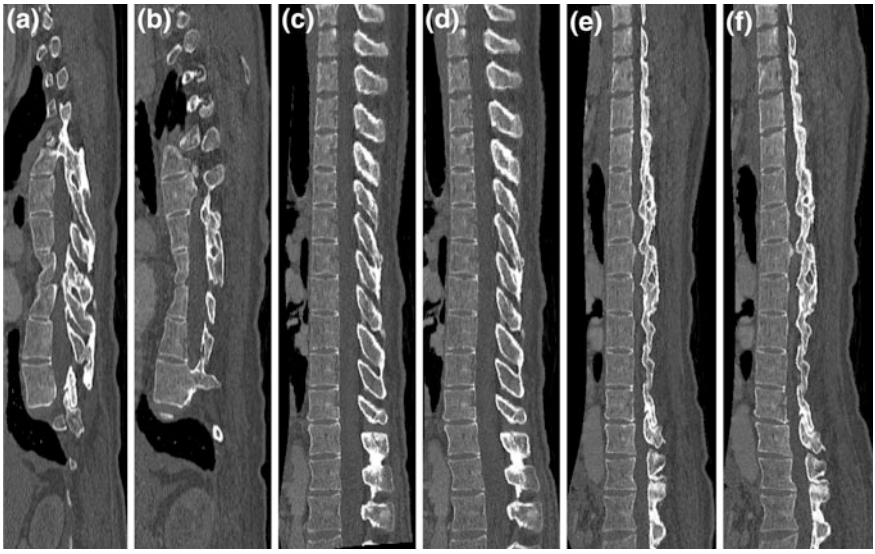


Fig. 27 A sagittal cross-section of a 3D CT image of a scoliotic spine, obtained by **a** orthogonal MPR, **b** oblique MPR, **c** orthogonal CPR (with anatomical deformations), **d** orthogonal CPR (without anatomical deformations), **e** oblique CPR (with anatomical deformations) and **f** oblique CPR (without anatomical deformations)

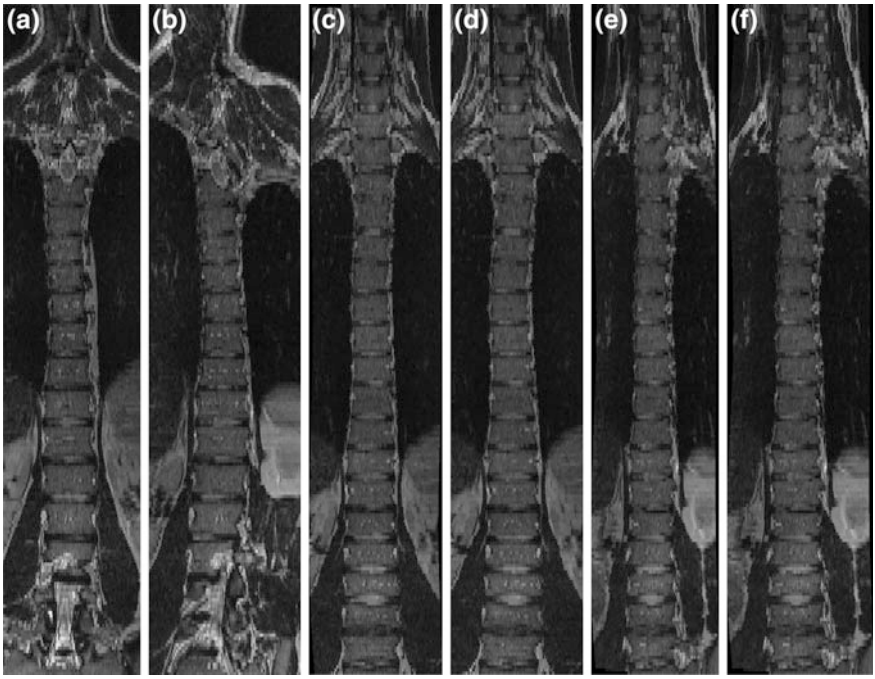


Fig. 28 A coronal cross-section of a 3D MR image of a normal spine, obtained by **a** orthogonal MPR, **b** oblique MPR, **c** orthogonal CPR (with anatomical deformations), **d** orthogonal CPR (without anatomical deformations), **e** oblique CPR (with anatomical deformations) and **f** oblique CPR (without anatomical deformations)

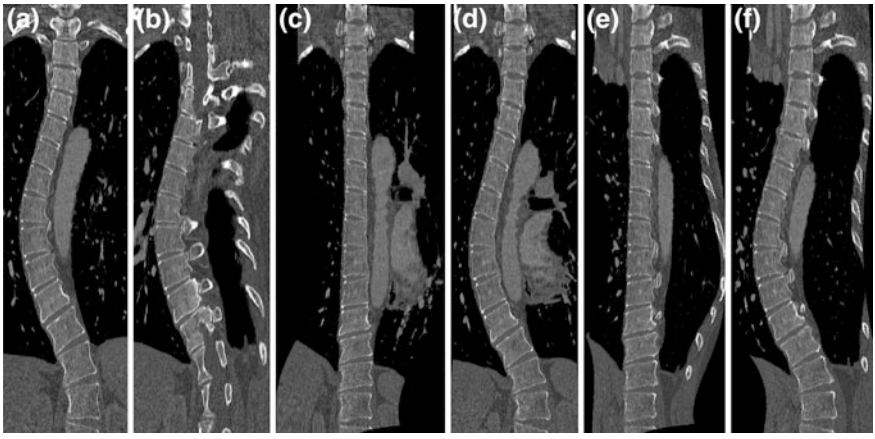


Fig. 29 A coronal cross-section of a 3D CT image of a scoliotic spine, obtained by **a** orthogonal MPR, **b** oblique MPR, **c** orthogonal CPR (with anatomical deformations), **d** orthogonal CPR (without anatomical deformations), **e** oblique CPR (with anatomical deformations) and **f** oblique CPR (without anatomical deformations)

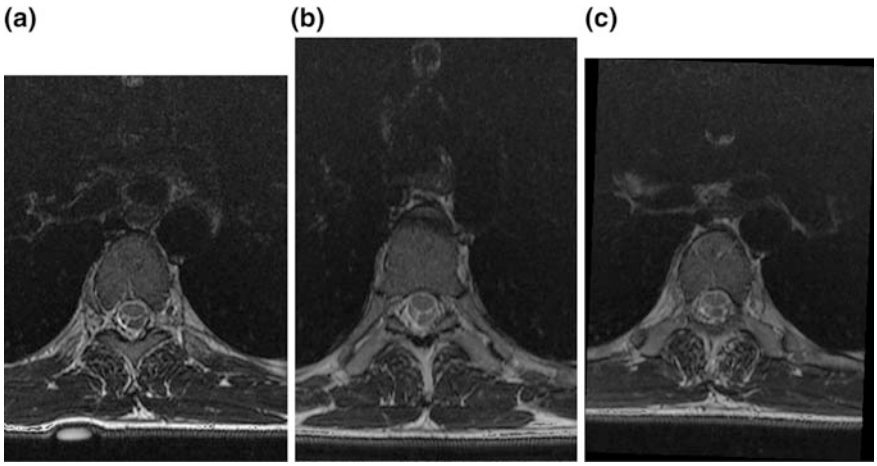


Fig. 30 An axial cross-section of a 3D MR image of a normal spine, obtained by **a** orthogonal MPR, **b** oblique MPR and **c** orthogonal CPR

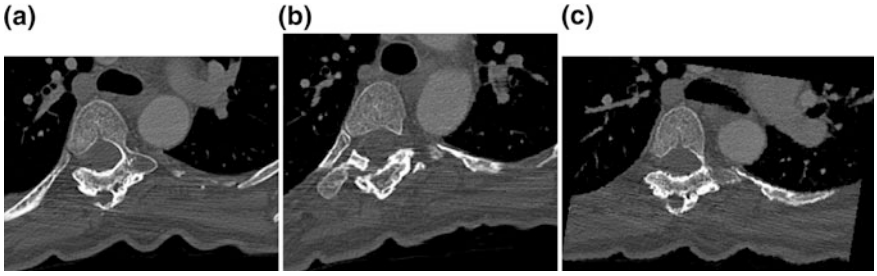


Fig. 31 An axial cross-section of a 3D CT image of a scoliotic spine, obtained by **a** orthogonal MPR, **b** oblique MPR and **c** orthogonal CPR

vertebral anatomical structures at corresponding anatomical locations. Although visualization at corresponding anatomical locations can be achieved by axial oblique multi-planar cross-sections (Figs. 30b and 31b), it is difficult to determine the correct angles of rotation so that the resulting cross-sections are orthogonal to the spine curve.

The above mentioned problems with visualization are solved by applying CPR. Sagittal and coronal orthogonal and oblique curved-planar cross-sections simultaneously display the anatomy of all vertebrae along the whole length of the spine, both in the case of the normal spine (Figs. 26c–f and 28c–e) and scoliotic spine (Figs. 27c–f and 29c–e). In the case the spine curve is displayed as a straight line in curved-planar cross-sections of both the normal spine (Figs. 26c, e and 28c, e) and scoliotic spine (Figs. 27c, e and 29c, e), then anatomical deformations are present in the resulting cross-sections due to the intersections of successive sagittal and

coronal profiles. In such cross-sections, however, morphometric measurements based on Euclidean metrics are not possible. On the other hand, curved-planar cross-sections that display the anatomy of all vertebrae along the whole length of the spine without anatomical deformations can be obtained, but in this case the spine curve is no longer always displayed as a straight line. As a result, the spinal curvature of the normal spine in the sagittal plane is visible in sagittal orthogonal and oblique curved-planar cross-sections (Fig. 26d, f), while the spinal curvature of the scoliotic spine in the coronal plane is visible in coronal orthogonal and oblique curved-planar cross-sections (Fig. 29d, f). However, a significant advantage of such visualization is that morphometric measurements based on Euclidean metrics can be performed directly from these cross-sections. The axial orthogonal curved-planar cross-sections are always orthogonal to the spine curve and aligned with the corresponding axial vertebral rotation, both in the case of the normal spine (Fig. 30c) and scoliotic spine (Fig. 31c).

4 Conclusion

Techniques for visualization and quantitative evaluation of medical images are in general valuable for the development of image-assisted diagnosis, planning of surgical interventions and assessment of medical treatment outcomes. In the field of spine image analysis, visualization and quantitative evaluation of spinal curvature and axial vertebral rotation is important not only for planning of orthopaedic surgical procedures and analysis of surgical results, but also for diagnosing and monitoring of the progression of spinal deformities. Computer-assisted visualization and quantitative evaluation of 3D spine images therefore remain challenging tasks in the field of medical image analysis.

In this chapter, automated determination of the spine-based coordinate system for an efficient cross-sectional visualization of 3D spine images was presented. The introduction of the spine-based coordinate system allows to determine curved-planar cross-sections that follow the spine curve and axial vertebral rotation along the whole length of the spine. The main purpose of the described image reformation technique is to reduce the structural complexity in favor of an improved feature perception of the spine, and to provide clinically relevant quantitative analysis of the 3D spinal anatomy. Displaying the whole length of the spine within a single 2D image makes the inspection of images quicker and more precise, while the probability of overlooking certain important features of the spine is reduced. The spine curve and rotation of vertebrae about the spine curve can be obtained automatically and used to transform 3D spine images from the image-based to the spine-based coordinate system. As the spine curve and axial vertebral rotation are inherent properties of the spine and therefore not affected by rigid body transformations, the generated curved-planar cross-sections are independent of the position of the patient in the scanner and of the orientation of the image acquisition plane. When visualizing and inspecting 3D images in the spine-based coordinate system,

pathological anatomy is oriented comparable to healthy anatomy, thus facilitating image interpretation and allowing a more objective evaluation and diagnosis of the abnormalities, especially in the case of significant coronal (e.g. scoliosis) or sagittal (e.g. hyper-kyphosis or hyper-lordosis) spinal curvatures. Furthermore, the knowledge on the location and orientation of the spine in 3D can be exploited by other image analysis techniques.

Acknowledgements The author would like to thank B. Ibragimov (University of Ljubljana, Faculty of Electrical Engineering, Slovenia) for manually segmenting the spine from a CT image that was used for rendering purposes.

References

1. Aaro S, Dahlborn M (1981) Estimation of vertebral rotation and the spinal and rib cage deformity in scoliosis by computer-tomography. Spine 6(5):460–467
2. Adam CJ, Askin GN (2006) Automatic measurement of vertebral rotation in idiopathic scoliosis. Spine 31(3):E80–E83
3. Angelelli P, Hauser H (2011) Straightening tubular flow for side-by-side visualization. IEEE Trans Visual Comput Graphics 17(12):2063–2070
4. Apicella PL, Mirowitz SA (1995) Interactive multiplanar reformation of conventional two-dimensional MR images. Clin Imaging 19(4):279–282
5. Aylward SR, Bullitt E (2002) Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction. IEEE Trans Med Imaging 21(2):61–75
6. Berthonnaud E, Dimnet J (2007) Analysis of structural features of deformed spines in frontal and sagittal projections. Comput Med Imaging Graph 31(1):9–16
7. Bhargavan M, Kaye AH, Forman HP, Sunshine JH (2009) Workload of radiologists in United States in 2006–2007 and trends since 1991–1992. Radiology 252(2):458–467
8. Birchall D, Hughes DG, Hindle J, Robinson L, Williamson JB (1997) Measurement of vertebral rotation in adolescent idiopathic scoliosis using three-dimensional magnetic resonance imaging. Spine 22(20):2403–2407
9. Bitter I, Sato M, Bender M, McDonnell KT, Kaufman A, Wan M (2000) CEASAR: a smooth, accurate and robust centerline extraction algorithm. In: Ertl T, Hamann B, Varshney A (eds) Proceedings of the IEEE visualization 2000 conference—Vis 2000, 8–13 Oct 2000. IEEE, Salt Lake City, pp 45–52
10. Börnert P (2003) 2D RF-pulse-encoded curved-slice imaging. Magn Reson Mater Phys, Biol Med 16(2):86–92
11. Börnert P, Schäffter T (1996) Curved slice imaging. Magn Reson Med 36(6):932–939
12. Cousins JP, Haughton VM (2009) Magnetic resonance imaging of the spine. J Am Acad Orthop Surg 17(1):22–30
13. Drerup B, Hierholzer E (1992) Evaluation of frontal radiographs of scoliotic spines—part I: measurement of position and orientation of vertebrae and assessment of clinical shape parameters. J Biomech 25(11):1357–1362
14. Drerup B, Hierholzer E (1992) Evaluation of frontal radiographs of scoliotic spines—part II: relations between lateral deviation, lateral tilt and axial rotation of vertebrae. J Biomech 25 (12):1443–1450
15. Drerup B, Hierholzer E (1996) Assessment of scoliotic deformity from back shape asymmetry using an improved mathematical model. Clin Biomech 11(7):376–383
16. Fischler MA, Bolles RC (1981) Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun ACM 24(6):381–395

17. Forsberg D, Lundström C, Andersson M, Knutsson H (2014) Model-based registration for assessment of spinal deformities in idiopathic scoliosis. Phys Med Biol 59(2):311–326
18. Forsberg D, Lundström C, Andersson M, Vavruch L, Tropp H, Knutsson H (2013) Fully automatic measurements of axial vertebral rotation for assessment of spinal deformity in idiopathic scoliosis. Phys Med Biol 58(6):1775–1787
19. Ge Y, Stelts DR, Vining DJ, Wang J, Ge Y (1999) Computing the centerline of a colon: a robust and efficient method based on 3D skeletons. J Comput Assist Tomogr 23(5):786–794
20. Göçen S, Havitçioglu H, Alici E (1999) A new method to measure vertebral rotation from CT scans. Eur Spine J 8(4):261–265
21. Gong J-S, Xu J-M (2004) Role of curved planar reformations using multidetector spiral CT in diagnosis of pancreatic and peripancreatic diseases. World J Gastroenterol 10(13):1943–1947
22. Hanaoka S, Nomura Y, Nemoto M, Masutani Y, Maeda E, Yoshikawa T, Hayashi N, Yoshioka N, Ohtomo K (2010) Automated segmentation method for spinal column based on a dual elliptic column model and its application for virtual spinal straightening. J Comput Assist Tomogr 34(1):156–162
23. Haughton VM, Rogers B, Meyerand E, Resnick DK (2002) Measuring the axial rotation of lumbar vertebrae in vivo with MR imaging. Am J Neuroradiol 23(7):1110–1116
24. Hay O, Hershkovitz I, Rivlin E (2009) Spine curve modeling for quantitative analysis of spinal curvature. In: Proceedings of the 31st IEEE annual international conference of the engineering in medicine and biology society (EMBS)—EMBC 2009, 2–6 Sept 2009. IEEE, Minneapolis, pp 6356–6359
25. He S, Dai R, Lu B, Cao C, Bai H, Jing B (2001) Medial axis reformation: a new visualization method for CT angiography. Acad Radiol 8(8):726–733
26. Ho EKW, Upadhyay SS, Chan FL, Hsu LCS, Leong JCY (1993) New methods of measuring vertebral rotation from computed tomographic scans: an intraobserver and interobserver study on girls with scoliosis. Spine 18(9):1173–1177
27. Huysmans T, Haex B, Van Audekercke R, Vander Sloten J, Van Der Perre G (2004) Three-dimensional mathematical reconstruction of the spinal shape, based on active contours. J Biomech 37(11):1793–1798
28. Jochimsen TH, Norris DG (2002) Single-shot curved slice imaging. Magn Reson Mater Phys, Biol Med 14(1):50–55
29. Kadoury S, Cheriet F, Labelle H (2010) Self-calibration of biplanar radiographic images through geometric spine shape descriptors. IEEE Trans Biomed Eng 57(7):1663–1675
30. Kadoury S, Labelle H, Paragios N (2011) Automatic inference of articulated spine models in CT images using high-order markov random fields. Med Image Anal 15(4):426–437
31. Kadoury S, Labelle H, Paragios N (2013) Spine segmentation in medical images using manifold embeddings and higher-order MRFs. IEEE Trans Med Imaging 32(7):1227–1238
32. Kadoury S, Shen J, Parent S (2014) Global geometric torsion estimation in adolescent idiopathic scoliosis. Med Biol Eng Comput 52:309–319
33. Kaminsky J, Klinge P, Rodt T, Bokemeyer M, Luedemann W, Samii M (2004) Specially adapted interactive tools for an improved 3D-segmentation of the spine. Comput Med Imaging Graph 28(3):119–127
34. Kanitsar A, Fleischmann D, Wegenkittl R, Felkel P, Gröller ME (2002) CPR—curved planar reformation. In: Moorhead R, Gross M, Joy KI (eds) Proceedings of the IEEE visualization 2002 conference—Vis 2002, 27 Oct–1 Nov 2002. IEEE, Boston, pp 37–44
35. Kanitsar A, Wegenkittl R, Fleischmann D, Gröller ME (2003) Advanced curved planar reformation: flattening of vascular structures. In: Turk G, van Vijk JJ, Moorehad R (eds) Proceedings of the IEEE visualization 2003 conference—Vis 2003, 19–24 Oct 2003. IEEE, Seattle, pp 43–50
36. Klinder T, Ostermann J, Ehm M, Franz A, Kneser R, Lorenz C (2009) Automated model-based vertebra detection, identification, and segmentation in CT images. Med Image Anal 13 (3):471–482
37. Kouwenhoven J-W, Bartels LW, Vincken KL, Viergever MA, Verbout AJ, Delhaas T, Castelein RM (2007) The relation between organ anatomy and pre-existent vertebral rotation

in the normal spine: magnetic resonance imaging study in humans with situs inversus totalis. Spine 32(10):1123–1128
38. Kouwenhoven JWM, Vincken KL, Bartels LW, Castelein RM (2006) Analysis of preexistent vertebral rotation in the normal spine. Spine 31(13):1467–1472
39. Krismer M, Sterzinger W, Christian H, Frischhut B, Bauer R (1996) Axial rotation measurement of scoliotic vertebrae by means of computed tomography scans. Spine 21 (5):576–581
40. Lam GC, Hill DL, Le LH, Raso JV, Lou EH (2008) Vertebral rotation measurement: a summary and comparison of common radiographic and CT methods. Scoliosis 3:16
41. Lampe OD, Correa C, Ma K-L, Hauser H (2009) Curve-centric volume reformation for comparative visualization. IEEE Trans Visual Comput Graphics 15(6):1235–1242
42. Law TY, Heng PA (2000) Automatic centerline extraction for 3D virtual bronchoscopy. In: Delp SL, DiGioia AM, Jaramaz B (eds) Proceedings of the 3rd international conference on medical image computing and computer-assisted intervention—MICCAI 2000, 11–14 Oct 2000. Lecture notes in computer science (LNCS), vol 1935. Springer, Pittsburgh, pp 786–795
43. Leonardi M, Righini A, Agati R, Brayda G, Zanotti B (1991) Curved CT reformatted images of head scans. J Comput Assist Tomogr 15(6):1074–1076
44. Li ALK, Yen D (2011) Effect of increased MRI and CT scan utilization on clinical decision-making in patients referred to a surgical clinic for back pain. Can J Surg 54(2):128–132
45. Liljenqvist UR, Allkemper T, Hackenberg L, Link TM, Steinbeck J, Halm HFH (2002) Analysis of vertebral morphology in idiopathic scoliosis with use of magnetic resonance imaging and multiplanar reconstruction. J Bone Joint Surg Am 84(3):359–368
46. Maddah M, Soltanian-Zadeh H, Afzali-Kusha A, Maddah M (2003) Snake modeling and distance transform approach to vascular centerline extraction and quantification. Comput Med Imaging Graph 27(6):503–512
47. Masharawi Y, Salame K, Mirovsky Y, Peleg S, Dar G, Steinberg N, Hershkovitz I (2008) Vertebral body shape variation in the thoracic and lumbar spine: characterization of its asymmetry and wedging. Clin Anat 21(1):46–54
48. Menten R, Mousny M, Saint-Martin C, Clapuyt P (2005) Planispheric multiplanar reformatted CT: a new method for evaluation of paediatric congenital spine abnormalities. Pediatr Radiol 35(6):627–629
49. Neubert A, Fripp J, Engstrom C, Schwarz R, Lauer L, Salvado O, Crozier S (2012) Automated detection, 3D segmentation and analysis of high resolution spine MR images using statistical shape models. Phys Med Biol 57(24):8357–8376
50. Newton PO, Hahn GW, Fricka KB, Wenger DR (2002) Utility of three-dimensional and multiplanar reformatted computed tomography for evaluation of pediatric congenital spine abnormalities. Spine 27(8):844–850
51. Ochi T, Shimizu K, Yasuhara Y, Shigesawa T, Mochizuki T, Ikezoe J (1999) Curved planar reformatted CT angiography: usefulness for the evaluation of aneurysms at the carotid siphon. Am J Neuroradiol 20(6):1025–1030
52. Panjabi MM, Duranceau J, Goel V, Oxland T, Takata K (1991) Cervical human vertebrae. Quantitative three-dimensional anatomy of the middle and lower regions. Spine 16 (8):861–869
53. Panjabi MM, Goel V, Oxland T, Takata K, Duranceau J, Krag M, Price M (1992) Human lumbar vertebrae. Quantitative three-dimensional anatomy. Spine 17(3):299–306
54. Panjabi MM, Takata K, Goel V, Federico D, Oxland T, Duranceau J, Krag M (1991) Thoracic human vertebrae. Quantitative three-dimensional anatomy. Spine 16(8):888–901
55. Patwardhan AG, Rimkus A, Gavin TM, Bueche M, Meade KP, Bielski R, Ibrahim K (1996) Geometric analysis of coronal decompensation in idiopathic scoliosis. Spine 21 (10):1192–1200
56. Peng Z, Zhong J, Wee W, Lee J-H (2005) Automated vertebra detection and segmentation from the whole spine MR images. In: Zhang YT, Xu LX, Roux C, Zhuang TG, Tamura T, Galiana H (eds) Proceedings of the 27th IEEE annual international conference of the

engineering in medicine and biology society (EMBS)—EMBC 2005, 1–4 Sept 2005. IEEE, Shanghai, pp 2527–2530
57. Perchet D, Fetita C, Preteux F (2004) Advanced navigation tools for virtual bronchoscopy. In: Dougherty ER, Astola JT, Egiazarian KO (eds) Proceedings of the SPIE medical imaging 2004: image processing conference, 14–19 Feb 2004, vol 5298. SPIE, San Jose, pp 147–158
58. Poncet P, Dansereau J, Labelle H (2001) Geometric torsion in idiopathic scoliosis: three-dimensional analysis and proposal for a new classification. Spine 26(20):2235–2243
59. Poncet P, Trochu F, Dansereau J (1999) Curvilinear three-dimensional modeling of spinal curves with dual kriging. Comput Methods Biomech Biomed Eng 2(4):295–308
60. Prokesch RW, Chow LC, Beaulieu CF, Nino-Murcia M, Mindelzun RE, Bammer R, Huang J, Jeffrey RB (2002) Local staging of pancreatic carcinoma with multi-detector row CT: use of curved planar reformations initial experience. Radiology 225(3):759–765
61. Prokesch RW, Coulam CH, Chow LC, Bammer R, Rubin GD (2002) CT angiography of the subclavian artery: utility of curved planar reformations. J Comput Assist Tomogr 26 (2):199–201
62. Rabassa AE, Guinto FC Jr, Crow WN, Chaljub G, Wright GD, Storey GS (1993) CT of the spine: value of reformatted images. Am J Roentgenol 161(6):1223–1227
63. Raman R, Napel S, Beaulieu CF, Bain ES, Jeffrey RB Jr, Rubin GD (2002) Automated generation of curved planar reformations from volume data: method and evaluation. Radiology 223(1):275–280
64. Raman R, Napel S, Rubin GD (2003) Curved-slab maximum intensity projection: method and evaluation. Radiology 229(1):255–260
65. Roberts CC, McDaniel NT, Krupinski EA, Erly WK (2003) Oblique reformation in cervical spine computed tomography: a new look at an old friend. Spine 28(2):167–170
66. Rogers B, Wiese S, Blankenbaker D, Meyerand E, Haughton V (2005) Accuracy of an automated method to measure rotations of vertebrae from computerized tomography data. Spine 30(6):694–696
67. Rogers BP, Haughton VM, Arfanakis K, Meyerand ME (2002) Application of image registration to measurement of intervertebral rotation in the lumbar spine. Magn Reson Med 48(6):1072–1075
68. Rothman SLG, Dobben GD, Rhodes ML, Jr Glenn WV, Azzawi Y-M (1984) Computed tomography of the spine: curved coronal reformations from serial images. Radiology 150 (1):185–190
69. Rousseeuw PJ, Leroy AM (2003) Robust regression and outlier detection. Wiley series in probability and statistics. Wiley, New York
70. Sakas G (2002) Trends in medical imaging: from 2D to 3D. Comput Graph 26(4):577–587
71. Samara Y, Fiebich M, Dachman AH, Kuniyoshi JK, Doi K, Hoffmann KR (1999) Automated calculation of the centerline of the human colon on CT images. Acad Radiol 6(6):352–359
72. Saroul L, Gerlach S, Hersch RD (2003) Exploring curved anatomic structures with surface sections. In: Turk G, van Vijk JJ, Moorehad R (eds) Proceedings of the IEEE visualization 2003 conference—Vis 2003, 19–24 Oct 2003. IEEE, Seattle, pp 27–34
73. Soni K, Bhargavan M, Forman HP, Sunshine JH (2010) Who's underworked and who's overworked now? An update on radiologist shortage and surplus. Am J Roentgenol 194 (3):697–703
74. Stokes IAF, Bigalow LC, Moreland MS (1987) Three-dimensional spinal curvature in idiopathic scoliosis. J Orthop Res 5(1):102–113
75. Summers RM (2003) Road maps for advancement of radiologic computer-aided detection in the 21st century. Radiology 229(1):11–13
76. Summers RM (2012) Evaluation of computer-aided detection devices: consensus is developing. Acad Radiol 19(4):377–379
77. Sunshine JH, Maynard CD, Paros J, Forman HP (2004) Update on the diagnostic radiologist shortage. Am J Roentgenol 182(2):301–305

78. Štern D, Likar B, Pernuš F, Vrtovec T (2010) Automated detection of spinal centrelines, vertebral bodies and intervertebral discs in CT and MR images of lumbar spine. Phys Med Biol 55(1):247–264
79. Štern D, Likar B, Pernuš F, Vrtovec T (2011) Parametric modelling and segmentation of vertebral bodies in 3D CT and MR spine images. Phys Med Biol 56(23):7505–7522
80. Štern D, Njagulj V, Likar B, Pernuš F, Vrtovec T (2013) Quantitative vertebral morphometry based on parametric modeling of vertebral bodies in 3D. Osteoporos Int 24(4):1357–1368
81. Tins B (2010) Technical aspects of CT imaging of the spine. Insights Imaging 1(5–6):349–359
82. Verdonck B, Nijlunsing R, Gerritsen FA, Cheung J, Wever DJ, Veldhuizen A, Devillers S, Makram-Ebeid S (1998) Computer assisted quantitative analysis of deformities of the human spine. In: Wells WM, Colchester A, Delp S (eds) Proceedings of the 1st international conference on medical image computing and computer-assisted intervention—MICCAI'98, 11–13 Oct 1998. Lecture notes in computer science (LNCS), vol 1496. Springer, Cambridge, pp 822–831
83. Vrtovec T, Likar B, Pernuš F (2005) Automated curved planar reformation of 3D spine images. Phys Med Biol 50(19):4527–4540
84. Vrtovec T, Likar B, Pernuš F (2005) Spine-based coordinate system. In: Zhang YT, Xu LX, Roux C, Zhuang TG, Tamura T, Galiana H (eds) Proceedings of the 27th IEEE annual international conference of the engineering in medicine and biology society (EMBS)—EMBC 2005, 1–4 Sept 2005. IEEE, Shanghai, pp 5120–5123
85. Vrtovec T, Likar B, Pernuš F (2008) Quantitative analysis of spinal curvature in 3D: application to CT images of normal spine. Phys Med Biol 53(7):1895–1908
86. Vrtovec T, Ourselin S, Lavier G, Likar B, Pernuš F (2007) Automated generation of curved planar reformations from MR images of the spine. Phys Med Biol 52(10):2865–2878
87. Vrtovec T, Pernuš F, Likar B (2008) A symmetry-based method for the determination of vertebral rotation in 3D. In: Metaxas D, Axel L, Davies B (eds) Proceedings of the 11th medical image computing and computer-assisted intervention—MICCAI 2008, 6–10 Sept 2008. Lecture notes in computer science (LNCS), vol 5241. Springer, New York, pp 942–950
88. Vrtovec T, Pernuš F, Likar B (2009) A review of methods for quantitative evaluation of axial vertebral rotation. Eur Spine J 18(8):1079–1090
89. Vrtovec T, Pernuš F, Likar B (2009) A review of methods for quantitative evaluation of spinal curvature. Eur Spine J 18(5):593–607
90. Vrtovec T, Pernuš F, Likar B (2010) Determination of axial vertebral rotation in MR images: comparison of four manual and a computerized method. Eur Spine J 19(5):774–781
91. Vrtovec T, Vengust R, Likar B, Pernuš F (2010) Analysis of four manual and a computerized method for measuring axial vertebral rotation in computed tomography images. Spine 35(12): E535–E541
92. Wan M, Liang Z, Ke Q, Hong L, Bitter I, Kaufman A (2002) Automatic centerline extraction for virtual colonoscopy. IEEE Trans Med Imaging 21(12):1450–1460
93. Weber H, Gallichan D, Schultz G, Cocosco CA, Littin S, Reichardt W, Welz A, Witschey W, Hennig J, Zaitsev M (2013) Excitation and geometrically matched local encoding of curved slices. Magn Reson Med 69(5):1317–1325
94. Weber H, Haas M, Kokorin D, Gallichan D, Hennig J, Zaitsev M (2014) Local shape adaptation for curved slice selection. Mag Reson Med 72(1):112–123
95. Weiss KL, Storrs JM, Banto RB (2006) Automated spine survey iterative scan technique. Radiology 239(1):255–262
96. Yao J, O'Connor SD, Summers RM (2006) Automated spinal column extraction and partitioning. In: Proceedings of the 3rd IEEE international symposium on biomedical imaging: from nano to macro—ISBI 2006, 6–9 Apr 2006. IEEE, Arlington, pp 390–393
97. Yushkevich PA, Zhang H, Gee JC (2006) Continuous medial representation for anatomical structures. IEEE Trans Med Imaging 25(12):1547–1564

Cross-Modality Vertebrae Localization and Labeling Using Learning-Based Approaches

Yiqiang Zhan, Bing Jian, Dewan Maneesh and Xiang Sean Zhou

Abstract Spine is one of the major organs in human body. It consists of multiple vertebrae and inter-vertebral discs. As the locations and labels of vertebrae provide a vertical reference framework to different organs in the torso, they play an important role in various neurological, orthopaedic and oncological studies. On the other hand, however, manual localization and labeling of vertebrae is often time consuming. Therefore, automatic vertebrae localization and labeling has drawn significant attentions in the community of medical image analysis. While some pioneer studies aim to localize and label vertebrae using domain knowledge, more recent studies tackle this problem via machine learning technologies. With the spirit of "data-driven", learning-based approaches are able to extract the appearance and geometric characteristics of vertebrae more efficient and effective than hand-crafted algorithms. More importantly, it facilitates cross-modality vertebrae localization, i.e., a generic algorithm working on different imaging modalities. In this chapter, we start with a review of several representative learning-based vertebrae localization and labeling methods. The key ideas of these methods are re-visited. In order to achieve a solution that is robust to severe diseases (e.g., scoliosis) and imaging artifacts (e.g., metal artifacts), we propose a learning-based method with two novel components. *First*, instead of treating vertebrae/discs as either repetitive components or completely independent entities, we emulate a radiologist and use a *hierarchial* strategy to learn detectors dedicated to *anchor* (distinctive) vertebrae, *bundle* (non-distinctive) vertebrae and inter-vertebral *discs*, respectively. At runtime, anchor vertebrae are detected concurrently to provide redundant and distributed appearance cues robust to local imaging artifacts. Bundle vertebrae detectors provide candidates of vertebrae with subtle appearance differences, whose labels are mutually determined by anchor vertebrae to gain additional robustness. Disc locations are derived from a cloud of responses from disc detectors, which is robust to sporadic voxel-level errors. *Second*, owing to the non-rigidness of spine anatomies, we employ a *local articulated* model to effectively model the spatial relations across vertebrae and discs. The local articulated model fuses appearance

Y. Zhan (✉) · B. Jian · D. Maneesh · X.S. Zhou
Siemens Medical Solutions USA, Inc., Malvern, PA 19355, USA
e-mail: yiqiang@gmail.com

© Springer International Publishing Switzerland 2015

S. Li and J. Yao (eds.), *Spinal Imaging and Image Analysis*,
Lecture Notes in Computational Vision and Biomechanics 18,
DOI 10.1007/978-3-319-12508-4_9

cues from different detectors in a way that is robust to abnormal spine geometry caused by severe diseases. Our method is validated on a large scale of CT (189) and MR (300) spine scans. It exhibits robust performance, especially to cases with severe diseases and imaging artifacts.

1 Introduction

Spine is one of the major organs in the human body. It includes vertebral column and spinal cord. A human vertebral column typically consists of 33 vertebrae. 24 of them are articulating (7 cervical, 12 thoracic and 5 lumbar vertebrae) and 9 of them are fused vertebrae in the sacrum and the coccyx. As spine strongly correlates to both neural and skeletal systems, various neurological, orthopaedic and oncological studies involve the investigations of spine anatomies. In addition, due to the strong spatial correlations between specific vertebrae and their surrounding organs, spine may also be used as a vertical reference framework to describe the locations of other organs in the trunk, e.g., transpyloric plane. Thereby, spine becomes one of the most frequently targeted anatomies in the interpretation of medical images.

In spine image analysis, localization and labeling of vertebrae is often the first step, which is tedious and time consuming for manual operators. This task becomes even more challenging for disease patients. For example, since vertebrae of a strong scoliotic spine may not be simultaneously visible in any single coronal and sagittal slice, a manual operators have to navigate multiple slices back and forth before localizing and labelling all vertebrae correctly. Accordingly, an automatic spine detection algorithm, i.e., localization and labeling of vertebrae and inter-vertebral discs, becomes an interesting research topic. A robust spine detection algorithm will not only benefit various clinical applications but also paves the way to other medical image analysis tasks, e.g., body part identification and cross-modality registration, etc.

2 Literature Review

The investigation of automatic spine detection can be traced back to the 1980s [1]. The study conducted by Chwialkowski et al. [1] aimed to detect lumbar spine discs in 2D MR slices. Prewitt edge detectors are employed to extract morphological information from the raw images. With the development of medical image analysis technologies, various algorithms have been applied on vertebrae or inter-vertebrae discs detection. In [2], Inesta et al. investigated the feasibility of identifying vertebrae levels using artificial neuron network. Active shape model was employed by Smyth et al. [3] to locate and measure vertebrae shapes in dual energy X-ray absorptiometry. In order to analyze intervertebral kinematics, Bifulco et al. [4]

proposed to recognize vertebral landmarks by comparing vertebral features in two adjacent fluoroscopic frames. Booth et al. [5] designed a system to construct a 3-D spinal image from axial MRI cross-sections. Key technologies in this system include vertebrae detection using anatomical symmetry estimation followed by a deformable template. Peng et al. [6] proposed to detect inter-vertebral discs in 2D MR slices with a polynomial function-based template followed by local postprocessing. Zheng et al. [7] combined congruency and hough transform to localize lumbar vertebrae in digital videofluoroscopy (DVF). Deschenes et al. [8] proposed to segment vertebrae in digital radiographs using multi-resolution wavelets. Naegel [9] proposed to segment the spine in CT images using mathematical morphology. Specifically, it consists of finding markers inside the vertebrae and computing the watershed from markers.

Compared to other human anatomies, spinal structures has some unique attributes, e.g., the repetitive appearance patterns of vertebrae and a characteristic geometry of spinal cord. To leverage these attributes, researchers start to use model-based approaches for spine detection. In the comprehensive spinal column extraction system proposed by Yao et al. [10], a four-part vertebra model is designed to separate vertebral region with surrounding anatomies in CT images. Alomari et al. [11] proposed a vertebrae labeling method for 2D lumbar MR scans. A two-level probabilistic model is designed to incorporate pixel-level (appearance) and object-level (geometrical) priors. Klinder et al. [12] proposed a method to detect and identify vertebrae in CT images, in which a set of models are constructed to encode shape, gradient and appearance priors.

With the success of machine learning technologies in medical imaging applications, learning-based approaches gain more attentions in spine detection as well. Schmidt et al. [13] proposed one of the first 3D MR whole spine detection methods. In their method, local appearance cues are learned by random trees. Vertebrae are localized by combining the responses of the random trees with non-local geometrical priors modeled by a parts-based graphical model. In Ma et al.'s [14] work for thoracic vertebrae identification in CT images, a discriminative classifier is trained to detect vertebrae edges and the shape of thoracic vertebrae are learned to identify their labels. In Kelm et al.'s [15] method, intervertebral disc detection in MR images is formulated as a classification problem in a nine dimensional transformation spaces. Iterative marginal space learning is proposed to generate candidates comprising position, orientation, and scale of the discs, which are further pruned by an anatomical network. Huang et al. [16] proposed a statistical learning approach based on AdaBoost algorithm to detect vertebrae centers in MR images. The detected locations are further refined to fit a spine curve. Glocker et al.'s [17] work aims to detect vertebrae in CT images using regression forest. The visible part of the spine are first roughly detected by a trained regression forest. Accurate localization and identification of individual vertebrae is then obtained through a generative shape and appearance model. The robustness of spine detection to pathological cases are further improved by using a discriminative centroid classifier using local and contextual features [18]. Major et al. [19] proposed an algorithm to label spine in both full and partial body CT scans. They employed probabilistic boosting trees

to detect spinal canal, intervertebral disks and three reference regions for landmark initialization. Final landmarks and labels are selected by Markov Random Field-based matches of a 3-disk models. Wu et al. [20] proposed a method to distinguish thoracic and lumbar vertebrae in CT images. To exploit the local context information, e.g., if a rib is attached to the vertebra, a dictionary-based classification method is designed. Specifically, a cascade of simultaneous orthogonal matching pursuit (SOMP) classifiers are applied on 2D vertebral regions extracted from the maximum intensity projection (MIP) images.

In general, spine detection algorithms are moving from heuristic rules/filters-based to learning-based approaches due to two reasons. (1) Learning-based approaches are able to effectively extract appearance/geometry/shape characteristics of spine anatomies. In learning-based approaches, spine detection is usually formulated as a classification [13, 16] or regression problem [17], in which image and geometry features are selected and combined to distinguish spine anatomies with others. Thanks to the development of machine learning technologies, learning-based algorithms are often able to find and optimally combine low level features that may not be easily designed by researchers to identify spine anatomies. (2) Learning-based approaches provide the scalability for extending the same algorithm to different imaging modalities, in which the appearance characteristics of spine may vary dramatically. Since learning-based approaches treat spine detection as a general classification/regression problem, they are purely data-driven and thus transparent to highly different image appearances. Given enough training dataset of an imaging modality, learning-based approaches are able to adapt themselves by selecting the most distinctive features from the specific imaging modalities.

In real clinical settings, patients with severe diseases may appear quite frequently and the imaging artifacts are sometimes unavoidable because of special patient conditions, e.g., metal implants (see Fig. 1). Thus, an auto-spine detection algorithm to be deployed in real clinical environment has to be highly robust to both imaging artifacts and spine diseases.

In this chapter, we introduce a spine detection method that is highly robust to severe imaging artifacts and spine diseases. In principle, our method also use machine learning technologies to capture the appearance and geometry characteristics of spine anatomies. Similar to [11, 13, 15], i.e., low-level appearance and high-level geometry information are combined to derive spine detection. In particular, our method leverages two unique characteristics of spine anatomies. First, although a spine seems to be composed by a set of repetitive components (vertebrae and discs), these components indeed have different distinctiveness. Hence, different anatomies provide different levels of reliability and should be employed *hierarchically* in spine detection. Second, spine is a *non-rigid* structure, in which local articulations exist in-between vertebrae and discs. This kind of articulation can be quite large in the presence of certain spine diseases, e.g., scoliosis. An effective geometry modeling should not consider vertebrae detections from scoliotic cases as errors just because of the abnormal geometry.

Our method includes two strategies to leverage these two characteristics effectively. First, instead of learning a general detector for vertebrae/discs, we use a

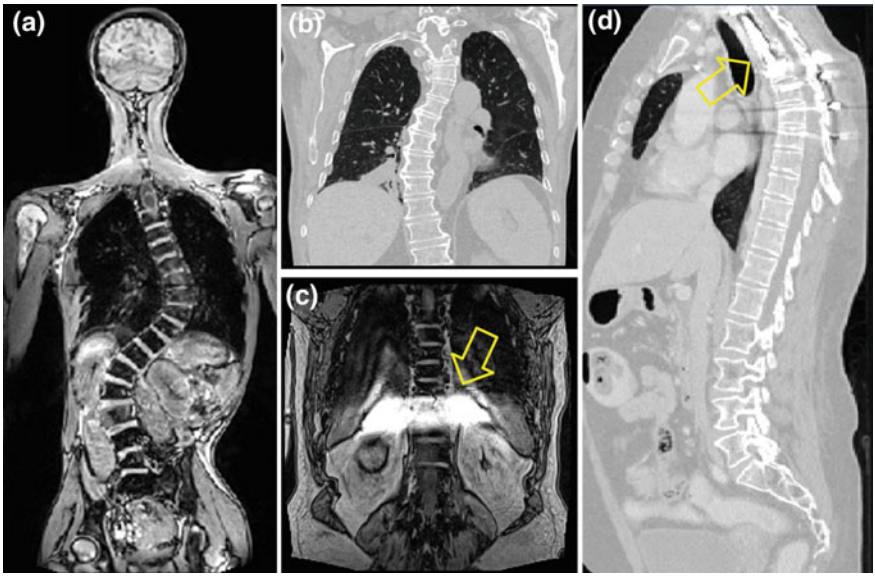


Fig. 1 Challenges of vertebrae labeling in clinical cases. **a** A MR scan of a scoliosis patient. **b** A CT scan of a scoliosis patient. **c** A MR scan with folding artifact. **d** A CT scan with metal implant

hierarchical strategy to learn detectors for anchor vertebrae, bundle vertebrae and inter-vertebral discs, respectively. Specifically, different learning strategies are designed to learn anchor, bundle and disc detectors considering different levels of "distinctiveness" of these anatomies. At run-time, these detectors are invoked in a hierarchial way. Second, a local articulation model is designed to describe spine geometries. It is employed to fuse the responses from hierarchical detectors. As the local articulation model satisfies the intrinsic geometric characteristics of both health and disease spines, it is able to propagate information from different detectors in a way that is robust to abnormal spine geometry. With the hallmarks of *hierarchical learning* and *local articulated model*, our method becomes highly robust to severe imaging artifacts and spine diseases.

3 Problem Statement

Notations: Human spine usually consists of 24 articulated vertebrae, which can be grouped as cervical (C_1–C_7), thoracic (T_1–T_{12}) and lumbar (L_1–L_5) sections. These 24 vertebrae plus the fused sacral vertebra (S_1) are the targets of spine labeling in our study.

We define vertebrae and inter-vertebral discs as $V = \{v_i | i = 1 \ldots N\}$ and $D = \{d_i | i = 1 \ldots N - 1\}$, where v_i is the i-th vertebra and d_i is the inter-vertebral disc between the i-th and i + 1-th vertebra. Here, $v_i \in \mathbb{R}^3$ is the vertebra center and

$d_i \in \mathbb{R}^9$ includes the center, orientation and size of the disc. It is worth noting that i is not a simple index but bears anatomical definition. In this paper, without loss of generality, v_i is indexed in the order of vertebrae from head to feet, e.g., v_1, v_{24}, v_{25} represents C_1, L_5 and S_1, respectively.

Formulation: Given an image I, spine detection problem can be formulated as the maximization of a posterior probability with respect to V and D as:

$$(V^*, D^*) = \arg\max_{V,D} P(V, D|I) \tag{1}$$

Certain vertebrae that appear either at the extremity of the entire vertebrae column, e.g., C_2, S_1, or at the transition regions of different vertebral sections, e.g., L_1, have much better distinguishable characteristics (red ones in Fig. 2a). The identification of these vertebrae helps in the labeling of others, and are defined as "*anchor vertebrae*". The remaining vertebrae (blue ones in Fig. 2a) are grouped into a set of continuous "bundles" and hence defined as "*bundle vertebrae*". Vertebrae characteristics are different across bundles but similar within a bundle, e.g., C_3–C_7 look similar but are very distinguishable from T_8–T_{12}.

Denoting $V_\mathcal{A}$ and $V_\mathcal{B}$ as anchor and bundle vertebrae, the posterior in Eq. (1) can be rewritten and further expanded as:

$$P(V, D|I) = P(V_\mathcal{A}, V_\mathcal{B}, D|I) = P(V_\mathcal{A}|I) \cdot P(V_\mathcal{B}|V_\mathcal{A}, I) \cdot P(D|V_\mathcal{A}, V_\mathcal{B}, I) \tag{2}$$

In this study, we use Gibbs distributions to model the probabilities. The logarithm Eq. (2) can be then derived as Eq. (3).

$$\begin{aligned} \log[P(V, D|I)] = \quad & A_1(V_\mathcal{A}|I) && \Leftarrow P(V_\mathcal{A}|I) \\ & + A_2(V_\mathcal{B}|I) + S_1(V_\mathcal{B}|V_\mathcal{A}) && \Leftarrow P(V_\mathcal{B}|V_\mathcal{A}, I) \\ & + A_3(D|I) + S_2(D|V_\mathcal{A}, V_\mathcal{B}) && \Leftarrow P(D|V_\mathcal{A}, V_\mathcal{B}, I) \end{aligned} \tag{3}$$

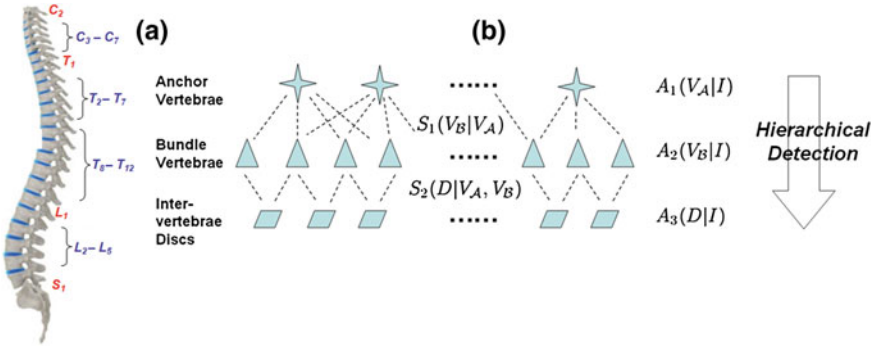


Fig. 2 **a** Schematic explanation of anchor (*red*) and bundle (*blue*) vertebrae. **b** Proposed spine detection framework

Here, A_1, A_2 and A_3 relate to the appearance characteristics of anchor, bundle vertebrae and inter-vertebral discs. S_1 and S_2 describe the spatial relations of anchor-bundle vertebrae and vertebrae-disc, respectively. It is worth noting that the posterior of anchor vertebrae solely depends on the appearance term, while those of bundle vertebrae and inter-vertebral discs depend on both appearance and spatial relations. This is in accordance to the intuition: while anchor vertebrae can be identified based on its distinctive appearance, bundle vertebrae and inter-vertebral discs have to be identified using both appearance characteristics and the spatial relations to anchor ones.

Figure 2b gives a schematic explanation of Eq. (3). Our framework consists of three layers of appearance models targeting to anchor, bundle vertebrae and discs. The spatial relations across different anatomies "bridge" different layers (lines in Fig. 2). Note that this framework is completely different from the two-level model of [11], which separates pixel- and object-level information. Instead, different layers of our framework target to anatomies with different appearance distinctiveness.

4 Hierarchical Learning Framework

4.1 Learning-Based Anatomy Detection

Before detailing hierarchical learning framework, we first introduce the basic learning modules for anatomy detection. Due to the complex appearance of vertebrae and discs, particularly in MR images, we resort to learning-based approach to model the appearance characteristics of vertebrae and inter-vertebral discs. Thanks to its data-driven nature, learning-based approaches also provide the scalability to extend our method on both CT and MR images. We formulate anatomy detection as a voxel-wise classification problem. Specifically, voxels within the anatomy primitive, i.e., vertebrae or inter-vertebral discs, are considered as positive samples and voxels away from the anatomy primitive are regarded as negative samples. To learn an anatomy detector, we first annotate vertebrae and inter-vertebral discs in training images. For each training sample (voxel), a set of elementary features are extracted in its neighborhood. Our elementary features are generated by a set of mother functions, $\{H_l(\mathbf{x})\}$ extended from Haar wavelet basis. As shown in Eq. (4) and Fig. 3, each mother function consists of one or more 3D rectangle functions with different polarities.

$$H(\mathbf{x}) = \sum_{i=1}^{N} p_i R(\mathbf{x} - \mathbf{a_i}) \tag{4}$$

where polarities $p_i = \{-1, 1\}$, $R(\mathbf{x}) = \begin{cases} 1, & \|\mathbf{x}\|_\infty \leq 1 \\ 0, & \|\mathbf{x}\|_\infty > 1 \end{cases}$ denotes rectangle functions and $\mathbf{a_i}$ is the translation.

By scaling the mother functions and convoluting them with the original image, a set of spatial-frequency spaces are constructed as Eq. (5).

$$F_l(\mathbf{x}, s) = H_l(s\mathbf{x}) * I(\mathbf{x}) \tag{5}$$

where s and l denote the scaling factor and index of mother functions, respectively.

Finally, for any voxel $\mathbf{x_0} \in \Re^3$, its feature vector $\tilde{\mathfrak{F}}(\mathbf{x_0})$ is obtained by sampling these spatial-frequency spaces in the neighborhood of $\mathbf{x_0}$ (Eq. 6). It provides cross-scale appearance descriptions of voxel $\mathbf{x_0}$.

$$\tilde{\mathfrak{F}}(\mathbf{x_0}) = \bigcup_{l=1\ldots L} \{F_l(\mathbf{x_i}, s_j) | \mathbf{x_i} \in \mathbb{N}(\mathbf{x_0}), s_{\min} < s_j < s_{\max}\} \tag{6}$$

Compared to standard Haar wavelet, the mother functions we employed are not orthogonal. However, they provide more comprehensive image features to characterize different anatomy primitives. For example, as shown in Fig. 3, mother function (a) potentially works as a smoothing filter, which is able to extract regional features. Mother functions (b) and (c) can generate horizontal or vertical "edgeness" responses, which are robust to local noises. More complicated mother function like (d) is able to detect "L-shape" patterns, which might be useful to distinguish some anatomy primitives. In addition, our features can be quickly calculated through integral image [21]. It paves the way to an efficient anatomy detection system.

All elementary features are then fed into a cascade classification framework [22] as shown in Fig. 4. The cascade framework is designed to address the highly unbalanced positive and negative samples. In fact, since only voxels around vertebrae centers or intervertebral discs are positives, the ratio of positives to negatives is often less than $1{:}10^5$. In the training stage, all positives but a small proportion of

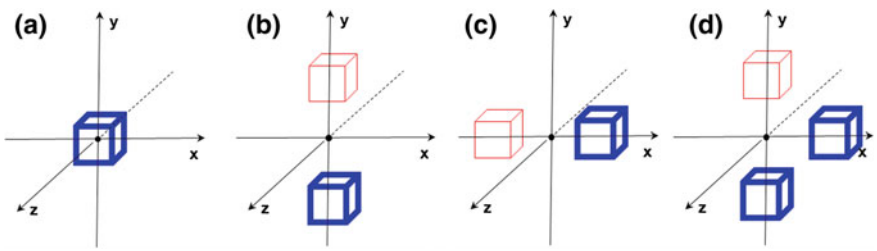


Fig. 3 Some examples of haar-based mother functions

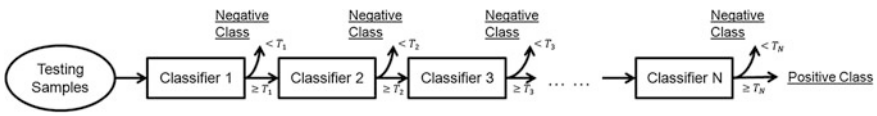


Fig. 4 Schematic explanation of cascade Adaboost classifers

negatives are used at every cascade level. The training algorithm is "biased" to positives, such that each positive have to be correctly classified but the negatives are allowed to be misclassified. These misclassified negatives i.e., False Positives, will be further trained in the following cascades. At run-time, while positives will go through all cascades, most negatives can be rejected in the first several cascades and do no need further evaluation. In this way, the run-time speed can be dramatically increased. In our study, we use Adaboost [23] as the basic classifier in the cascade framework. The output of the learned classifier $\mathcal{A}(\widetilde{\mathfrak{F}}(\mathbf{x}))$ indicates the existence likelihood of a landmark at $\mathbf{x}$.

Our learning-based framework is general to detect different anatomical structures in different imaging modalities due to two reasons. (1) The extended Haar wavelet generates thousands of features. In this large feature pool, there are always some features that are distinctive to specific anatomies. (2) The cascade learning framework is able to select the most distinctive features for a specific anatomy in a specific imaging modality.

In spine detection, a straightforward way to use this general detection framework is to train detectors for each vertebrae independently. However, these trained detectors might be confused by the similar appearances of neighboring vertebrae, particularly in the presence of diseases or imaging artifacts. An alternative way is to train a general detector to all vertebrae. However, due to the large shape and appearance variability across different vertebrae, e.g., the shape and size of cervical vertebrae are very different from lumbar vertebrae, it is very different to capture the common characteristics of all vertebrae with one detector/classifier. By observing these two limitations, we design a hierarchical learning scheme, which essentially categorize vertebrae and discs into different groups and use different training strategies based on their different characteristics. Specifically, our learning scheme consists of three layers, anchor vertebrae, bundle vertebrae and inter-vertebral discs.

4.2 Anchor Vertebrae

Anchor vertebrae (red ones in Fig. 2a) are vertebrae with distinctive characteristics. They are usually the vertebrae located at the extremes of vertebral column (e.g., C2, S1) or at the transition border of different spine sections (e.g., C7, L1). In radiology practices, anchor vertebrae usually provide critical evidences for labels of other vertebrae. To leverage the distinctive characteristics of anchor vertebrae, we build anchor vertebrae detectors as the first layer of our hierarchical learning scheme. The learning scheme is designed to achieve two goals. First, since anchor vertebrae have distinctive characteristics and can be identified exclusively, the detectors of anchor vertebrae should be very discriminative and only have high responses around the specific vertebrae centers. Second, as anchor vertebrae will be used to derive the labels of other vertebrae, the detection of anchor vertebrae should be highly robust.

To achieve the first goal, we train anchor vertebra detectors in a very discriminative way. Specifically, we only select voxels close to the specific anchor vertebra

as the positives and all others are treated as negatives. Although the numbers of positive and negative samples become highly balanced, thanks to the cascade learning framework, we can still learn a discriminative detector for each anchor vertebra. To reach the second goal, for each anchor vertebrae, we define a set of "supporting landmarks" surrounding it. For example, for S1 vertebra, we use tip of coccyx, center of sacrum and spinous process of L5, etc., as its supporting landmarks. The strong spatial correlation between the supporting landmarks and the anchor vertebra are exploited to achieve robust detection. Mathematically, we employ a linear model to capture the spatial correlation between the anchor vertebra v_i and its supporting landmarks $\mathcal{S}(v_i)$ as Eq. (7):

$$v_i = \mathfrak{C} \cdot \mathbb{U} \tag{7}$$

Here, $\mathbb{U}$ is a vector concatenated by coordinates of supporting landmarks and $\mathfrak{C}$ denotes the linear correlation matrix. Given a set of training samples, $\mathfrak{C}$ can be learned by solving a least squares problem. At run-time, besides detecting anchor vertebrae, we also detect its supporting landmarks. The learned linear correlation matrix are then used to verify the detected anchor vertebrae. In principle, we resort to "redundancy" [24] for highly robust anchor vertebrae detection.

4.3 Bundle Vertebrae

Different from anchor vertebrae, other vertebrae have less distinctive shapes and appearances. These vertebrae look similar to their neighbors but different from remote ones. On one hand, training a general detector for *all* bundle vertebrae is almost infeasible due to the large variations across distal ones. On the other hand, an attempt to learn the subtle differences between a bundle vertebra and its neighborhoods also adversely affects the robustness. For example, T_9 and T_{10} are two neighboring vertebrae with similar appearance and shape characteristics. For normal cases (see Fig. 5a), the two detectors are still possible to distinguish the subtle differences between T_9 and T_{10}. However, when the appearance of T_9 becomes abnormal due to imaging artifacts or diseases, T_9 detector might have higher responses at T_{10} than T_9 (Fig. 5b), which induces wrong/miss labeled vertebrae. In fact, this problem is also observed in [15], where "(standard) MSL approach may end up with detections for the most salient disks only".

To avoid this problem, we propose to group neighboring vertebrae as "bundles" (Fig. 2a). Vertebrae within the same bundle are treated as equivalent positives in the learning algorithm. In this way, each bundle has one detector that encodes the commonality of corresponding vertebrae and distinguishes them from other bundles. This kind of detectors are learnable as in-bundle vertebrae often have much more similar characteristics than cross-bundle ones. Compared to anchor vertebrae detectors, bundle detectors are less discriminative. Bundle detectors are not able to identify specific vertebrae, e.g., T_9 or T_{10}, but specific vertebrae bundle, e.g., T_6–T_8

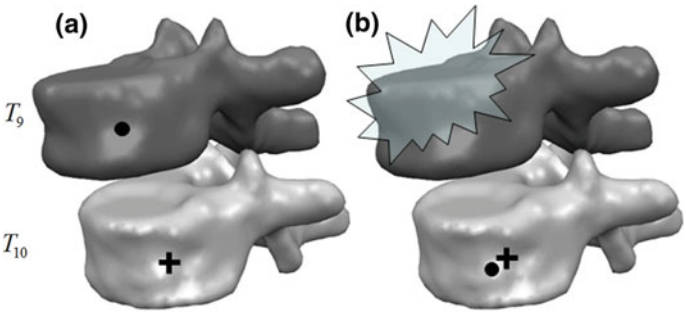


Fig. 5 An example of "over-discriminative" vertebrae detectors. "*Filled circle*" and "*Plus sign*" denote the highest response of T_9 and T_{10} detectors, respectively

or T_9–T_{11}. At run-time, bundle vertebrae detectors are expected to return multiple peak responses at vertebrae centers belong to the bundle. The locations of these responses will be further verified based on anchor vertebrae and the local articulated spine model (see Sect. 5). The specific vertebrae labels are assigned in the same way.

4.4 Inter-vertebral Discs

To detect an inter-vertebral disc, not only its center but also its orientation and size should be determined. Compared to vertebrae center detection (3-D hypothesis space), disc detection has a higher dimensional hypothesis space with 9 spatial parameters. In [15], this problem is tackled by marginal space learning (MSL), which detect the center, orientation and sequentially. In principle, MSL treats the inter-vertebral disc as a whole and aims to sequentially determine the spatial parameters through sub-space learning. The sequential nature, however, might influence the robustness of the algorithm. For example, a gross error of disc center detection may not be corrected by the following orientation/size detection.

Instead of treating the disc as a whole, we propose to formulate the disc detection as a voxel-wise classification problem. Specifically, for each voxel, the disc detector aims to predict the likelihood of this voxel belonging to the inter-vertebral disc. Therefore, in the training stage, each voxel is considered as an individual training sample. On-disc and off-disc voxels are used as positives and negatives, respectively. At runtime, the learned disc detector will derive a response map. The 9 spatial parameters of the disc is then derived by fitting disc response maps with an elliptical cylinder using principal component analysis. This strategy brings robustness in twofold. First, since each voxel is labelled independently, the classification errors of one voxel will not influence any others. Second, since principal component analysis is robust to outliers, sporadic classification errors at voxel-level will not dramatically change the derived disc positions.

It is worth noting that this learning scheme follows the same trend from anchor to bundle vertebrae. As previously discussed, neighboring vertebrae are "bundled" since they are not easy to be distinguished. Hence, the bundle vertebrae detectors are not as distinctive as anchor vertebrae detectors and have multiple peaky responses at runtime. Since voxels on the same disc are even more indistinguishable, all of them are "bundled" in the training stage. Hence, the learned detectors becomes even less distinctive and expected to have high responses at any voxels located on the disc.

To further enhance the robustness, we leverage the information from vertebrae centers in learning disc detectors. Specifically, the training images of a disc are aligned by the line connecting its two neighboring vertebrae centers. Since the inter-vertebral discs are roughly perpendicular to the line connecting its, this alignment will effectively remove the variations of the appearance features resulting from different disc orientations. (Note that Haar-like features are not rotation invariant.) At runtime, based on the previously detected vertebrae centers, the subject image is aligned in the same way before applying the disc detectors.

4.5 Summary

The differences in training anchor vertebrae, bundle vertebrae and inter-vertebral discs detectors primarily exist in the selection of positive/negative samples and image alignment before feature extraction (see Table 1). Moving down the table from anchor vertebrae to inter-vertebral discs, as the targeted anatomies become less and less distinctive, more positive samples are extracted, i.e., voxels around the center of a specific anchor vertebra → voxels around centers of several neighboring vertebrae → any voxels located at a specific inter-vertebral disc. This results in less discriminative detectors that are expected to return high responses at more voxels. On the other hand, the image alignment becomes more and more sophisticated.

Table 1 Training scheme of detectors for anchor vertebrae, bundle vertebrae and inter-vertebral discs

Detector	Positive samples	Negative samples	Image alignment
Anchor vertebrae	Voxels close to the center of the *specific* vertebrae	Remaining voxels in the *entire* volume image	No alignment
Bundle vertebrae	Voxels close to the centers of *any* vertebrae within the bundle	Remaining voxels in the *local* volume image covering neighboring bundles	Aligned by anchor vertebrae
Inter-vertebral Discs	Voxels located on the disc	Remaining voxels in the *local* volume image covering the two neighboring vertebrae	Aligned by two neighboring vertebrae

The alignment essentially removes the feature variations from spatial transformation, hence, the learning task becomes easier.

Using the above strategy, we train detectors for anchor vertebrae, bundle vertebrae and inter-vertebral discs as $\mathcal{A}_i(\mathfrak{F}(p))$, $\mathcal{B}_j(\mathfrak{F}(p))$, and $\mathcal{D}_k(\mathfrak{F}(p))$. Here, $\mathfrak{F}(p)$ denotes the over-complete Haar features extracted around voxel p, and $\mathcal{A}_i$, $\mathcal{B}_j$ and $\mathcal{D}_k$ are the trained cascade Adaboost classifiers, which select and combine a small proportion of $\mathfrak{F}(p)$ to achieve best anatomy detection. The appearance terms in Eq. (3) are eventually concretized as $A_1(V_\mathcal{A}|I) = \sum_{v_i \in V_\mathcal{A}} \mathcal{A}_i(\mathfrak{F}(v_i))$, $A_2(V_\mathcal{B}|I) = \sum_{v_j \in V_\mathcal{B}} \mathcal{B}_j(\mathfrak{F}(v_j))$ and $A_3(D|I) = \sum_{d_k \in D} \sum_{p \in d_k} \mathcal{D}_k(\mathfrak{F}(p))$.

5 Local Articulated Spine Model

Recall the definition of Eq. (3), $S_1(V_\mathcal{B}|V_\mathcal{A})$ and $S_2(D|V_\mathcal{A}, V_\mathcal{B})$ model the spatial relations between anchor-bundle vertebrae and vertebrae-discs, respectively. In our spine detection method, spatial relations are exploited in threefold. First, it determines where the detectors should be invoked. For example, after the anchor vertebrae are detected, we can predict the positions of bundle vertebrae and only invoke the bundle vertebrae detectors in these local regions. Second, spatial relations can be used to verify the detection. For example, if a detected disc is almost parallel to the line connecting its two neighboring vertebrae center, it is highly probable that either vertebrae centers or disc are erroneously detected. Third, since bundle vertebrae detectors only determines which bundle the vertebra belongs to, spatial relations should be employed to assign exact labels to bundle vertebrae. Accordingly, a proper modeling of spatial relations across vertebrae and inter-vertebral discs becomes critical to the success of spine detection.

Various methodologies, including principal component analysis [25] and sparse representation [26], have been investigated for anatomy shape/geometry modeling and achieved tremendous success. However, since spine is a flexible structure where each vertebra has freedom of local articulation (see Fig. 6). those methods that treat the object as a whole may not model the characteristics of spine geometry properly. For example, for a patient with severe scoliosis (see Fig. 1a), the spine geometry appears as an outlier in eigen space. Hence, even when the vertebrae detection is correct, they will be "mis-corrected" to follow a normal spine geometry using standard active shape model [25].

To deal with the specialty of spine geometry, local articulated model is designed in [27, 28]. In our study, we employ similar model to describe the spatial relations across vertebrae. The key idea is to decompose the spatial transformation of a spine into a set of local transformations between neighboring vertebrae. Instead of enforcing constraints on the global transformation, we constrains local transformations based on learned statistics. In addition, smoothness across neighboring local transformations is applied as another constraints. Note that this constraint still holds for patients with severe scoliosis, since the spinal cord usually forms a smoothing curve even for scoliosis patients.

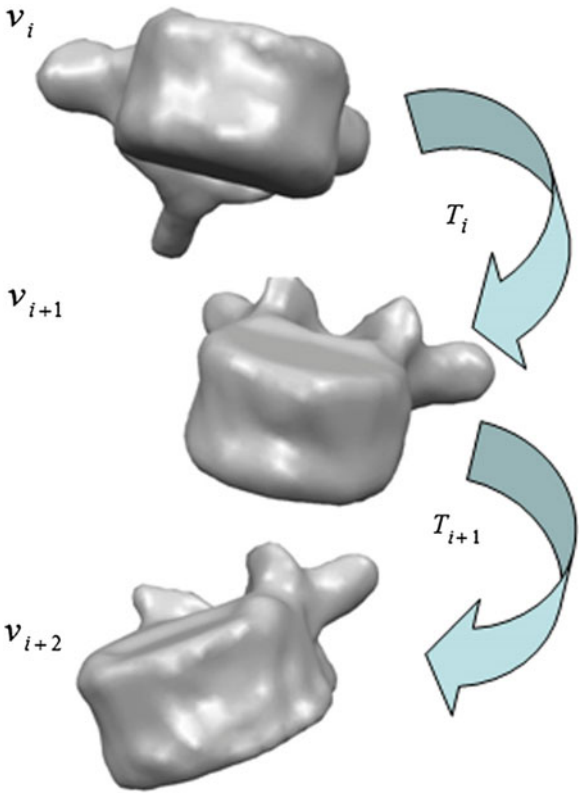


Fig. 6 Local articulation model

Assume v_i is an anchor vertebra and $\{v_{i+1}, \ldots, v_{i+M}\}$ are the subsequent bundle vertebrae. As shown in Fig. 6, the spatial relations between anchor and bundle vertebrae are modeled as $[T_i, T_i \circ T_{i+1}, \ldots, T_{i,} \circ T_{i+1} \circ \ldots \circ T_{i+M-1}]$, where T_i defines a local similarity transformation between v_i and v_{i+1}. $S_1(V_\mathcal{B}|V_\mathcal{A})$ is defined as:

$$S_1(V_\mathcal{B}|V_\mathcal{A}) = \sum_i e^{-(\psi(T_i)-\mu_{T_i})^T \Xi_{T_i} (\psi(T_i)-\mu_{T_i})} + 2/(1 + e^{\gamma \|\psi(T_i)-\psi(T_{i+1})\|^2}) \quad (8)$$

Here, $\psi(.)$ is an operator that converts T_i to a vector space, i.e., the rotation part of T_i is converted to its quaternion. μ_{T_i} and Ξ_{T_i} are the Frechet mean and generalized covariance of local transformation T_i, calculated as [27]. The first term contains the prior information of local transformations across *population*. The second term evaluates the difference between local T_i across the same *spine*. These two terms complement each other, such that a scoliotic spine still get a high value of S_1, due to the continuity of its local transformations.

Spatial configurations between vertebrae and discs, $S_2(D|V_\mathcal{A}, V_\mathcal{B})$, is modeled with two assumptions: (1) A vertebral disc is roughly perpendicular to the line connecting its neighboring vertebrae centers; and (2) Center of a vertebral disc is

close to the mid point of the two neighboring vertebrae centers. $S_2(D|V_{\mathcal{A}}, V_{\mathcal{B}})$ is then defined as:

$$S_2(D|V_{\mathcal{A}}, V_{\mathcal{B}}) = \sum_i [e^{-(1-U(v_i - v_{i+1})\cdot D_{d_i})^2/\lambda_1^2} + e^{-\|(\frac{v_i+v_{i+1}}{2}) - C_{d_i}\|^2/\lambda_2^2}] \quad (9)$$

where v_i and v_{i+1} denote the centers of the two neighboring vertebrae of disc d_i, whose center and norm are D_{d_i} and C_{d_i}. $U(.)$ is the normalization operator. The first and second terms of Eq. (9) in fact reflects the two aforementioned spatial constrains between neighboring vertebrae and discs, respectively. The modeling of spatial correlations between vertebrae and discs allow our system to propagate information of vertebrae layer to disc layer for robust detection.

6 Hierarchical Spine Detection

Based on the descriptions in Sects. 4 and 5, we have all terms in Eq. (3) defined. At runtime, spine detection becomes an optimization procedure of Eq. (3). As Eq. (3) is a high-dimensional and non-linear function, we design a multi-stage algorithm to optimize it.

Different stages target to anchor vertebrae, bundle vertebrae and inter-vertebral discs, respectively. In each stage, we alternatively optimize the appearance terms and spatial terms. More specifically, optimization starts from the concurrent detection of anchor vertebrae, which is equivalent to the maximization of A_1. Based on the detected anchor vertebrae, S_1 is maximized to predict the positions of subsequent bundle vertebrae. It determines the local regions where bundle vertebrae detectors will be invoked to maximize A_2. S_1 is then further maximized. In this step, the responses of bundle vertebrae detectors are verified and the exact spine labels are assigned. Subsequently, A_3 and S_2 are optimized in the same fashion.

Figure 2a gives a schematic explanation of the optimization procedure. This hierarchial detection scheme emulates a manual operator and achieves the robustness in three aspects: (1) Anchor vertebrae are detected *concurrently* to provide redundant and distributed appearance cues. With the redundant detection of supporting landmarks, the detection of anchor vertebrae is very robust. Even when some anchor vertebrae are missed due to severe local imaging artifacts, others still provide reliable clues for spine detection. (2) Detectors of bundle vertebrae and discs provide supporting cues. More specifically, instead of trying to directly derive vertebrae labels, bundle vertebrae detectors provide a set of candidates whose labels are *mutually* assigned according to relative positions to anchor vertebrae. Note that labels assigned by different anchor vertebrae might be different, and are fused through the maximization of S_1. Disc detectors return a cloud of responses for disc localization, which is robust to individual false classifications as well. (3) Local articulated model propagates these appearance cues in a way robust to abnormal spine geometry resulting from severe diseases.

7 Results

7.1 Data Descriptions

We evaluated our method on both CT and MR datasets. Our CT data includes 189 randomly selected CT cases with partial or whole spine coverage. A wide range of imaging parameters are used for these dataset. For example, slice thickness ranges from 0.75 to 20 mm and different reconstruction kernel (B30s, B31f, B35f, B60f, H20f, H31s, H70h) are used. Our MR data includes 300 T1-weighted 3D MR scout scans. These scout scans have relatively low but isotropic resolution 1.7 mm and large fields of view. It is not designed for diagnostic purpose but for vertebrae and disc detections. Our MR data also covers partial or whole spines. These datasets come from different clinical sites and were generated by different types of Siemens MR Scanners (Avanto 1.5T, Verio 3T, Skyra 3T, etc.).

7.2 Evaluations

The automatic spine detection results are shown to experienced radiologists and rated as "perfect" (no manual editing required), "acceptable" (minor manual editing required) and "rejected" (major manual editing required).

As shown in Table 2, our method reaches "perfect" detection in 95 + % cases. It is worth noting that our testing set includes CT/MR scans with severe diseases or imaging artifacts. However, our method can still detect vertebrae and discs robustly. Figures 7 and 8 show the results of our method on challenging CT and MR scans. As shown in these cases, our method is robust to different kinds of imaging artifacts (Figs. 7a, b and 8c, g), large imaging noises (Fig. 7f), metal implants (Figs. 7d, e and 8d, f), severe scoliosis (Fig. 8a), pathologies (Figs. 7c and 8e) congenital abnormality (Fig. 8b) and scans that have anchor vertebrae out of field of view (Fig. 8h).

We also conduct quantitative evaluation on 355 discs and 340 vertebrae from 15 WholeSpine MR scans. The average translation errors of discs and vertebrae are 1.91 and 3.07 mm. The average rotation errors of discs is 2.33°.

Table 2 Evaluations of spine detections in CT and MR scans

	Number of cases	Perfect	Acceptable	Reject
CT	189	180 (95.2 %)	6 (3.2 %)	3 (1.6 %)
MR	300	293 (97.7 %)	4 (1.3 %)	3 (1.0 %)

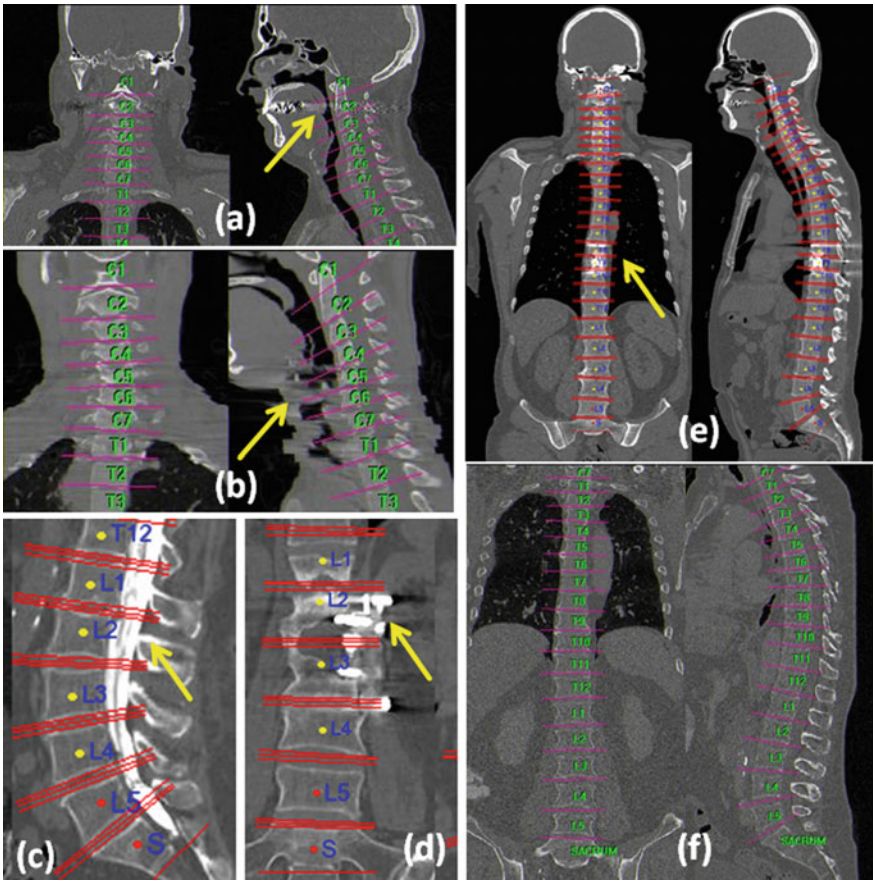


Fig. 7 Examples of spine detection in challenging CT scans. **a** C-Spine scan with metal artifacts. **b** C-Spine scan with motion artifacts. **c** L-Spine scan with spinal cord disease. **d** L-Spine scan with metal implant. **e** Whole spine scan with metal artifacts. **f** Whole spine scan with large imaging noises

7.3 Comparisons

To illustrate the importance of hierarchical learning and the local articulation model, we also evaluate results from two adapted versions of the proposed method. In *Method1*, we take out the hierarchical learning part. Specifically, dedicated detectors are trained for each vertebra and inter-vertebral disc. In *Method2*, we take out local articulated model and use the standard PCA-based method to model the spine geometry (Table 3).

In Table 3, we list the qualitative results of these three methods on 300 MR scout scans. The proposed method generates "perfect" results in more than 97 % cases, which is significantly better than the others. In general, Method2 is better than

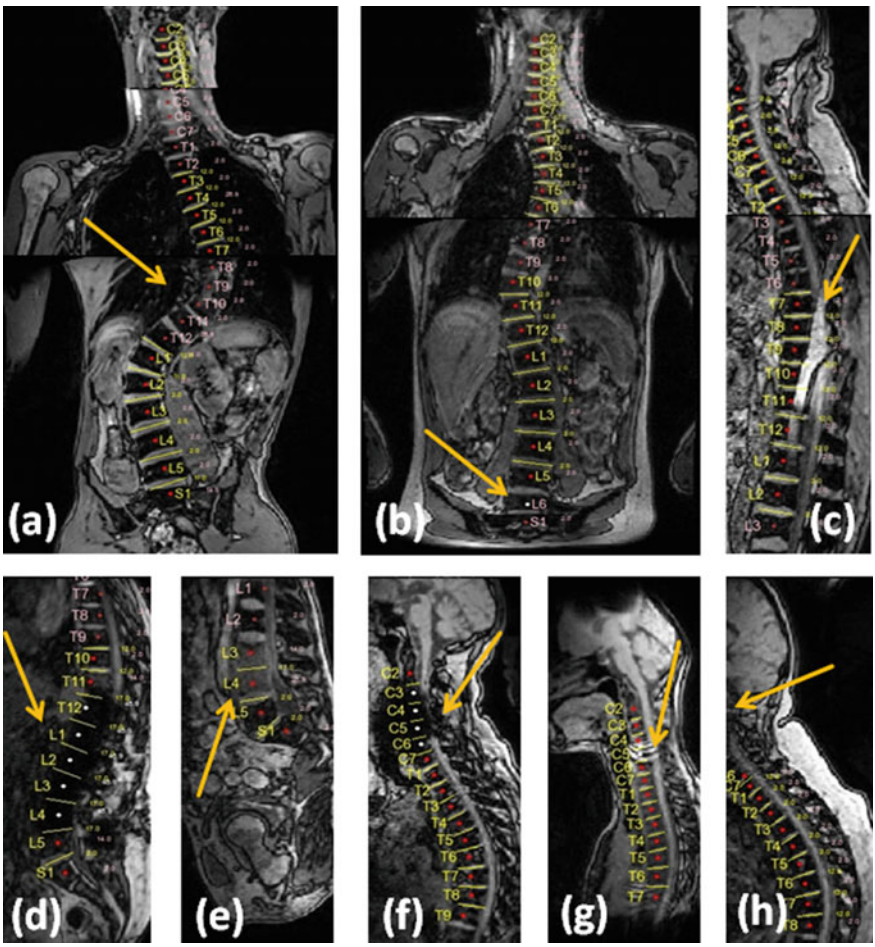


Fig. 8 Examples of spine detection in challenging MR scans. **a** Whole-spine scan with strong scoliosis. **b** Whole-spine scan with congenital abnormality (6 lumbar vertebra). **c** Whole-spine scan with folding artifact. **d** L-spine scan with metal implant. **e** L-spine scan with vertebra pathology. **f** C-spine scan with metal artifact. **g** C-spine scan with ring artifact. **h** C-spine scan where anchor vertebra is out of field of view

Method1, since the lack of articulated model mainly affects scoliosis cases, which has a small proportion in clinically representative dataset. Another interesting observation is that Method1 has larger impacts on CSpine than LSpine, but Method2 is in the other way around. This phenomenon in fact results from the different sizes of cervical and lumbar vertebrae. Due to the smaller size of cervical vertebrae, it is prone to erroneously detection using non-hierarchical detectors. On the other hand, the larger size of lumbar vertebrae L-Spine makes the detection more sensitive to abnormal spine geometry, that only can be tackled with the local articulated model. Two representative failure cases of Method1 and Method2 are

Table 3 Evaluations of spine detections in 3D MR scout scans

	Without hierarchical learning			Without articulated model			Proposed method		
	Perfect	Accept	Reject	Perfect	Accept	Reject	Perfect	Accept	Reject
LS	85 (85.0 %)	1 (10.0 %)	5 (5.0 %)	93 (93.0 %)	5 (5.0 %)	2 (2.0 %)	98 (98.0 %)	2 (2.0 %)	0 (0.0 %)
CS	65 (81.2 %)	11 (13.8 %)	4 (5.0 %)	76 (95.0 %)	3 (3.7 %)	1 (1.2 %)	79 (98.8 %)	0 (0.0 %)	1 (1.2 %)
WS	97 (80.8 %)	16 (13.3 %)	7 (5.8 %)	106 (88.3 %)	8 (6.7 %)	6 (5.0 %)	116 (96.7 %)	2 (1.6 %)	2 (1.6 %)
All	247 (82.4 %)	37 (12.3 %)	16 (5.3 %)	275 (91.7 %)	16 (5.3 %)	9 (3.0 %)	293 (97.7 %)	4 (1.3 %)	3 (1.0 %)

LS L-Spine, *CS* C-Spine, *WS* Whole-Spine

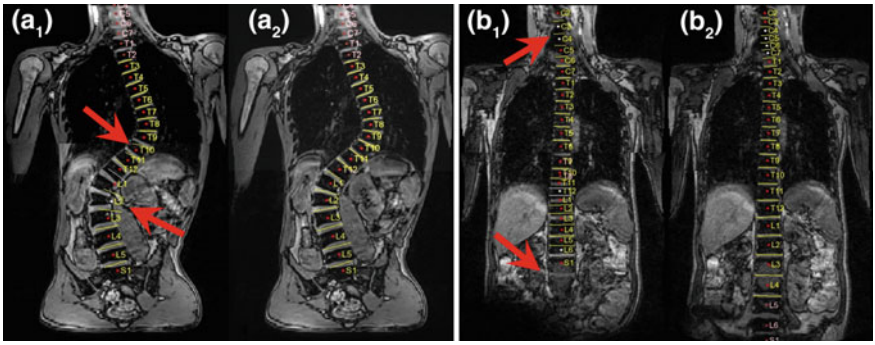


Fig. 9 Comparisons of spine detection using different methods. **a** A scoliotic case using Method2 (*a1*) and the proposed method (*a2*). **b** An artifact case using Method1 (*b1*) and the proposed method (*b2*)

shown in Fig. 9. As shown in Fig. 9a, the detection results of a scoliotic patient is wrong if the spine geometry is modeled with the standard PCA-based approach. In Fig. 9b, the specifically trained detectors are "confused" by the imaging artifacts and pathology in both C-spine and L-spine regions. Hence, the vertebrae labels are wrong. In contrast, our method can robustly detect spine in both scenarios.

8 Conclusions

In this chapter, we introduced a method to automatically detect and label vertebrae and inter-vertebral discs in medical images. By employing learning-based technologies, our method is generic for different imaging modalities. In order to achieve a solution that is robust to severe diseases (e.g., scoliosis) and imaging artifacts (e.g., metal artifacts), two novel components are designed in our system. *First*, we emulate a radiologist and use a *hierarchial* strategy to learn detectors dedicated to *anchor* (distinctive) vertebrae, *bundle* (non-distinctive) vertebrae and inter-vertebral *discs*, respectively. At run-time, anchor vertebrae are detected concurrently to provide redundant and distributed appearance cues robust to local imaging artifacts. Bundle vertebrae detectors provide candidates of vertebrae with subtle appearance differences, whose labels are mutually determined by anchor vertebrae to gain additional robustness. Disc locations are derived from a cloud of responses from disc detectors, which is robust to sporadic voxel-level errors. *Second*, owing to the non-rigidness of spine anatomies, we employ a *local articulated* model to effectively model the spatial relations across vertebrae and discs. The local articulated model fuses appearance cues from different detectors in a way that is robust to abnormal spine geometry resulting from severe diseases.

We tested our method on a large scale of CT (189) and MR (300) spine scans. Verified by experienced radiologists, our method reaches "perfect" labeling

in 95 %+ cases. In particular, our method exhibits robust performance, especially to cases with severe diseases and imaging artifacts. This method opens the door to improve the speed and quality of spine imaging workflow. It can also benefits various spine applications, e.g., quantitative measurements of spine geometry for scoliosis diagnosis.

It is worth noting that the validation of the chapter is conducted on an in-house dataset. Recently, a public dataset becomes available at *SpineWeb* http://spineweb.digitalimaginggroup.ca. This public dataset opens the window to compare different methods in a much more fair way. Interested readers may test different methods or develop new ones on this public dataset.

References

1. Chwialkowski M, Shile P, Peshock R, Pfeifer D, Parkey R (1989) Automated detection and evaluation of lumbar discs in MR images. In: IEEE EMBS, pp 571–572
2. Inesta J, Sarti M, Buendia M (2005) On the possibility of objective identification of human vertebrae through pattern recognition algorithms. In: 5th international conference on image processing and its applications
3. Smyth PP, Taylor CJ, Adams JE (1997) Automatic measurement of vertebral shape using active shape models. Image Vis Comput 15(8):575–581
4. Bifulco P, Cesarelli M, Allen R, Sansone M, Bracale M (2001) Automatic recognition of vertebral landmarks in fluoroscopic sequences for analysis of intervertebral kinematics. Med Biol Eng Compu 39(1):65–75
5. Booth S, Clausi DA (2001) Image segmentation using MRI vertebral cross-sections. Canadian conference on electrical and computer engineering, vol 2, IEEE, pp 1303–1307
6. Peng Z, Zhong J, Wee W, Lee J (2005) Automated vertebra detection and segmentation from the whole spine MR images. In: IEEE EMBS, pp 2527–2530
7. Zheng Y, Nixon MS, Allen R (2001) Automatic lumbar vertebrae segmentation in fluoroscopic images via optimised concurrent hough transform. In: Proceedings of the 23rd annual international conference of the IEEE Engineering in Medicine and Biology Society, vol 3, IEEE, pp 2653–2656
8. Deschênes S, De Guise J (2002) Wavelet-based automatic segmentation of the vertebral bodies in digital radiographs. In: 2002 IEEE international conference on acoustics, speech, and signal processing (ICASSP), vol 4, IEEE, IV–3868
9. Naegel B (2007) Using mathematical morphology for the anatomical labeling of vertebrae from 3D CT-scan images. Comput Med Imaging Graph 31(3):141–156
10. Yao J, O'Connor SD, Summers RM (2006) Automated spinal column extraction and partitioning. In: 3rd IEEE international symposium on biomedical imaging: nano to macro, IEEE, pp 390–393
11. Alomari R, Corso J, Chaudhary V (2011) Labeling of lumbar discs using both pixel- and object-level features with a two-level probabilistic model. IEEE Trans Med Imaging 30:1–10
12. Klinder T, Ostermann J, Ehm M, Franz A, Kneser R, Lorenz C (2009) Automated model-based vertebra detection, identification, and segmentation in CT images. Med Image Anal 13:471–482
13. Schmidt S, Kappes J, Bergtholdt M, Pekar V, Dries S, Bystrov D, Schnörr C (2007) Spine detection and labeling using a parts-based graphical model. In: IPMI, pp 122–133
14. Ma J, Lu L, Zhan Y, Zhou X, Salganicoff M, Krishnan A (2010) Hierarchical segmentation and identification of thoracic vertebra using learning-based edge detection and coarse-to-fine

deformable model. In: Medical image computing and computer-assisted intervention—MICCAI 2010, pp 19–27
15. Kelm M, Zhou S, Sühling M, Zheng Y, Wels M, Comaniciu D (2010) Detection of 3D spinal geometry using iterated marginal space learning. In: MCV, pp 96–105
16. Huang SH, Chu YH, Lai SH, Novak CL (2009) Learning-based vertebra detection and iterative normalized-cut segmentation for spinal MRI. IEEE Trans Med Imaging 28 (10):1595–1605
17. Glocker B, Feulner J, Criminisi A, Haynor DR, Konukoglu E (2012) Automatic localization and identification of vertebrae in arbitrary field-of-view CT scans. In: Medical image computing and computer-assisted intervention—MICCAI 2012. Springer, Berlin, pp 590–598
18. Glocker B, Zikic D, Konukoglu E, Haynor DR, Criminisi A (2013) Vertebrae localization in pathological spine CT via dense classification from sparse annotations. In: Medical image computing and computer-assisted intervention—MICCAI 2013. Springer, Berlin, pp 262–270
19. Major D, Hladuvka J, Schulze F, Buhler K (2013) Automated landmarking and labeling of fully and partially scanned spinal columns in CT images. Med Image Anal 17(8):1151–1163
20. Wu T, Jian B, Zhou XS (2011) Automated identification of thoracolumbar vertebrae using orthogonal matching pursuit. In: Machine learning in medical imaging. Springer, Berlin, pp 126–133
21. Crow F (1984) Summer-area tables for texture mapping. In: SIGGRAPH
22. Viola P, Jones MJ (2004) Robust real-time face detection. Int J Comput Vision 57:137–154
23. Freund Y, Schapire RE (1997) A decision-theoretic generalization of on-line learning and an application to boosting. J Comput Syst Sci 55:119–139
24. Zhou XS, Peng Z, Zhan Y, Dewan M, Jian B, Krishnan A, Tao Y, Harder M, Grosskopf S, Feuerlein U (2010) Redundancy, redundancy, redundancy: the three keys to highly robust anatomical parsing in medical images. In: Proceedings of the international conference on multimedia information retrieval, ACM, pp 175–184
25. Cootes TF, Taylor CJ, Cooper DH, Graham J (1995) Active shape models-their training and application. Comput Vis Image Underst 61(1):38–59
26. Zhang S, Zhan Y, Dewan M, Huang J, Metaxas DN, Zhou XS (2012) Towards robust and effective shape modeling: sparse shape composition. Med Image Anal 16:265–277
27. Boisvert J, Cheriet F, Pennec X, Labelle H, Ayache N (2008) Geometric variability of the scoliotic spine using statistics on articulated shape models. IEEE Trans Med Imaging 27:557–568
28. Kadoury S, Labelle H, Paragios N (2011) Automatic inference of articulated spine models in CT images using high-order markov random fields. Med Image Anal 15:426–437

Articulated Statistical Shape Models of the Spine

Jonathan Boisvert

Abstract The spine is a complex assembly of rigid vertebrae surrounded by various soft tissues (ligaments, spinal cord, intervertebral discs, etc.). Its motion for a given individual and its shape variations across a population are greatly influenced by this fact. We show in this chapter how statistical shape models can be constructed, used, and analyzed while taking into account the articulated nature of the spine. We begin by defining what articulated models are and how they can be extracted from existing 3D reconstructions or segmented models. As an example, we use data from scoliotic patients that have been reconstructed in 3D using bi-planar radiographs. Articulated models naturally belong to a manifold where conventional statistical tools are not applicable. In this context, a few key concepts allowing the computation of statistical models on Riemannian manifolds are presented. When properly visualized, the resulting statistical models can be quite useful to analyze and compare the shape variations in different groups of patients. Two different approaches to visualization are demonstrated graphically. Finally, another important use of statistical models in medical imaging is to constrain the solution of inverse problems. Articulated models can readily be used in this context, we illustrate this in the context of 3D model reconstruction using partial data. More precisely, we will show the benefits of integrating a simple regularization term based on articulated statistical models to well known algorithms.

1 Introduction

The human spine is naturally curved, and its exact shape varies from one individual to another. These variations may be normal variations between healthy individuals, but they may also be a sign of pathology. Unfortunately, healthy variations in shape are large enough that they may be hard to distinguish from problematic deformations.

J. Boisvert (✉)
National Research Council, 1200 Montreal Road, Ottawa, Canada
e-mail: jonathan.boisvert@nrc-cnrc.gc.ca

© Springer International Publishing Switzerland 2015

S. Li and J. Yao (eds.), *Spinal Imaging and Image Analysis*,
Lecture Notes in Computational Vision and Biomechanics 18,
DOI 10.1007/978-3-319-12508-4_10

In order to better analyze spine shapes and potentially compare different groups of patients to find commonalities, one has to aggregate the data from a large number of patients and create statistical shape models. These models can then be used in a variety of ways, but we will concentrate on two important classes of applications in this chapter.

First, statistical shape models can be used in a descriptive fashion. In this context, they are used to describe, visualize, or summarize a large number of complex 3D models. They enhance healthcare workers' or researchers' understanding of how the spine shape varies for different groups of people and allow them to act accordingly. As an example, one can compare the spine shape of patients with and without a brace to slow down the progression of scoliosis. It is then possible to fine-tune the brace itself based on the comparison results.

Second, statistical shape models can be used to assist in image analysis tasks. Tasks such as 3D model reconstruction, registration, or the labeling of anatomical parts can be tedious to perform manually and difficult to perform automatically without a prior shape model. Because they implicitly encode what constitute a valid spine model, statistical shape models can be used to constrain the possible solutions. This reduces the solution space for image analysis algorithms, which translates into better accuracy or faster algorithms that necessitate fewer human interventions.

There are several ways to represent the shape of the spine. Therefore, there are several ways to create statistical shape models of the spine. One possibility is to have clinicians derive clinically relevant indices from the 3D models and then compute statistics based on these indices. This avenue was exploited in the context of studies on scoliosis. Several studies [11–13, 30, 41] examined the variations in the clinical indices used by physicians to quantity the severity of the deformations.

These indices have the advantage of enabling physicians to quickly and easily assess the severity of the scoliosis. However, they also present many problems. First, most clinical indices are global to the whole spine, and thus do not provide spatial insight about the local geometry. Furthermore, most of the indices (including the Cobb angle) are computed using 2D projections, where a significant part of the curvature could be hidden (since the deformity is three-dimensional). In addition, they describe the characteristics of the deformation, rather than the shape itself. It is usually not possible to move backward from clinical indices and compute a 3D model that could be compared with radiographs. This situation makes clinical indices far from being ideal for assisting image analysis algorithms. Finally, clinical indices are created based on the experience of clinical practitioners in order to describe certain types of deformations. Thus, clinical indices that are relevant to one pathology may be useless for another.

An alternative to clinical indices is to describe the geometry of the spine directly and not indirectly via the characteristics of a pathology. For instance, one can build statistical shape models based on a collection of predefined 3D points located on the spine. This idea is attractive because conventional multivariate probabilities and statistics can be leveraged to analyze and use the data. Thus, there are a large

number of statistical shape models based on the idea of using a set of 3D points that are anatomically comparable to build a statistical model. This is the basic idea behind the popular active shape models published by Cootes et al. [7]. Point-based statistical shape models were applied to spine applications on many occasions (for instance in [2, 3, 9, 16, 18, 21, 39]). The fact that the spine is composed of interconnected vertebrae can be integrated with this approach by making explicit the notion of neighboring vertebrae [10, 36, 37].

However, points do have limitations when one is interested in studying the whole spine. In general, points can be effectively adopted when the statistical relationship between different locations on the 3D model can be described by a linear combination of the points' coordinates. However, the spine is a complex articulated structure. The position and orientation of one vertebra affects the positions and orientations of its neighbors. These vertebrae then influence their own neighbors' orientations and positions. The successive small changes in orientations associated with a series of vertebrae result in large rotations. Unfortunately, these large rotations introduce a non-linear element into the statistical relationships, which needs to be taken into account when modeling the shape of the whole spine.

Graphical models can be used to integrate more complex probabilistic relationships between the modeled entities [8, 14, 38]. In theory, these could provide a framework in which the nonlinearities caused by the cumulative vertebral rotation would be explicitly taken into account, but they do not solve the problem by themselves.

It is possible to take into account a certain level of nonlinearity by using more complex statistical models. For instance, Kadoury et al. [23] used a locally linear embedding model, and Kirschner et al. [24] used a statistical shape model based on kernel principal component analysis (kernel-PCA) [27]. However, a more natural choice is to simply describe the spine as an articulated object. Then, one can compute statistics for this object, and the nonlinearities caused by the vertebrae rotations can be handled explicitly. Once this cause of nonlinearities is explicitly handled, it is still possible to use non-parametric methods in combination with articulated models [26] to analyze other causes of nonlinearities (pathology progression, various types of deformations, etc.).

In this chapter, we will investigate how articulated models can be used to compute statistics on the shape of the spine, and how these statistical models can be used in the context of 3D shape inference from images. Section 2 will therefore introduce articulated models and briefly describe how they can be extracted from existing data. Section 3 is then dedicated to the computation of statistical shape models from a series of articulated models. Visual methods to analyze the statistical models are also discussed. Various results based on a cohort of patients suffering from adolescent idiopathic scoliosis are shown as early as possible in the text to demonstrate the different concepts. Finally, Sect. 4 analyzes the use of an articulated statistical shape model to perform inference on 3D spine models.

2 Articulated Models

We define an articulated model as a collection of simpler models that all have an associated frame of reference and which normally move rigidly with respect to each other. In order to keep things simple, we will consider that each sub-model is in fact a point-based model (although there is nothing preventing someone from using more complex representations). More precisely, we will represent an articulated model as a series of rigid transformations T_i and 3D points p_i^j. The rigid transformations T_i describe the relative position and orientation of each sub-model, and the points p_i^j describe the local shape in the coordinate system associated with the *i*th sub-model. An example of anatomical landmarks that can be used to describe the shape of a vertebra is shown in Fig. 1. In this case, only the vertebral end-plates and pedicles are used, which leads to a compact model. However, more landmarks could be used to better describe the shape of the vertebral body or processes. In order to obtain the position and orientation of a vertebra with respect to a global

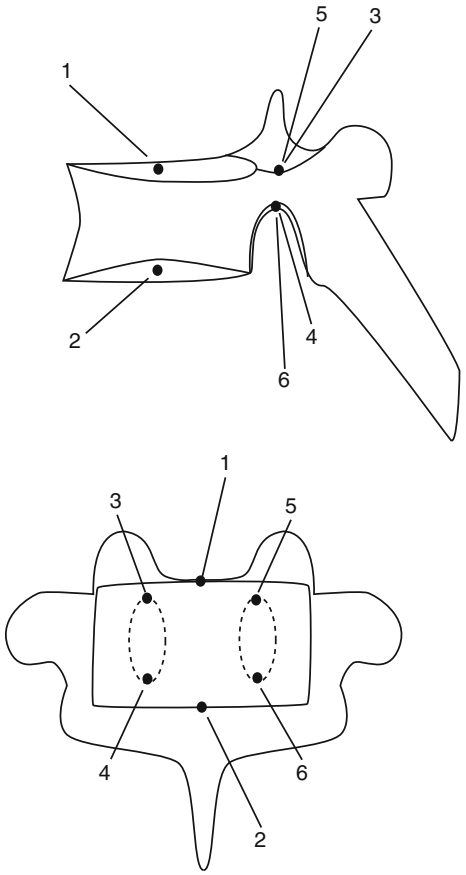


Fig. 1 Example of local anatomical landmarks that can be used to model shape of vertebra. *Top* lateral view. *Bottom* anterior view. *1* Center of the superior end plate. *2* Center of the inferior end plate. *3* Top of the right pedicle. *4* Bottom of the right pedicle. *5* Top of the left pedicle. *6* Bottom of the left pedicle

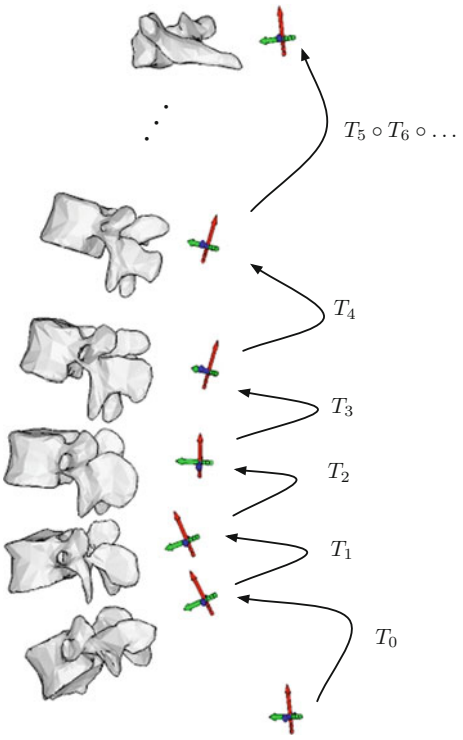


Fig. 2 Relative rigid transformations are used to represent a spine as an articulated object. To obtain absolute positions and orientations with respect to a global frame of reference, one would need to compose a series of inter-vertebral transformations. For instance, the position and orientation of the third vertebra from the bottom would be obtained by applying the rigid transformation $T_0 \circ T_1 \circ T_2$ to the global frame of reference

frame of reference, several transformations will need to be composed (see Fig. 2, for instance).

This representation of the spine is both intuitive and useful. The spine is a complex anatomical structure, which supports a large portion of our body weight and yet remains flexible enough to allow for complex motions. It includes the vertebrae as well as a variety of soft-tissues (inter-vertebral discs, ligaments, etc.). Vertebrae do not deform significantly in typical circumstances, but because they sit on top of one another; the orientation and position of one vertebra influence its direct neighbors. These relative transformations between neighboring vertebrae would not be captured correctly by a simple set of points and would result in an overestimation of the variability of the spine's geometry.

2.1 Reconstructing Articulated Models from Existing 3D Data

Articulated representations can be obtained from several sources. Most medical imaging modalities that are routinely applied to the spine can be used to produce articulated spine models. The exact procedure used to extract the articulated

models, as well as their accuracy will vary. However, clinical imperatives are likely to decide the imaging modality and posture of the patient.

Volumetric modalities such as computed tomography (CT), magnetic resonance (MR), and even 3D ultrasound can be processed by registering template vertebral models, manually identifying a predefined set of anatomical landmarks, or computing spine-based coordinate systems as part of a curve planar reformation approach [42]. The vast majority of CT or MR scanners require the patient to lie down. Therefore, one has to be very conscious that the variability encoded in the articulated statistical model will represent the anatomical variations for this particular posture.

Traditional bidimensional modalities can also be used to recreate articulated models of the spine. More than one image will generally be necessary to remove the ambiguities that are inherent to 2D images. The most common case in this category is the reconstruction of 3D models of the spine from multiple radiographs (see Chap. 5 of this book for an in-depth discussion). In essence, a human expert provides a computer program with indications about the matching locations in the different radiographs. A three-dimensional model can then be reconstructed by performing triangulation on the matched image coordinates. The way the human expert interacts with the computer program and how the computer program uses these interactions to perform matching varies significantly from one system to another.

Many methods that use an articulated statistical shape model can be used to extract new articulated models with relative ease. For instance, Klinder et al. [25] and Rasoulian et al. [35] used articulated models as part of segmentation methods of the spine applied to CT (computed tomography). The three-dimensional reconstruction of the spine from multiple radiographs was also performed with the help of an articulated statistical shape model [4, 29], and even ultrasound segmentation has been considered for an articulated biomechanical shape prior [20].

In order to illustrate the different concepts presented in this chapter and demonstrate the use of certain techniques, we will use a database of approximately 300 scoliotic patients that was collected at the Sainte-Justine Hospital (Montreal, Canada). These cases were all examined using stereo-radiographs of the spine. Six anatomical landmarks were manually identified by a skilled technician (the same landmarks presented in Fig. 1) on each vertebra from T1 (first thoracic vertebra) to L5 (last lumbar vertebra) on the two radiographs (a posterior-anterior and a lateral radiograph). Then, the 3D coordinates of the landmarks were computed using a triangulation procedure. The accuracy of this method was previously established to be 2.6 mm [1]. Once the landmarks are reconstructed in 3D, each vertebra can be rigidly registered to its upper neighbor. The resulting rigid transformations can then be used to build the articulated representation.

3 Statistics on Articulated Models

An articulated model is well adapted to represent the spine because it intuitively describes its natural degrees of freedom. Building statistical models based on this type of articulated representation is therefore very attractive. However, there are theoretical complications. Because articulated models rely on rigid transformations to encode inter-vertebral transformations, it is necessary to compute the statistics on these rigid transformations.

Rigid transformations are special because they cannot be added together like real numbers. Unfortunately, concepts as simple as the mean or standard deviation require summing the measurements as part of their computation. This calls for the generalization of a few basic concepts.

These generalizations are performed using a few mathematical tools borrowed from the field of Riemannian geometry. These tools will be introduced as simply as possible; thus, no prior knowledge of Riemannian geometry is needed. Nevertheless, interested readers can find a more complete introduction in [6].

We will first discuss a few properties of the rigid transformations related to Riemannian geometry that will be needed to build a statistical model of the spine. Then, we will present the generalization of the mean and covariance that are used in articulated statistical models of the spine. Finally, the visualization of the mean and covariance of the articulated spine models will be discussed.

3.1 Riemannian Geometry and Rigid Transformations

The most common method for numerically representing a rigid transformation T is to use the combination of a rotation matrix R and translation vector $t (T = \{R, t\})$. Using this representation, the action of T on a 3D point x can be written as $y = Rx + t$, and the composition of two rigid transformations T_2 and T_1 is given by $T_2 \circ T_1 = \{R_2R_1, R_2t_1 + t_2\}$.

The composition and action on points have very simple and efficient expressions using this representation. Thus, it may be tempting to compute statistics directly on R and t. However, naively summing rotation matrices and dividing the number of transformations in an attempt to compute an average rotation matrix will most likely result in a matrix that is not a rotation matrix. It could even lead to a singular matrix.

Fortunately, there are several other ways to represent rotations beyond the conventional rotation matrix. For instance, Euler angles are a compact notation that can be useful in certain applications. Unit quaternions require less mathematical operations to perform multiple compositions. Unfortunately, computing statistics directly on these representations leads to problems because the results depend on the orientation of the global frame of reference. Another, perhaps less known representation, called the rotation vector, offers certain advantages. This representation is defined using an axis of rotation n and an angle of rotation θ (see Fig. 3). The rotation vector r is then simply defined as the product of the unit vector n and θ.

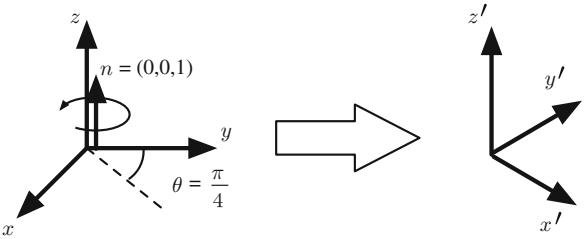


Fig. 3 Rotation vector is defined as a rotation of θ radians around an axis of rotation n and represented as their product $(\theta \frac{n}{\|n\|})$

A rotation vector can be converted into a rotation matrix using Rodrigues' formula:

$$R = I + \sin(\theta)S(n) + (1 - \cos(\theta))S^2(n) \tag{1}$$

$$\text{where} \quad S(n) = \begin{bmatrix} 0 & -n_z & n_y \\ n_z & 0 & -n_x \\ -n_y & n_x & 0 \end{bmatrix}.$$

It is also possible to compute the rotation vector from a rotation matrix using the following equations:

$$\theta = \arccos(\frac{Tr(R) - 1}{2}) \quad \text{and} \quad S(n) = \frac{R - R^T}{2\sin(\theta)}. \tag{2}$$

Let $\overrightarrow{T}$ be an alternate representation of rigid transformation T that uses the rotation vector instead of a rotation matrix; thus, $\overrightarrow{T} = \{r, t\}$. One of the main advantages of this representation is that it enables us to define a left-invariant distance $(d(\overrightarrow{T_1}, \overrightarrow{T_2}) = d(\overrightarrow{T_3} \circ \overrightarrow{T_1}, \overrightarrow{T_3} \circ \overrightarrow{T_2}))$ between two rigid transformations. This distance is defined as follows:

$$\begin{aligned} & d(\overrightarrow{T_1}, \overrightarrow{T_2}) = N_\lambda(\overrightarrow{T_2}^{-1} \circ \overrightarrow{T_1}) \\ \text{with:} \quad & N_\lambda(\overrightarrow{T})^2 = N_\lambda(\{r, t\})^2 = \|r\|^2 + \|\lambda t\|^2. \end{aligned} \tag{3}$$

The parameter λ makes it possible to change the relative importance of rotational changes in comparison to translational changes. It is a parameter worth considering carefully, because the rotation and translation are measured in different units. It would be easy to almost completely discard one in favor of the other without a well chosen value. In our experience, choosing a value that leads to variabilities that are approximately equally distributed in the translational and rotational parts of the transformations works well for descriptive studies.

The distance d actually comes from the fact that rigid transformations constitute a Riemannian manifold equipped with a metric. Because of this Riemannian structure, it is possible to locally define tangent planes to the manifold and map vectors from these tangent planes to the manifold itself in such a way that the magnitude of the vectors are consistent with the distances on the manifold.

We can express this mapping (called the exponential map) and its inverse (the logarithmic map) around the identity transformation as follows:

$$\mathrm{Exp}_{Id}(\vec{T}) = \begin{vmatrix} R(r) \\ t \end{vmatrix} \quad \text{and} \quad \mathrm{Log}_{Id}(T) = \begin{vmatrix} r(R) \\ \lambda t \end{vmatrix}, \tag{4}$$

where $R(r)$ and $r(R)$ are the conversion from the rotation vector to rotation matrix and vice versa, respectively.

The exponential and logarithmic map around any rigid transformation μ can then be related to the map around the identity transformation. Because of the left-invariance property of the distance d, that relation can then be expressed as follows:

$$\begin{aligned} \mathrm{Exp}_{\mu}(\overrightarrow{x}) &= \mu \circ \mathrm{Exp}_{Id}(J_L(\mu)^{(-1)}\overrightarrow{x}) \\ \mathrm{Log}_{\mu}(T) &= J_L(\mu)\mathrm{Log}_{Id}(\mu^{-1} \circ T), \end{aligned} \tag{5}$$

where $J_L(\overrightarrow{T_2})$ is the Jacobian of the composition ($J_L(\overrightarrow{T_2}) = \frac{\partial}{\partial \overrightarrow{T_2}} \overrightarrow{T_2} \circ \overrightarrow{T_1}|_{\overrightarrow{T_1}=Id}$). Its detailed derivation can be found in [32]. Even though these exponential and logarithmic maps are advanced mathematical concepts, they will be precious tools to generalize statistical concepts for rigid transformations and, by extension, for articulated models of the spine.

3.2 Mean and Centrality

As mentioned earlier, rigid transforms cannot be added together. They can, however, be composed, inversed, and compared using a valid distance. We should therefore use a generalization of the conventional mean that takes advantage of these properties.

It can be observed that the conventional mean minimizes the Euclidian distance of the measures with the mean. Thus, given a general distance, a generalization of the conventional mean would be to define the mean as the element μ of a manifold $\mathcal{M}$ that minimizes the sum of the distances with a set of elements $x_{0\ldots N}$ of the same manifold $\mathcal{M}$. Thus, μ is given by the following:

$$\mu = \underset{x \in M}{\arg\min} \sum_{i=0}^{N} d(x, x_i)^2. \tag{6}$$

This generalization of the mean is called the Fréchet mean [19]. It is equivalent to the conventional mean for vector spaces with a Euclidian distance.

The computation of the Fréchet mean directly from the definition is difficult because of the presence of a minimization operator. Fortunately, a simple gradient descent procedure can be used to compute the mean [31] when the exponential and logarithmic maps are known. This procedure is summarized by the following recurrent equation:

$$\mu_{n+1} = \mathrm{Exp}_{\mu_n}\left(\frac{1}{N}\sum_{i=0}^{N}\mathrm{Log}_{\mu_n}(x_i)\right). \tag{7}$$

This equation is guaranteed to converge. Moreover, in practice it converges rather quickly. We observed that it generally converged in less than five iterations for articulated models of the spine.

To use Eq. (7), it is necessary to initialize the mean to start the procedure. The initial value can be one of the points of the set from which the mean is computed. Furthermore, more than one starting point can be tried to test the uniqueness of the mean and escape local minimums.

The Fréchet mean is not unique in general, and the starting point of the iterative procedure is theoretically important. However, in the case of an articulated model of the spine, multiple strategies were tried, but generally produced the exact same results.

3.3 Covariance and Variability

In conventional statistics, variations would first be studied with the covariance matrix. However, because the covariance matrix is defined using the conventional mean, we need to define its generalization around the Fréchet mean.

The exponential map can be intuitively understood as a local linearized vector space around a given point on a manifold. Thus, as a simple generalization, it is possible to simply compute a covariance matrix around the Fréchet mean on the exponential map. Mathematically, this can be expressed by the following:

$$\begin{aligned}\Sigma &= E\left[\mathrm{Log}_{\mu}(x)^{T}\mathrm{Log}_{\mu}(x)\right] \\ &= \frac{1}{N}\sum_{i=0}^{N}\mathrm{Log}_{\mu}(x_i)^{T}\mathrm{Log}_{\mu}(x_i).\end{aligned} \tag{8}$$

This generalized covariance computed in the tangent space of the mean and associated variance are connected because $Tr(\Sigma) = \sigma^2$, which is also the case for the usual vector space definitions.

Articulated models contain several rigid transformations and sub-models, which can all be correlated to each other. To explore the covariance of a whole articulated

model, one only needs to consider the logarithmic map of the Cartesian product of its components. In other words, it is possible to concatenate the logarithmic map of individual inter-vertebral transformations and sub-models to obtain the logarithmic map of the whole articulated model and use it to compute the generalized covariance matrix.

3.4 Graphical Visualization

It is crucial to be able to visualize a statistical shape model in an intuitive way in order to efficiently communicate results. Large tables filled with numbers are far from being ideal. Fortunately, there is an easier and more intuitive method to do so using an articulated statistical model.

Visualizing the mean model is rather simple; one can simply render the mean model, as it would be done for any of the individual models that were used to build the statistical models. For instance, Fig. 5 was produced by rendering a template model that was deformed based on the positions and orientations of the Fréchet mean computed over a large set of individual spine models from scoliotic patients.

However, it is more challenging to properly illustrate the variabilities. One relatively easy way to do so is to focus on the individual covariance matrices associated with the rotations and translations. Each three-by-three matrix can then be visualized as a 3D ellipsoid. The ellipsoid geometric description is obtained by performing an eigenvalue decomposition on the covariance matrix. The eigenvectors become the principal axes of the ellipsoid, and the eigenvalues are used to scale the length of the ellipsoid along the principal axes.

In the case of translations, the ellipsoid can be interpreted by thinking of the lengths along the principal axes as the standard deviations of the translations measured along these three-dimensional directions. For rotations, the interpretation relies on the fact that the exponential map is given by the rotation vector. Therefore, the lengths of the ellipsoid along the principal axes may be interpreted as the standard deviations of the rotations measured around these axes (see Fig. 4).

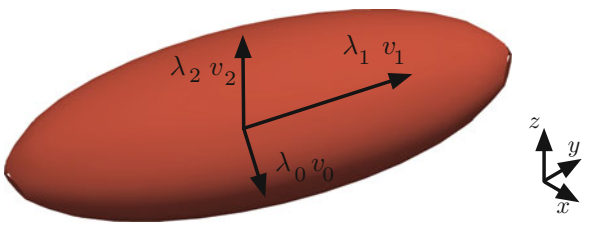


Fig. 4 Three-dimensional ellipsoids can be used to visualize covariance matrices associated with translations or rotations. The generalized covariance matrices of these transformations are decomposed into their eigenvectors v and eigenvalues λ, which are used to determine the directions and lengths of the ellipsoid's principal axes, respectively

3.5 Graphical Visualization of Scoliotic Patient Variability

Adolescent idiopathic scoliosis is a pathology that causes three-dimensional deformations of the spine. Because these deformations have various shapes and severities, a dataset comprised of scoliotic patients is ideal to illustrate the different concepts introduced so far. We thus used a database of approximately 300 scoliotic patients who were diagnosed with scoliosis, but had never before received orthopedic treatment. The mean model and variability associated with the relative rigid transformations are illustrated in Fig. 5.

It can be observed that an average scoliotic patient has a curvature in the frontal plane that a healthy individual would not have. Because right thoracic curvatures

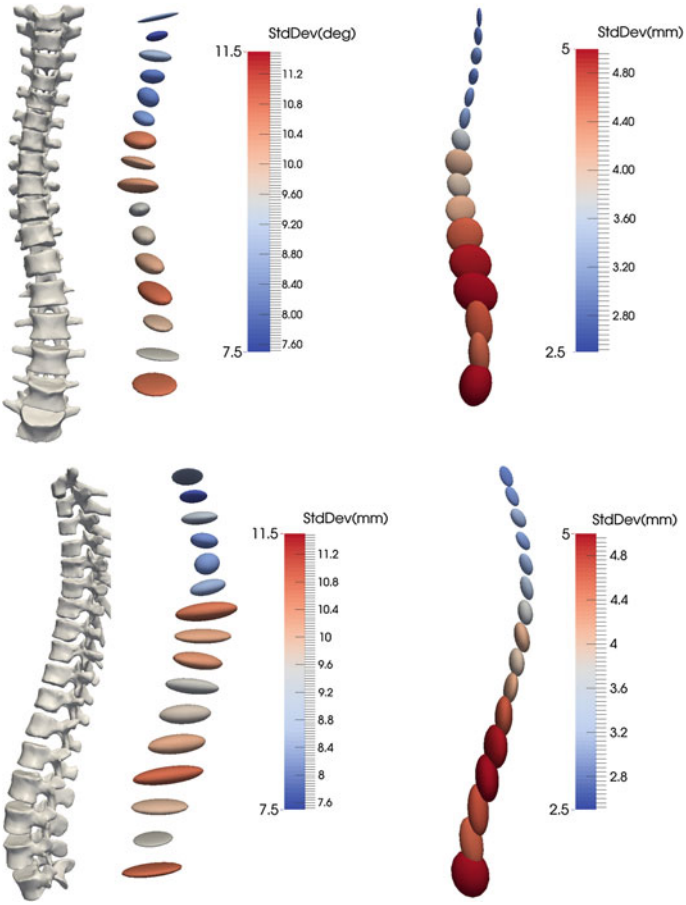


Fig. 5 Frontal view (*above*) and lateral view (*below*) of a statistical model of the relative poses for a group of scoliotic patients. From *left* to *right* mean spine model, rotation covariance, and translation covariance

are known to be more prevalent than left thoracic curvatures among scoliotic patients, this imbalance is reflected in the mean spine shape. The normal healthy kyphosis and lordosis are also present in the average model.

The variability of the relative rigid transformations between vertebral levels is also interesting to observe. Its largest departure from the mean shape in terms of orientation appears to be along the posterior–anterior axis. This means that the variability would be most noticeable in a posterior–anterior radiograph, which is the one on which the Cobb angle is generally measured. Finally, one can note that the most important translational variability appears to be in the axial direction, which is explained by the fact that most patients are growing adolescents and, consequently, the variations in the patients' heights are important.

As an alternative to relative rigid transformations between neighboring vertebrae, one might instead consider the absolute transformations, where one vertebra serves as a global reference. This may seem like an appealing option, because it would reduce the number of rigid transformation compositions needed to use the model in inference applications. However, as can be seen in Fig. 6 (where *L5* is used as a reference), the variabilities become much larger and their values then depend on the arbitrary choice of the reference vertebra. For these reasons, absolute transformations should be used with caution.

The tools presented so far can also be used to study the effect of orthopedic treatment on the spine shape of patients. The computation of the mean shape before and after treatment remains identical to the procedure used in Figs. 5 and 6. However, the variability is computed on the rigid transformations that transform the articulated model before treatment into the articulated model recorded after treatment. Figure 7 illustrates this procedure on a cohort of patients who received orthopedic braces to slow down the progression of scoliosis. It is also possible to test for differences between the effect and a control group [5] to locate significant effects and help optimize treatments.

3.6 Component Analysis

Another useful way to visualize the variability in a large dataset is to find unidimensional axes along which the variability is particularly strong. Then, a series of models can be reconstructed along these axes and viewed either as an animation or side-by-side (which has obvious advantages for printed media).

The best-known method in this family is called principal component analysis [15]. This method was developed for multi-dimensional vector spaces. However, it was shown that it could also be applied to manifolds under certain conditions [17].

The general idea is that, unlike the manifold itself, the tangent plane around the mean is a vector space, and its basis can be changed by applying a linear transformation. Thus, we seek an orthonormal matrix $A(AA^T = I)$ to linearly transform the tangent plane ($\mathrm{Log}_\mu(g) = A\,\mathrm{Log}_\mu(f)$) so that the resulting components would be uncorrelated to each other and have decreasing variances. In other words, the

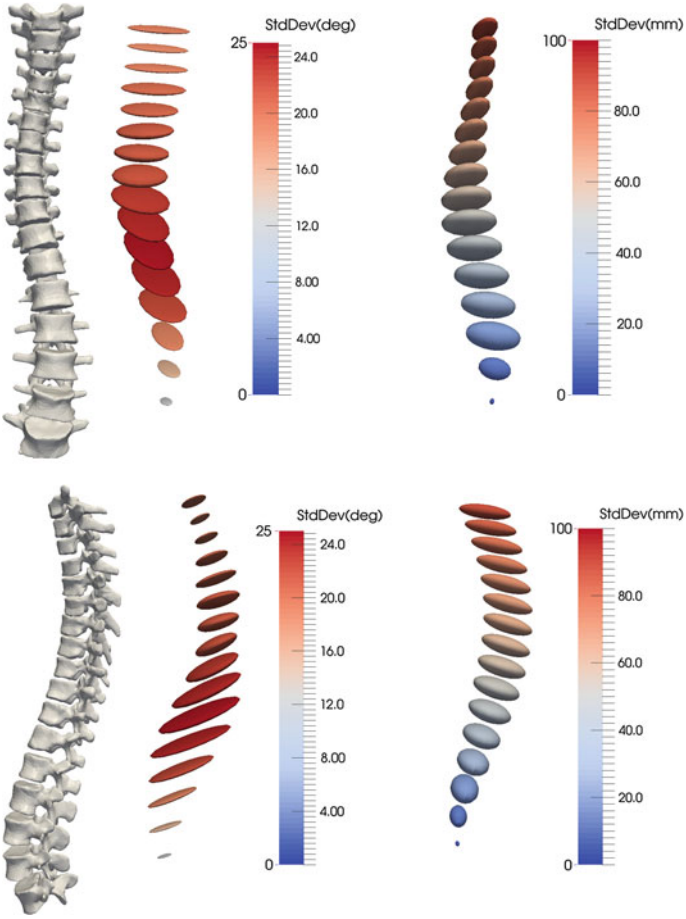


Fig. 6 Frontal view (*above*) and lateral view (*below*) of a statistical model of the absolute poses of the vertebrae for a group of scoliotic patients. From *left* to *right* mean spine model, rotation covariance, and translation covariance

generalized covariance in the transformed tangent space would be a diagonal matrix with decreasing values on the diagonal ($\Sigma_{gg} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_k)$ with $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k$). The covariance matrix in the transformed tangent space can be computed from the original covariance matrix Σ and the transformation matrix A:

$$\Sigma_{gg} = \text{diag}(\lambda_1, \lambda_2, \ldots, \lambda_k) = A\Sigma_{ff}A^T. \tag{9}$$

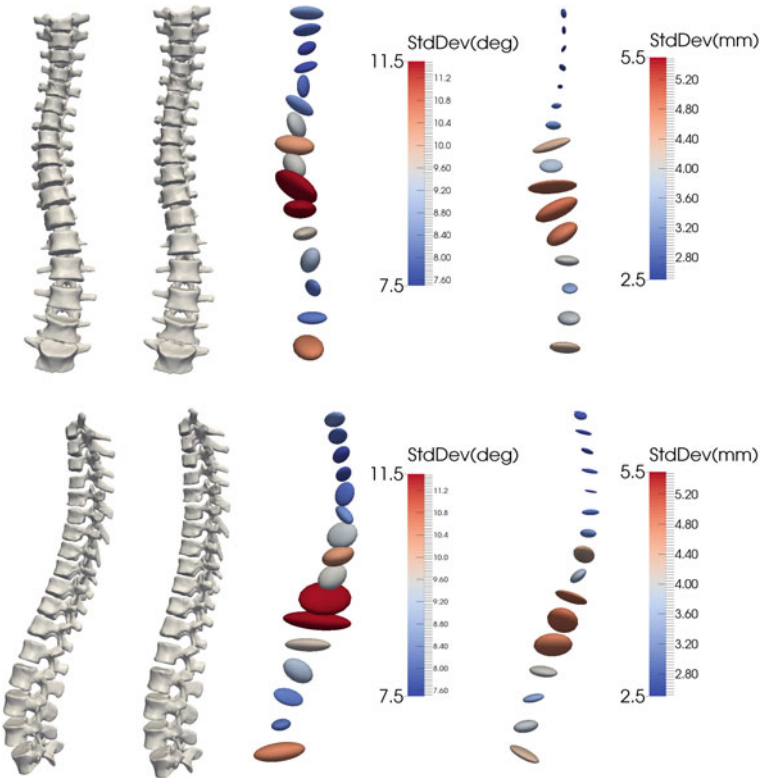


Fig. 7 Frontal view (*above*) and lateral view (*below*) of a statistical model of the spine shape deformations associated with the Boston brace. From *left* to *right* mean shape prior treatment, mean shape with the brace, rotation covariance of the deformations, and translation covariance of the deformations

If A is rewritten as $A = [a_1, a_2, \ldots, a_k]^T$, then it is easy to show that

$$[\lambda_1 a_1, \lambda_2 a_2, \ldots, \lambda_k a_k] = \left[\Sigma_{ff} a_1, \Sigma_{ff} a_2, \ldots, \Sigma_{ff} a_k\right]. \tag{10}$$

The line vectors of matrix A are therefore the eigenvectors of the original covariance matrix, and the elements of the covariance matrix in the transformed space are the eigenvalues of the original covariance. This is exactly the same procedure that is used to perform PCA (principal component analysis) in real vector spaces. As in real vector spaces, the variance is left unchanged because $\sigma^2 = \mathrm{Tr}(\Sigma_{ff}) = \mathrm{Tr}(\Sigma_{gg})$, and the cumulative fraction of the variance explained by the first n components is

$$p = \frac{1}{\sigma^2} \sum_{i=1\ldots n} \lambda_i. \tag{11}$$

A shape model can be recreated from the coordinates of the transformed tangent space simply by going back to the original tangent space and projecting the model on the manifold using the exponential map. Thus, if α_i is the coordinate associated with the *i*th principal component, the following equation can be used to re-create a shape model:

$$S = \mathrm{Exp}_{\mu}(\sum_{i=1}^{k} \alpha_i a_i). \tag{12}$$

The visualization of the components allows an analysis of not only the variations of individual inter-vertebral transformations, but also of the variations of the spine shape as a whole. In a sense, the method presented in Sect. 3.4 allowed us to analyze the variations of individual translations and rotations in the articulated models of the spine, and component analysis can be used for the major modes of variation of the whole spine. These visualization methods are thus complementary.

Component analysis helps identifying and analyzing deformation trends that changes the shape of the whole spine. This can often be used to better understand different concurrent factors that contribute to the variation in the spine's shape. For instance, the first principal deformation mode of our database of scoliotic patients illustrated in Fig. 8 shows an elongation of the spine combined with the development of a thoracic curve. This component explains the largest amount of variance in the spine shapes. It could also be analyzed more generally as being a combination

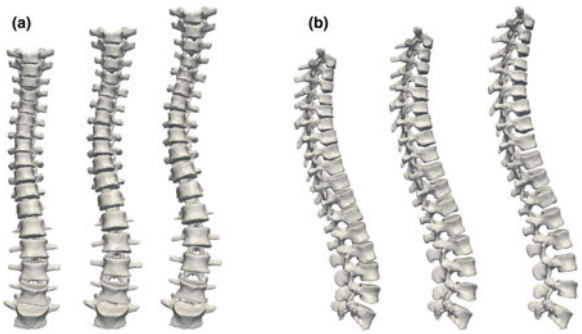


Fig. 8 First principal deformation modes for a dataset of adolescent idiopathic scoliosis patients. Spine models were rendered (from *left* to *right*) for −3, 0, and 3 times the standard deviation explained by the corresponding deformation mode. **a** Front view. **b** Lateral view

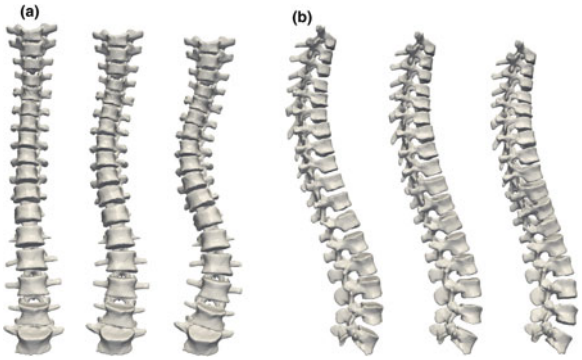


Fig. 9 Second principal deformation modes for a dataset of adolescent idiopathic scoliosis patients. Spine models were rendered (from *left* to *right*) for −3, 0, and 3 times the standard deviation explained by the corresponding deformation mode. **a** Front view. **b** Lateral view

of adolescent patients being at various stages in their growth and the development of a scoliotic curve.

The second mode (as seen in Fig. 9) can be described as a double thoraco-lumbar curve. In this case, there are two opposing curves: one in the thoracic segment (upper spine) and another in the lumbar segment (lower spine). The third mode of deformation (illustrated by Fig. 10) is another thoracic curve, but it affects a more important portion of the spine than the curve observed in the first mode. It is also interesting to note that, in addition to the curves visible on the posterior–anterior view, the second and third principal deformation modes are also associated with the development of a kyphosis (back hump) that can be observed in the lateral view.

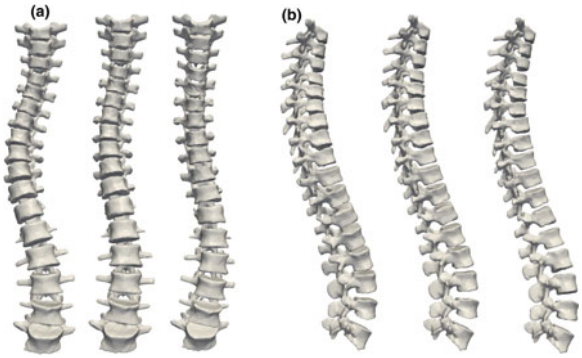


Fig. 10 Third principal deformation modes for a dataset of adolescent idiopathic scoliosis patients. Spine models were rendered (from *left* to *right*) for −3, 0, and 3 times the standard deviation explained by the corresponding deformation mode. **a** Front view. **b** Lateral view

4 Shape Inference

Articulated models of the spine can be used beyond their role in descriptive statistics. They model inter-patient variability quite well and offer a compact representation of what is and is not a valid 3D spine model. Furthermore, most common manipulations of articulated models have closed-form solutions and are differentiable. These two qualities make an articulated model well suited for integration with different shape estimation algorithms. Shape estimation systems are generally large and complex systems where image processing, system calibration, and a user interface all have to work in synergy to produce suitable results. We will, however, focus our attention only on the integration of the articulated statistical models to simplify our exposition, make it more understandable, and keep it concise.

4.1 Articulated Shape Prior for 3D Reconstruction from 2D Correspondences

A common and simple 3D reconstruction problem is the computation of the three-dimensional coordinates of points based on image coordinates in multiple images. This basic problem can be solved by calibrating the projective geometry of the system and then performing triangulation on the image coordinates corresponding to the same 3D point in multiple images. The corresponding points can be obtained automatically using an elaborate image processing system or manually defined by an expert in spinal anatomy. This general idea has been applied to 3D spine reconstruction from multiple radiographs for many years [33].

The performances deteriorate quickly as the image calibration and image correspondences are degraded, either by human error or simply by lower quality radiographs. However, an articulated statistical shape model prior can be integrated quite simply to mitigate some of these problems. Thus, let $p_{2D}^{i,j,k}$ be the image coordinates of an anatomical landmark identified in a radiograph. The index i associates a landmark with a vertebra. The index j indicates the position of the anatomical landmark within the set of landmarks used for the ith vertebra. Finally, k denotes the index of the radiograph on which the coordinates were measured. In addition, let S be the departure from the Fréchet of an articulated model, which is defined as follows:

$$S = (s_1, s_2, \ldots, s_N) \text{ with } s_i = \left(\bar{T}_i^{-1} \circ T_i, p_{i,1} - \bar{p}_{i,1}, p_{i,2} - \bar{p}_{i,2}, \ldots, p_{i,M} - \bar{p}_{i,M}\right). \tag{13}$$

A simple but effective way to combine the similarity between $p_{2D}^{i,j,k}$ and S with prior knowledge of possible spine shapes is to sum the Mahalanobis distance and the quadratic error on the anatomical landmarks. The following equation summarizes this operation:

$$C(S) = S\Sigma^{-1}S^T + \alpha \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{o}\left\| p_{2D}^{i,j,k}(S) - \hat{p}_{2D}^{i,j,k} \right\|^2, \tag{14}$$

where α is the relative weight of the anatomical landmark error with respect to the prior spine shape knowledge. If identification errors come from a Gaussian distribution with known variance and the departures from the Fréchet mean can be approximated by another Gaussian distribution, then α can be chosen to obtain the *maximum* a posteriori. However, the distribution and variance of the identification errors are rarely known in advance, and α has to be adjusted manually. A non-linear optimization is then performed to estimate the best articulated model given $\hat{p}_{2D}^{i,j,k}$. The initial solution from which the optimization procedure starts can be provided by a direct linear transform.

To illustrate the benefits of this approach, we numerically simulated a radiographic setup with a posterior–anterior and a lateral radiograph. A set of six anatomical landmarks per vertebra were then identified in the radiographs and corrupted using different noise levels. The conventional DLT (direct linear transform) algorithm was applied to the noise corrupted correspondences, as well as the algorithm summarized by Eq. (14). The procedure was repeated for 50 spine models for each noise level.

The results of the simulations are shown in Fig. 11. It appears that the methods yield comparable results when the errors on the correspondences are low. The direct linear method even seems to enjoy a small advantage, which may have to do with the articulated reconstruction being prone to reach a local minimum with low noise levels. However, the articulated reconstruction appears to cope better with high levels of noise. The reconstruction errors obtained with the articulated reconstruction method were better when the standard deviation of the noise was more than approximately 3 pixels.

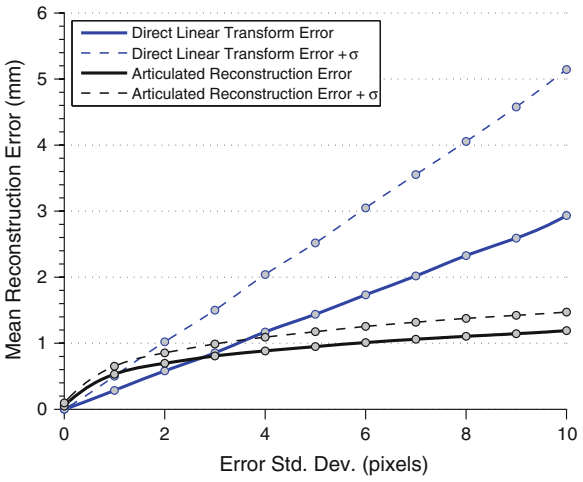


Fig. 11 Three-dimensional reconstruction performed by applying direct linear transform to corresponding image coordinates and by maximizing the posterior probability of the articulated model

This example illustrates that powerful prior models such as a statistical model of the spine based on articulated modeling make it possible to mitigate the effect of noise. In the particular case of the reconstruction of the spine from two radiographs, it appears that using an articulated model of the spine may be interesting only when there is a high level of noise.

However, a more important problem for 3D spine reconstruction using radiographs is the amount of user interaction required and the impossibility of finding meaningful correspondences for certain anatomical landmarks.

4.2 Completing Partial Models

Thus, another interesting problem would be to complete a three-dimensional spine model that has been partially reconstructed. Incomplete models can be caused by several factors. Surgical instrumentation could have occulted parts of the spine. Certain vertebrae may be outside the radiograph's borders, or a user may just want to save time by reconstructing only a few vertebrae.

The problem then involves reconstructing an articulated model of the whole spine based on a few vertebrae. The idea is to use the statistical model to fill the gaps by computing the most likely model given a set of constraints. This can be done, in the case of a Gaussian distribution, by minimizing the Mahalanobis distance while preserving certain constraints. This idea may be formalized by the following equation:

$$\tilde{S} = \arg\min_{S} S\Sigma^{-1}S^{T} \tag{15}$$

$$\begin{aligned} \text{Subject to}: \quad & \tilde{T}_i^{absolute} = T_{i-1}^{absolute} \circ \bar{T}_i \circ T_i \quad \forall i \in K \\ & p_{i,j} = \tilde{p}_{i,j} \quad \forall \quad (i,j) \in L, \end{aligned}$$

where $\tilde{T}_i^{absolute}$ are the absolute poses of known vertebrae, $\tilde{p}_{i,j}$ are known anatomical landmarks, K is the set of all known vertebrae, and L is the set of known landmarks.

In summary, Eq. (15) states that we seek the closest model to the mean with respect to the Mahalanobis distance. This model should, however, match the poses and shapes of the already available vertebrae.

The optimization method used to minimize Eq. (15) is sequential quadratic programming [40]. It was selected because the cost function is quadratic and the constraints are usually close to linear constraints when an initial solution sufficiently close to the optimum is provided. Sequential quadratic programming is a generalization of Newton's method for unconstrained optimization, which iteratively solves a quadratic model of the problem using linear approximations of the constraints. Like Newton's method, it is a local optimization method, and it is subject to entrapments in local minima.

A good starting point that satisfies all the constraints is thus necessary. A simple solution that worked well in practice was to use the one-parameter subgroup (computed from the matrix exponential) of the rigid transformations to split a rigid transformation into smaller equal transformations when more than two consecutive vertebrae were missing.

To get an idea of the achievable performances of this type of method, we selected 50 cases from a database of 3D spine models from the Sainte-Justine hospital. We then removed a certain number of vertebrae from the models (complete models include both lumbar and thoracic vertebrae, which means a total of 17 vertebrae), and reconstructed them by minimizing Eq. (15). The removed vertebrae were equally distributed along the spine. The reconstructed vertebrae were then compared with the original model, and the results are illustrated in Fig. 12.

These results are a great demonstration of the power of statistical models based on an articulated modeling of the spine. It was possible to remove eight vertebrae from a complete model and still expect a mean absolute error of just less than 1 mm. In this experiment, the known vertebrae were not corrupted by calibration errors. Thus, these results should be considered as a minimum bound for real-world experiments.

Nonetheless, these interesting results demonstrate that representing a spine using a large number of anatomical landmarks (102 in this case) is very redundant from a statistical point of view. A statistical model based on an articulated modeling of the spine can harness this redundancy and reduce the amount of information needed to recreate a valid 3D model.

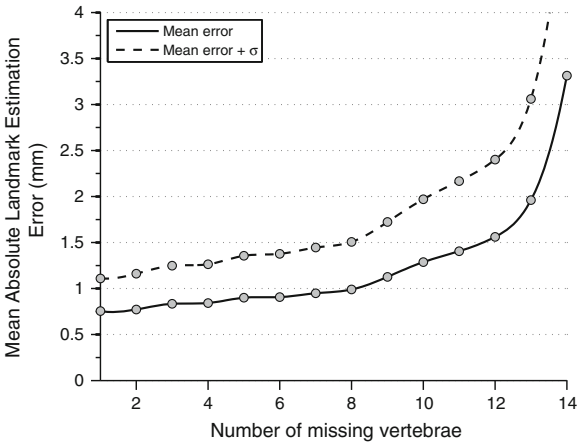


Fig. 12 Three-dimensional spine models with missing vertebrae reconstructed using an articulated model prior. The mean absolute error and standard deviation of the individual landmark reconstruction errors are shown for different numbers of missing vertebrae

4.3 Fast Reconstruction with Reduced Human Intervention

Producing complete 3D reconstructions based on a large number of manually identified anatomical landmarks is a long and labor intensive process that is subject to human error. This task can be greatly accelerated by reducing the amount of human intervention required for one reconstruction. Section 4.2 provided a hint that an efficient way to do so would be to use a statistical shape model to find the most likely spine given reduced human input.

However, instead of completing partial models, it is possible for a human user to provide more general cues about the shape of the whole spine. Then, a complete 3D model can be generated based on these cues. For instance, a human expert can quickly define a spline that crosses the center of the vertebral bodies from a few control points on each radiograph. Then, a computer program can find the most likely articulated model corresponding to that input [29].

Figure 13 shows an example where a posterior–anterior radiograph and a lateral radiograph were used. A human expert then selected five or six control points to define the splines shown in yellow. An articulated model was then computed, and its associated anatomical landmarks are shown in blue.

The method optimizes a combination of the prior probability of the reconstructed model and that of the distances between the user-defined splines and the centers of the vertebral bodies. This method is obviously not the only fast reconstruction method available (see for instance [22, 28, 34], Chap. 5, and the references therein).

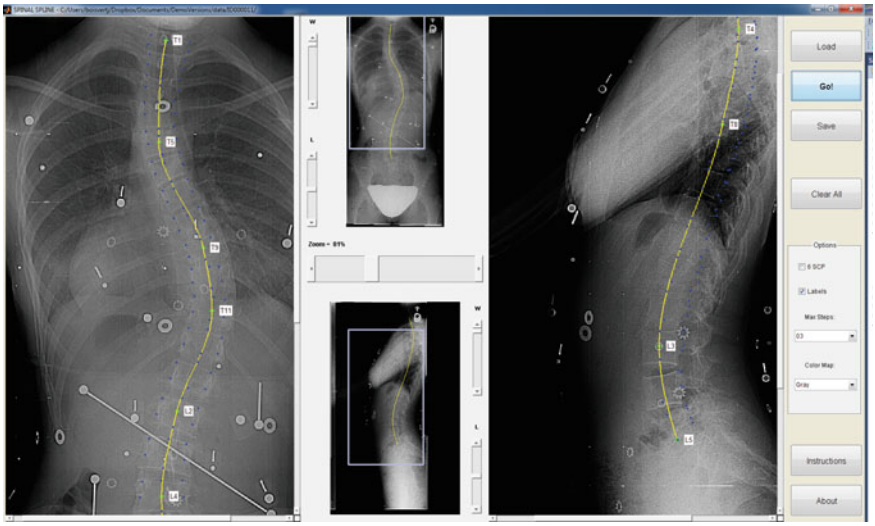


Fig. 13 Articulated spine models can be used as a priori shape models to perform three-dimensional spine reconstruction from multiple radiographs. In this example, a spline is defined on each radiograph, and the most likely articulated model is computed based on an optimization scheme

This family of methods is of great importance because it transforms the 3D reconstruction of the spine from radiographs from a long process that had to be performed by a specialist to a relatively fast procedure that can be performed by a skilled user (who is not necessarily an expert in either spine anatomy or 3D reconstruction). In [29], mild cases were reconstructed in 90 s, on average, and severe cases in 110 s (these times include both user interactions and computations).

5 Future Directions and Conclusion

Articulated statistical models are powerful tools for the analysis of the three-dimensional shape of the human spine. Their use for descriptive statistics of the spine shape of large patient groups was demonstrated, and the interpretation of the results proved to be intuitive thanks to an appealing visualization scheme. We also provided overviews of the methods needed to build articulated statistical models and use these as a valuable part of more complex systems that perform 3D shape inference.

It cannot be denied that articulated models are more complex to comprehend and handle than unstructured point clouds. However, this chapter provides the basis of a framework to handle articulated models rigorously. Fortunately, this does not mean overly complex methods. In most cases, it means computing the Fréchet mean instead of the traditional mean and using the rotation vector instead of other representations in the computations. Once these operations are isolated in dedicated computer methods, the additional complexity associated with articulated models becomes remarkably easy to manage.

Articulated models, however, do necessitate more mathematical operations when they need to be compared to absolute 3D points or reprojected on images. This can be important when comparisons have to be performed multiple times as part of an optimization scheme. These additional computations are, however, partly balanced by the strong constraints that a statistical model based on an articulated model can place on the solution space.

We most certainly have not yet explored all the scenarios in which these articulated statistical models could be useful. For instance, in this chapter, we discussed in length three-dimensional shape modeling of the spine, but we did not tackle the fourth dimension. However, time is the key to numerous challenging new applications. For instance, tracking and analyzing the effect of a disease or the effect of a treatment on the geometry of the spine are important endeavors. Studying the deformation of the spine under different types of strain may also reveal important information about the spine's biomechanics. Interpreting multiple images of the spine taken at different times and in different postures using multiple modalities would be needed to model the influence of time in a sensible way. It therefore seems that incorporating temporal variations will offer numerous opportunities and challenges for further developing statistical articulated models of the spine.

Acknowledgments We wish to thank M.-A. Drouin whose astute comments greatly improved the manuscript. We would also like to acknowledge Dr. H. Labelle and the Sainte-Justine Hospital (Montreal, Canada) staff for meticulously collecting a large number of high quality 3D reconstructions of the spine over many years and for giving us access to some of these reconstructions. Statistical methods are quite powerful, but they are useless without good data.

References

1. Aubin CE, Dansereau J, Parent F, Labelle H, de Guise J (1997) Morphometric evaluations of personalised 3D reconstructions and geometric models of the human spine. Med Bio Eng Comp 35
2. Becker M, Kirschner M, Fuhrmann S, Wesarg S (2011) Automatic construction of statistical shape models for vertebrae. In: Medical image computing and computer-assisted intervention, Springer, Berlin, pp 500–507
3. Benameur S, Mignotte M, Parent S, Labelle H, Skalli W, de Guise J (2003) 3D/2D registration and segmentation of scoliotic vertebrae using statistical models. Comput Med Imaging Graph 27(5):321–337
4. Boisvert J, Cheriet F, Pennec X, Labelle H, Ayache N (2008) Articulated spine models for 3-D reconstruction from partial radiographic data. IEEE Trans Biomed Eng 55(11):2565–2574
5. Boisvert J, Cheriet F, Pennec X, Labelle H, Ayache N (2008) Geometric variability of the scoliotic spine using statistics on articulated shape models. IEEE Trans Med Imaging 27 (4):557–568
6. Boothby WM (1986) An introduction to differentiable manifolds and Riemannian geometry, vol 120. Academic Press, Massachusetts
7. Cootes TF, Taylor CJ, Cooper DH, Graham J (1995) Active shape models-their training and application. Comput Vis Image Underst 61(1):38–59
8. Corso JJ, Raja'S A, Chaudhary V (2008) Lumbar disc localization and labeling with a probabilistic model on both pixel and object features. In: Medical image computing and computer-assisted intervention, Springer, Berlin, pp 202–210
9. Cresson T, Chav R, Branchaud D, Humbert L, Godbout B, Aubert B, Skalli W, De Guise JA (2009) Coupling 2D/3D registration method and statistical model to perform 3D reconstruction from partial X-rays images data. In: IEEE engineering in medicine and biology conference, pp 1008–1011
10. de Bruijne M, Lund MT, Tankó LB, Pettersen PC, Nielsen M (2007) Quantitative vertebral morphometry using neighbor-conditional shape models. Med Image Anal 11(5):503–512
11. Delorme S, Labelle H, Aubin CE, de Guise JA, Rivard CH, Poitras B, Coillard C, Dansereau J (1999) Intraoperative comparison of two instrumentation techniques for the correction of adolescent idiopathic scoliosis. Rod rotation and translation. Spine 24:2011–2011
12. Delorme S, Labelle H, Aubin CE, de Guise JA, Rivard CH, Poitras B, Dansereau J (2000) A three-dimensional radiographic comparison of Cotrel-Dubousset and colorado instrumentations for the correction of idiopathic scoliosis. Spine 25:205–210
13. Delorme S, Labelle H, Poitras B, Rivard CH, Coillard C, Dansereau J (2000) Pre-, intra-, and postoperative three-dimensional evaluation of adolescent idiopathic scoliosis. J Spinal Disord 13
14. Dong X, Zheng G (2010) Automated vertebra identification from X-ray images. In: Image analysis and recognition, Springer, Berlin, pp 1–9
15. Duda RO, Hart PE, Stork DG (2012) Pattern classification. Wiley, New York
16. Elias de Oliveira M, Reutlinger C, Zheng G, Hasler C, Buchler P (2010) Statistical shape modeling of pathological scoliotic vertebrae: a comparative analysis. In: 2010 annual international conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, pp 5939–5942

17. Fletcher PT, Lu C, Pizer SM, Joshi S (2004) Principal geodesic analysis for the study of nonlinear statistics of shape. IEEE Trans Med Imaging 23(8)
18. Fleute M (2001) Shape reconstruction for computer assisted surgery based on non-rigid registration of statistical models with intra-operative point data and X-ray images. These de l'Université Joseph Fourier, Grenoble I
19. Fréchet M (1948) Les éléments aléatoires de nature quelconque dans un espace distancié. Ann Inst H Poincaré 10:215–310
20. Gill S, Abolmaesumi P, Fichtinger G, Boisvert J, Pichora D, Borshneck D, Mousavi P (2012) Biomechanically constrained groupwise ultrasound to CT registration of the lumbar spine. Med Image Anal 16(3):662–674
21. Heimann T, Meinzer HP (2009) Statistical shape models for 3D medical image segmentation: a review. Med Image Anal 13(4):543–563
22. Humbert L, De Guise J, Aubert B, Godbout B, Skalli W (2009) 3D reconstruction of the spine from biplanar X-rays using parametric models based on transversal and longitudinal inferences. Med Eng Phys 31(6):681–687
23. Kadoury S, Cheriet F, Labelle H (2009) Personalized X-ray 3-D reconstruction of the scoliotic spine from hybrid statistical and image-based models. IEEE Trans Med Imaging 28 (9):1422–1435
24. Kirschner M, Becker M, Wesarg S (2011) 3D active shape model segmentation with nonlinear shape priors. In: Medical image computing and computer-assisted intervention, Springer, Berlin, pp 492–499
25. Klinder T, Wolz R, Lorenz C, Franz A, Ostermann J (2008) Spine segmentation using articulated shape models. In: Medical image computing and computer-assisted intervention, Springer, Berlin, pp 227–234
26. Lecron F, Boisvert J, Mahmoudi S, Labelle H, Benjelloun M (2013) Three-dimensional spine model reconstruction using one-class SVM regularization. IEEE Trans Biomed Eng 60(11)
27. Mika S, Schölkopf B, Smola AJ, Müller KR, Scholz M, Rätsch G (1998) Kernel PCA and de-noising in feature spaces. Neural Inf Process Syst 11:536–542
28. Mitulescu A, Semaan I, Guise JAD, Leborgne P, Adamsbaum C, Skalli W (2001) Validation of the non-stereo corresponding points stereoradiographic 3D reconstruction technique. Med Biol Eng Comput 39(2):152–158
29. Moura DC, Boisvert J, Barbosa JG, Labelle H, Tavares JMR (2011) Fast 3D reconstruction of the spine from biplanar radiographs using a deformable articulated model. Med Eng Phys 33 (8):924–933
30. Papin P, Labelle H, Delorme S, Aubin CE, de Guise JA, Dansereau J (1999) Long-term three-dimensional changes of the spine after posterior spinal instrumentation and fusion in adolescent idiopathic scoliosis. Eur Spine J 8:16–16
31. Pennec X (1999) Probabilities and statistics on riemannian manifolds: basic tools for geometric measurements. Proc IEEE-EURASIP Workshop Nonlinear Signal Image Process 1:194–198
32. Pennec X, Thirion JP (1997) A framework for uncertainty and validation of 3D registration methods based on points and frames. Int J Comput Vis 25(3):203–229
33. Plamondon A, Gagnon M, Maurais G (1988) Application of a stereoradiographic method for the study of intervertebral motion. Spine 13(9):1027–1032
34. Pomero V, Mitton D, Laporte S, Guise JAD, Skalli W (2004) Fast accurate stereoradiographic 3D-reconstruction of the spine using a combined geometric and statistic model. Clin Biomech 19(3):240–247
35. Rasoulian A, Rohling R, Abolmaesumi P (2013) Lumbar spine segmentation using a statistical multi-vertebrae anatomical shape + pose model. In: IEEE transactions on medical imaging
36. Roberts MG, Cootes TF, Pacheco E, Oh T, Adams JE (2009) Segmentation of lumbar vertebrae using part-based graphs and active appearance models. In: Medical image computing and computer-assisted intervention, Springer, Berlin, pp 1017–1024

37. Roberts M, Oh T, Pacheco E, Mohankumar R, Cootes T, Adams J (2012) Semi-automatic determination of detailed vertebral shape from lumbar radiographs using active appearance models. Osteoporos Int 23(2):655–664
38. Schmidt S, Kappes J, Bergtholdt M, Pekar V, Dries S, Bystrov D, Schnorr C (2007) Spine detection and labeling using a parts-based graphical model. Inf Process Med Imaging 20:122–133
39. Smyth PP, Taylor CJ, Adams JE (1999) Vertebral shape: automatic measurement with active shape models. Radiology 211(2):571–578
40. Sun W, Xiang Yuan Y (2006) Optimization theory and methods: nonlinear programming. Springer, Berlin
41. Vijvermans V, Fabry G, Nijs J (2004) Factors determining the final outcome of treatment of idiopathic scoliosis with the boston brace: a longitudinal study. J Pediatr Orthop B 13:143–149
42. Vrtovec T, Likar B, Pernuš F (2005) Automated curved planar reformation of 3D spine images. Phys Med Biol 50(19):4527

Reconstruction of 3D Vertebral Models from a Single 2D Lateral Fluoroscopic Image

Guoyan Zheng and Lutz-P. Nolte

Abstract Accurate three-dimensional (3D) models of lumbar vertebrae are required for image-based 3D kinematics analysis. MRI or CT datasets are frequently used to derive 3D models but have the disadvantages that they are expensive, time-consuming or involving ionizing radiation (e.g., CT acquisition). In this chapter, we present an alternative technique that can reconstruct a scaled 3D lumbar vertebral model from a single two-dimensional (2D) lateral fluoroscopic image and a statistical shape model. Cadaveric studies are conducted to verify the reconstruction accuracy by comparing the surface models reconstructed from a single lateral fluoroscopic image to the ground truth data from 3D CT segmentation. A mean reconstruction error between 0.7 and 1.4 mm was found.

1 Introduction

Several studies have shown that fluoroscopy is well-suited to in vivo lumbar spine kinematics analysis due to its capability of screening patients during free motion with an acceptably low radiation dosage [1, 2]. The disadvantage of this technique, however, lies in its limitation to planar motion analysis. To enable fluoroscopic image-based 3D kinematic analysis, accurate three-dimensional (3D) models are needed [3, 4]. If kinematics of an implanted prosthesis is the interest, a Computer Aided Design (CAD) model can be used [5]. However, this is not the case for analyzing in vivo lumbar spine kinematics. Thus, MRI or CT datasets are frequently used to derive 3D models, but have the disadvantages that they are expensive, time-consuming or involving ionizing radiation (e.g., CT acquisition). In this paper, we present a technique to reconstruct a scaled 3D lumbar vertebral model from a single two-dimensional (2D) lateral fluoroscopic image.

G. Zheng (✉) · L.-P. Nolte
Institute for Surgical Technology and Biomechanics, University of Bern,
Stauffacherstrasse 78, 3014 Bern, Switzerland
e-mail: guoyan.zheng@ieee.org; guoyan.zheng@istb.unibe.ch

© Springer International Publishing Switzerland 2015

S. Li and J. Yao (eds.), *Spinal Imaging and Image Analysis*,
Lecture Notes in Computational Vision and Biomechanics 18,
DOI 10.1007/978-3-319-12508-4_11

Constructing a 3D surface model of the vertebra from 2D calibrated fluoroscopic image(s) is a challenging task. A priori information is often required to handle this otherwise ill-posed problem. Previously, kriging-based methods [6–11] as well as statistical shape model (SSM)-based methods [12–14] have been proposed. Unlike the methods in the former category, where one generic object is used as the prior information, the methods in the latter category use statistical shape models obtained from statistical shape analysis. Statistical shape analysis is an important tool for understanding anatomical structures from medical images [15–17, 32, 33]. Statistical shape models give efficient parameterization of the shape variations found in a collection of sample models of a given population. Model-based approaches are popular due to their ability to robustly represent objects [18, 19]. In Benameur et al. [12, 13], a SSM of scoliotic vertebrae was fitted to two radiographic views by simultaneously optimizing both shape and pose parameters. The optimal estimation was obtained by iteratively minimizing a combined energy function, which is the sum of a likelihood energy term measured from an edge potential field on the images and a prior energy term measured from the statistical shape model. Boisvert et al. [14] used a statistical articulated model of the spine for 3D reconstruction from partial radiographic data. Previously, we proposed a 2D-3D reconstruction scheme combining statistical instantiation and regularized shape deformation with an iterative image-to-model correspondence establishing algorithm, and showed its application to reconstruct a surface model of the proximal femur [20, 21].

Common to all these previous works is that at least two images are used as the input. Recently, we proposed a technique that could reconstruct a scaled, patient-specific 3D surface model from a standard X-ray radiograph and showed its application to reconstruct a surface model of the pelvis [22]. Based on this work, this paper presents an improved technique that combines a landmark-to-ray registration with a statistical shape model-based 2D/3D reconstruction scheme for reconstructing a scaled, patient-specific 3D surface model of the lumbar vertebra from a single fluoroscopic image. The landmark-to-ray registration is used to find an initial scale and an initial rigid transformation between the fluoroscopic image and the statistical shape model. The estimated scale and rigid transformation are then used to initialize the statistical shape model-based 2D/3D reconstruction scheme.

This chapter is organized as follows. Section 2 briefly presents the construction of the statistical shape model. Section 3 describes the statistically deformable 2D/3D reconstruction approach. Section 4 describes the experimental design and results, followed by the discussion and conclusions in Sect. 5.

2 Construction of a Statistical Shape Model of the Lumber Vertebrae

In this step, our goal was to construct a statistical shape model of the lumbar vertebrae, simultaneously considering shape information from all five lumbar levels, and thereby to determine the principal modes of shape variation. We chose

the Point Distribution Model (PDM) [19] as the representation of the SSMs of the lumbar vertebrae. The PDM was constructed from a training database consisting of 39 CT-segmentation based binary volumes of lumbar vertebrae (according to the spine level, the distribution of these 39 binary volumes are as follows, L1 level: 3; L2 level: 5; L3 level: 9; L4 level: 14; L5 level: 8). The PDM was constructed from the training data based on following procedure. First, a binary volume of a L3 level vertebra was chosen as the reference. Demon's algorithm [23] was used to estimate the deformation fields between the chosen reference binary volume and the other floating volumes. Each estimated deformation field was then used to displace the positions of the vertices on the reference surface model to the associated target volume. We thus obtained a set of aligned surface models with established correspondences.

Following the alignment, the PDM were constructed as follows. Let $x_i, i = 0, 1, \ldots .m-1$, be m (m = 39) members of the aligned training surface models. Each member is described by a vector $\mathrm{X_i}$ with N (N = 5000) vertices:

$$\mathrm{x}_i = \{x_0, y_0, z_0, x_1, y_1, z_1, \ldots, x_{N-1}, y_{N-1}, z_{N-1}\} \tag{1}$$

A PDM was then obtained by applying principal component analysis [24] to the aligned training surface models:

$$\begin{aligned} \mathbf{D} &= ((m-1)^{-1}) \cdot \sum_{i=0}^{m-1} (\mathrm{x}_i - \bar{\mathbf{x}})(\mathrm{x}_i - \bar{\mathbf{x}})^T \\ P &= (\mathbf{p}_0, \mathbf{p}_1, \ldots); \quad \mathbf{D} \cdot \mathbf{p}_k = \sigma_k^2 \cdot \mathbf{p}_k \end{aligned} \tag{2}$$

where $\bar{\mathbf{x}}$ and $\mathbf{D}$ are the mean vector and the covariance matrix of the PDM, respectively. $\{\sigma_k^2\}$ are non-zero eigenvalues of the covariance matrix $\mathbf{D}$, and $\{\mathbf{p}_k\}$ are the corresponding eigenvectors. The descendingly sorted eigenvalues σ_k^2 and the corresponding eigenvector $\mathbf{p}_k$ are the principal directions spanning a shape space with $\bar{\mathbf{x}}$ representing its origin. Figure 1 shows the variability captured by the first two modes of variations of the PDM.

3 Statistically Deformable 2D/3D Reconstruction

Without loss of generality, here we assume that the input image is calibrated and image distortion is corrected. For more details about fluoroscopic image calibration, we refer to our previous work [25]. Thus, for a pixel in the input image we can always find a projection ray emitting from the focal point of the image through the pixel.

The single image based surface model reconstruction technique proposed in this paper is based on a hybrid 2D/3D deformable registration process coupling a landmark-based scaled rigid registration with an adapted SSM-based 2D/3D reconstruction algorithm [20, 21]. Different from the situation in our previous works

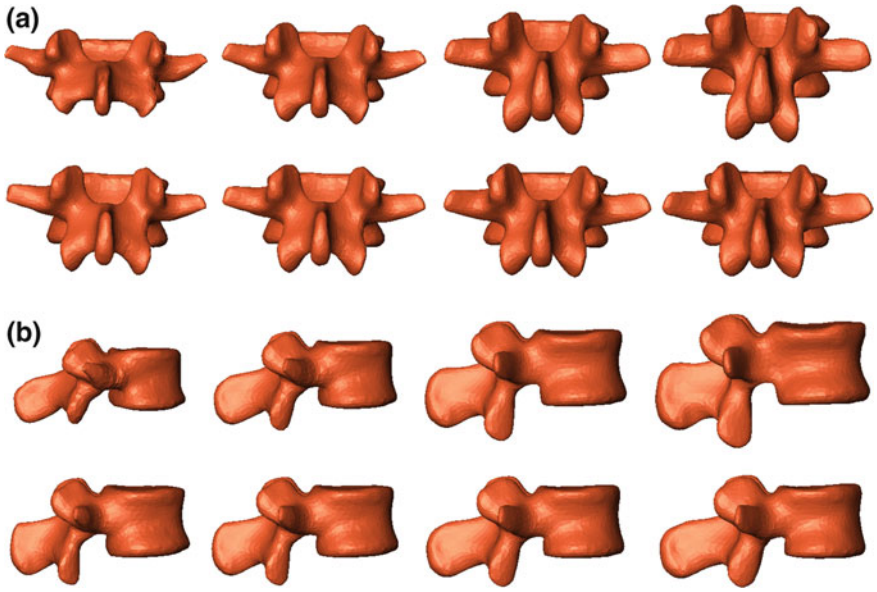


Fig. 1 The first two principal modes of variation of the PDM used in this investigation. The shape instances from *left* to *right* at each row were generated by evaluating $\bar{\mathbf{x}} + \alpha\sigma_k\mathbf{p}_k$, $\alpha \in \{-2, -1, 1, 2\}$ with (**a**) Posterior view of the PDM (k = 1: the *1st row*; k = 2: the *2nd row*); and (**b**) Lateral view of the PDM (k = 1: the *1st row*; k = 2: the *2nd row*)

[20, 21], where two or more calibrated X-ray images were required as the input for a successful reconstruction, here only a single lateral fluoroscopic image is available. Similar to the situation when multiple images are used, the convergence of the single image based 2D/3D reconstruction also depends on the initialization and on the image contour extraction. Thus, in the following we focus on the image contour extraction and on a landmark-based scaled rigid registration for initializing the single image based 2D/3D reconstruction.

3.1 Image Contour Extraction

As a feature-based 2D/3D reconstruction approach, our technique requires a prerequisite image contour extraction. Explicit and accurate contour extraction is a challenging task, especially when the shapes involved become complex or when the background of the image becomes complex. In this paper, we feel that it is a far better choice to provide the user with a tool that supports interactive segmentation but at the same time speeds up the tedious manual segmentation process and makes the results repeatable. This leads us to developing a semi-automatic segmentation tool.

Our semi-automatic segmentation tool is based on the Livewire algorithm introduced by Mortensen and Barrett [26]. In their paper, graph edges are defined as

the connection of two 8-adjacent image pixels. A local cost function is assigned to the graph edges to weight their probability of being included in an optimal path. In this work, we use two static feature components to form this cost function. The first component f_G is calculated from Canny edge detectors [27] at three different scales (the standard deviations of the Gaussian smoothing operator in these three scales are 1.0, 2.0, and 3.0, respectively) as follows.

Let's denote the edges extracted by the Canny edge detector at three different scales as $E^1(\mathbf{q})$, $E^2(\mathbf{q})$, and $E^3(\mathbf{q})$, respectively. $\{E^i(\mathbf{q});\ i = 1, 2, 3\}$ are defined as follows: if pixel $\mathbf{q}$ is a detected edge pixel at the ith scale, then $E^i(\mathbf{q}) = 1$; otherwise, it equals to zero. Let's further denote the gradient magnitudes at different scales as $G^1(\mathbf{q})$, $G^2(\mathbf{q})$, and $G^3(\mathbf{q})$, respectively. Then we have,

$$f_G(\mathbf{q}) = (1.0 - \frac{G^1(\mathbf{q})}{\max(G^1(\mathbf{q}))} \cdot E^1(\mathbf{q})) + (1.0 - \frac{G^2(\mathbf{q})}{\max(G^2(\mathbf{q}))} \cdot E^2(\mathbf{q})) + (1.0 - \frac{G^3(\mathbf{q})}{\max(G^3(\mathbf{q}))} \cdot E^3(\mathbf{q})) \quad (3)$$

According to Eq. (3), if $\mathbf{q}$ is not a detected edge pixel at the ith scale, a constant cost of 1.0 will be added to the cost function. Otherwise, the cost depends on the gradient magnitude: the bigger the magnitude, the smaller the cost.

The second component, the gradient direction $f_D(\mathbf{p}, \mathbf{q})$, is calculated according to the form proposed in the original paper [26], which is used to add a smoothness term to the contour definition by assigning high costs to sharp changes.

Finally, these two static features are combined by weighted summation to form a single statistic local cost function as follows

$$l(\mathbf{p}, \mathbf{q}) = 0.6 f_G(\mathbf{q}) + 0.4 f_D(p, \mathbf{q}) \quad (4)$$

where the weights for these two terms are empirically determined.

Based on the Livewire algorithm, the semi-automatic contour extraction starts with a seed point, which is interactively placed by the user with a click of the left mouse button. During the extraction, the user can add more seed points by clicking the left mouse button. A click of the right mouse button will finish the definition of one contour. After that, clicking the left mouse button again starts the extraction of a new contour. Figure 2 shows an example of how the livewire segmentation technique is used to extract contours from the input image.

3.2 Landmark-Based Scaled Rigid Registration for Initialization

Initialization here means to estimate the initial scale and the rigid transformation between the mean model of the PDM and the input fluoroscopic image. For this purpose, we have adopted an iterative landmark-to-ray scaled rigid registration. The

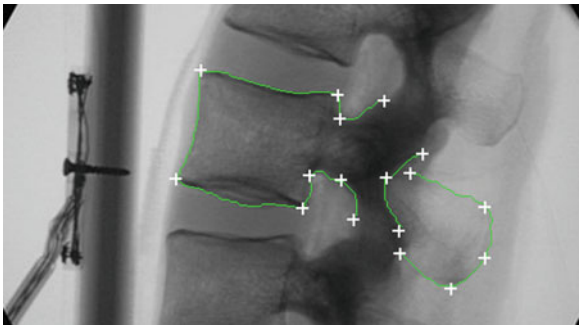


Fig. 2 Example of using livewire segmentation algorithm to extract image contours. The *white crosses* show where the user clicks the mouse button

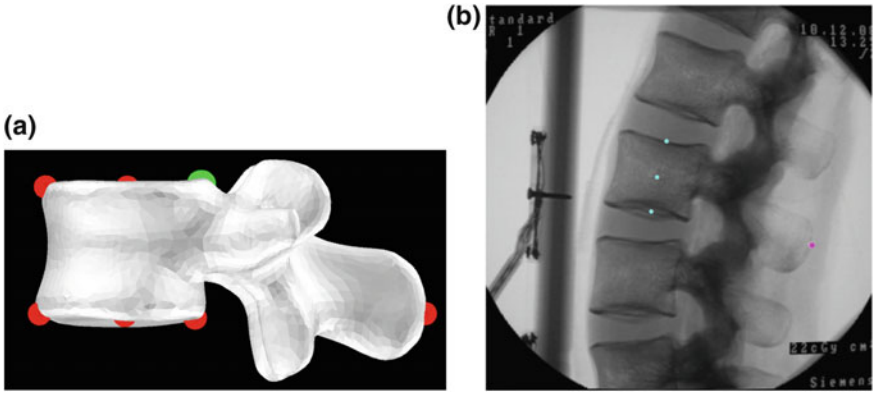


Fig. 3 Definition of initialization landmarks. **a** Landmarks extracted from the mean model of the PDM. **b** Landmarks extracted from the fluoroscopic images

four anatomical landmarks that we used here are the center of the top surface of the vertebra body, the center of the bottom surface of the vertebra body, the geometrical center of the vertebra body, and the center of the spinal process tip. Their positions on the mean model of the PDM are obtained through point picking or center calculation (the center of the vertebra body is computed as the center of four boundary landmarks along the anterior-posterior direction as shown in Fig. 3a, while their positions on the fluoroscopic image are defined through interactive picking (see Fig. 3a, b for details).

Let us denote those landmarks defined on the mean model of the PDM, i.e., the vertebra body center, the center of the top surface of the vertebra body, the center of the bottom surface of the vertebra body and the center of the spinal process tip, as v^1_{Mean}, v^2_{Mean}, v^3_{Mean}, and v^4_{Mean}, respectively; and their corresponding landmarks interactively picked from the fluoroscopic image as $v^1_{\mathrm{X-ray}}$, $v^2_{\mathrm{X-ray}}$, $v^3_{\mathrm{X-ray}}$, and

$v^4_{\mathrm{X-ray}}$, respectively. And for each X-ray landmark, we can calculate a projection ray emitting from the focal point to the landmark. We then calculate the length between v^1_{Mean} and v^4_{Mean} and denote it as $l^{1,4}_{Mean}$. Using the known image scale, we also calculate the length $l^{1,4}_{\mathrm{X-ray}}$ between $v^1_{\mathrm{X-ray}}$ and $v^4_{\mathrm{X-ray}}$. Then, we do:

Data Preparation. In this step, we assume that the line connecting the centers of the vertebra body and the center of the spinal process tip is parallel to the input fluoroscopic image and is certain distance away from the imaging plane. Using this assumption and the correspondences between the landmarks defined in the CT volume and those from the fluoroscopic image, we can compute two points $\bar{v}^1_{\mathrm{X-ray}}$ and $\bar{v}^4_{\mathrm{X-ray}}$ on the projection rays of $v^1_{\mathrm{X-ray}}$ and $v^4_{\mathrm{X-ray}}$, respectively (see Fig. 4a), which satisfy:

$$\begin{aligned} &\bar{v}^1_{\mathrm{X-ray}}\bar{v}^4_{\mathrm{X-ray}} // v^1_{\mathrm{X-ray}}v^4_{\mathrm{X-ray}};\ \textit{and} \\ &|\bar{v}^1_{\mathrm{X-ray}} - \bar{v}^4_{\mathrm{X-ray}}| = l^{1,4}_{\mathrm{X-ray}} \cdot \frac{F-d}{F} \end{aligned} \tag{3}$$

where "//" symbol indicates that the two lines are parallel; F is the calibrated distance from the focal point to the imaging plane and d is the assuming distance from the line connecting the center of the vertebra body and the center of the spinal process tip to the imaging plane.

The current scale s between the mean model and the input image is then estimated as,

$$s = |\bar{v}^1_{\mathrm{X-ray}} - \bar{v}^4_{\mathrm{X-ray}}| / l^{1,4}_{Mean} \tag{4}$$

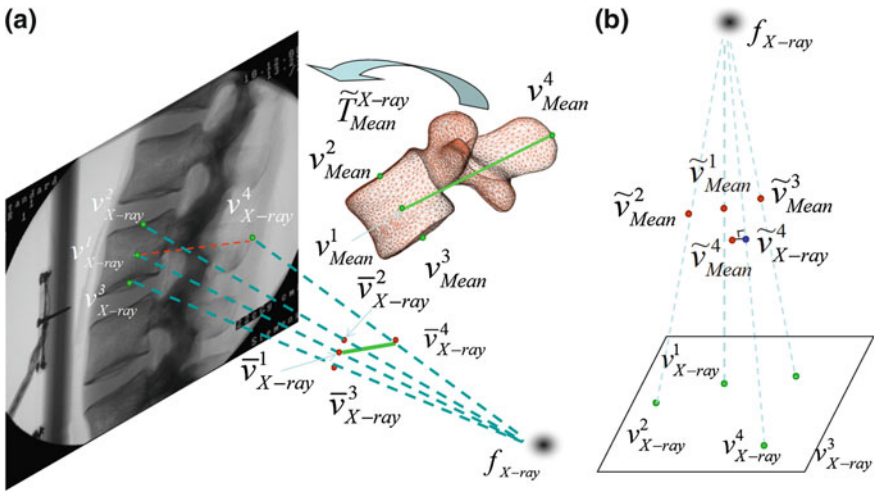


Fig. 4 Iterative landmark-to-ray registration. **a** Schematic view of data preparation. **b** Schematic view of finding 3D point pairs

Using s, we scale all landmark positions on the mean model and denote them as $\{\bar{v}^i_{Mean};\ i = 1,2,3,4\}$. We then calculate the distances from $\bar{v}^2_{Mean}$ and $\bar{v}^3_{Mean}$ to the line $\bar{v}^1_{Mean}\bar{v}^4_{Mean}$ and denote it as $\bar{l}^{2,1-4}_{Mean}$ and $\bar{l}^{3,1-4}_{Mean}$, respectively.

Next we find two points, point $\bar{v}^2_{\mathrm{X-ray}}$ on the projection ray of $v^2_{\mathrm{X-ray}}$ whose distance to the line $\bar{v}^1_{\mathrm{X-ray}}\bar{v}^4_{\mathrm{X-ray}}$ is equal to $\bar{l}^{2,1-4}_{Mean}$, and point $\bar{v}^3_{\mathrm{X-ray}}$ on the projection ray of $v^3_{\mathrm{X-ray}}$ whose distance to the line $\bar{v}^1_{\mathrm{X-ray}}\bar{v}^4_{\mathrm{X-ray}}$ is equal to $\bar{l}^{3,1-4}_{Mean}$. A paired-point matching based on $\{\bar{v}^i_{Mean}; i = 1,2,3,4\}$ and $\{\bar{v}^i_{\mathrm{X-ray}}; i = 1,2,3,4\}$ is used to calculate an updated scale s_0 and a rigid transformation $\bar{T}^{\mathrm{X-ray}}_{Mean}$ (see Fig. 4a for details). From now on, we assume that all information defined in the mean model coordinate frame has been transformed into the fluoroscopic image coordinate frame using s_0 and $\bar{T}^{\mathrm{X-ray}}_{Mean}$. We denote the transformed mean model landmarks as $\{\tilde{v}^i_{Mean};\ i = 1,2,3,4\}$.

Iteration. The following steps are iteratively executed until convergence:

- For a point $\tilde{v}^i_{Mean}$, we find a point on the corresponding projection ray of $v^i_{\mathrm{X-ray}}$ which has the shortest distance to the point $\tilde{v}^i_{Mean}$ and denote it as $\tilde{v}^i_{\mathrm{X-ray}}$ (see Fig. 4b). We then perform a paired-point matching using the extracted point pairs to compute a scale $\tilde{s}$ and a rigid transformation update $\Delta\tilde{T}^{\mathrm{X-ray}}_{Mean}$.
- We update the mean model coordinate frame using $\tilde{s}$ and $\Delta\tilde{T}^{\mathrm{X-ray}}_{Mean}$.

3.3 Statistical Shape Model-Based 2D/3D Reconstruction

The estimated scale and the rigid transformation between the mean model and the input image are then treated as the starting values for the PDM-based 2D/3D reconstruction scheme [20, 21], which depends on an iterative image-to-model correspondence establishing algorithm that we introduced previously [28]. The image-to-model correspondence is established using a non-rigid 2D point matching process, which iteratively uses a symmetric injective nearest-neighbor mapping operator and 2D thin-plate splines-based deformation to find a fraction of best matched 2D point pairs between those contours extracted from the x-ray image as we described above and the projections of the apparent contours extracted from the 3D model. The apparent contours of a statistically instantiated 3D model are extracted using the approach introduced by Hertzmann and Zorin [29]. Previously, we mathematically proved that the proposed non-rigid 2D point matching process could automatically eliminate the cross-matching event [28], which was defined as the interactions between the lines linking any matched point pair. Figure 5a shows the mean mode of the complete-PDM initialized with respect to the input image using the landmark-based scaled rigid registration, and the apparent contours extracted from the mean model. An example of building 2D/2D correspondences between the image contours and the projections of the apparent contours of the

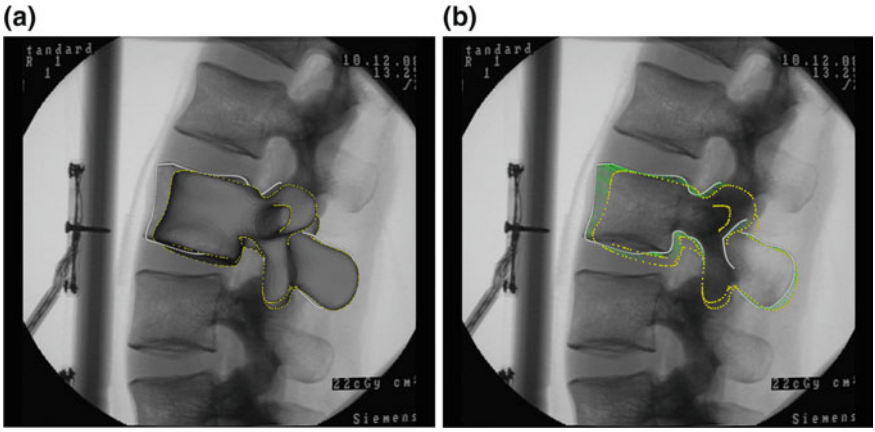


Fig. 5 Screenshots of establishing image-to-model correspondences. **a** the apparent contours (*yellow dots*) of the mean model (*dark grey*) of the complete-PDM after the landmark-based initialization. **b** 2D/2D correspondences (*green lines*) between the image contours (*white*) and the projections of the apparent contours

mean model as shown in Fig. 5a is presented in b. The obtained 2D point pairs are then used to set up a set of 3D point pairs so that we turn a 2D/3D reconstruction problem to a 3D/3D one. For details about how the proposed non-rigid 2D point matching process works and about the mathematic proof of how the proposed process eliminates the cross-matching event, we refer to our previous work [28]. In the following, the details about how to convert the 2D/3D reconstruction problem to a 3D/3D one and how the latter problem is solved are given for completeness.

3.3.1 Converting a 2D/3D Problem to a 3D/3D One

Assume that a set of 2D matched point pairs $\{(A_b, I_b); b = 0, 1, \ldots, n-1\}$ have been found, where A_b is the projection of a point on the apparent contours of a 3D model that is instantiated from the PDM and I_b is a point on the image contours that is matched to A_b; n is the number of point pairs. The corresponding 3D point pairs are then constructed as follows (see Fig. 6 for a schematic illustration). For a 2D point I_b, one can find a projection ray r_b emitting from the focal point of the X-ray image through the point I_b. Additionally, for its matched point A_b, one knows the associated 3D point Ω_b on the apparent contours of the model whose projection onto the image is A_b. By computing a point v_b on the rayr_b that has the shortest distance to Ω_b, a 3D point pair (Ω_b, v_b) can be obtained. Combining all these 3D point pairs, one can establish 2D/3D correspondence between the input image and a 3D model instantiated from the PDM, and thus convert a 2D/3D reconstruction problem to a 3D/3D one.

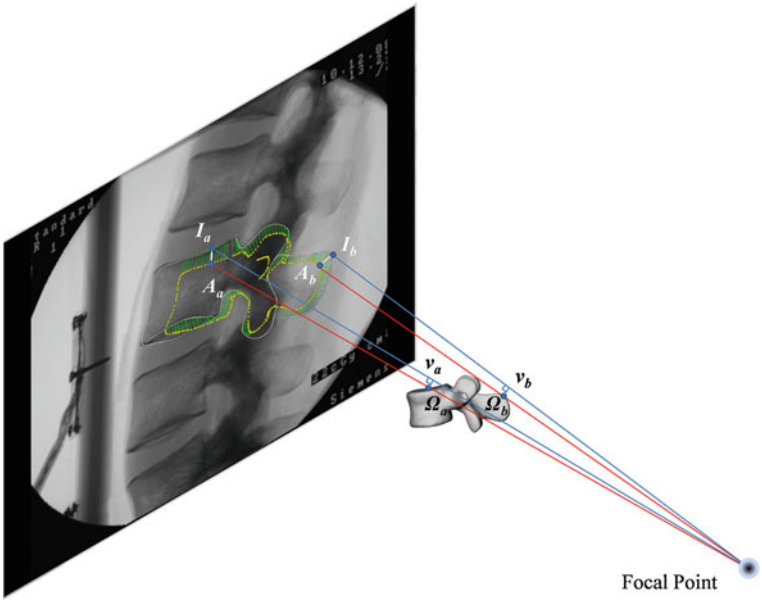


Fig. 6 Schematic illustration of computing 3D point pairs between a model and the input image from the established 2D/2D correspondences

3.3.2 3D/3D Reconstruction

As soon as a set of 3D point pairs are available, the problem of surface reconstruction is then solved optimally in three sequential stages using the algorithm presented in [28]: *scaled rigid registration*, *statistical instantiation*, and *regularized shape deformation*. For details about how to implement these three sequential stages, we refer to our previous work [28]. Figure 7 shows different stages of the reconstruction process, where a scaled surface model of the L2 lumbar vertebra was reconstructed from a single lateral fluoroscopic image.

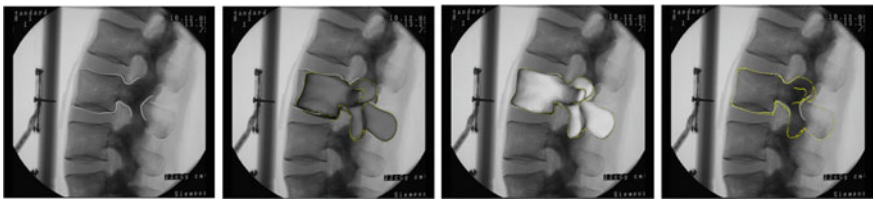


Fig. 7 Different stages of the reconstruction process. *Left* the smoothed contours; *left middle* the landmark-based initialization of the mean model (*grey*) of the PDM; *right middle* the reconstructed model (*white*); *right* the apparent contours (*yellow*) extracted from the reconstructed model versus the image contours (*white*)

4 Experimental Results

4.1 Experimental Design

We validated the present method on a single lateral fluoroscopic image of one cadaveric lumbar spine segment (there are totally four vertebrae in this segment but only three of them are visible in the image). All the binary volumes of the lumbar vertebrae contained in the test spine segment were semi-automatically segmented from the associated CT datasets using the commercially available software package Amira 5.0 (TGS Europe, Paris, France). To evaluate the reconstruction accuracy, a surface model derived from the binary volume of each test lumbar vertebra was used as the ground truth. As we only reconstructed a scaled surface model of the lumbar vertebra from each lateral fluoroscopic image, we had to first recover the unknown scale factor of the reconstructed model with respect to the ground truth surface model (or vice versa) before we could evaluate the reconstruction accuracy. For this purpose, we proposed to estimate the unknown scale factor of the reconstructed models by performing surface-based registrations [30]. After the registration, the open source tool MESH [31] was used to compute the distances between the reconstructed surface models and their associated ground truth surface models, which were regarded as the reconstruction errors. We adapted this tool to include the computation of different error statistics.

Two studies were conducted to evaluate the robustness and the accuracy of the present technique. Due to the reason that all 4 surface models of the lumbar vertebrae contained in the test spine segment were part of the training database for constructing the PDM, we named the first study as the leave-all-in study. In this study, each time the PDM as described in Sect. 2 was used together with the single lateral fluoroscopic image of the test spine segment to reconstruct a scaled surface model of a test vertebra. In the second study, all 4 aligned training surface models corresponding to the lumbar vertebrae in the test spine segment were removed from the training database and a PDM constructed from the rest 35 training surface models was used to reconstruct a scaled surface model of each test vertebra. We thus called the second study as the leave-four-out study. In both studies, two different surface-based registration techniques, i.e., a surface-based anisotropically-scaled rigid registration and a surface-based isotropically-scaled rigid registration, are used to estimate the unknown scale factors between the reconstructed surface models and the associated ground truth surface models.

For all experiments, we used an Intel Duo Core 2.4-GHz laptop with 4 GB of RAM. All programming was done using Visual C++ 2005 on Windows Vista.

4.2 Experimental Results

In both studies, the present technique could successfully reconstruct 3D surface models of all 3 test lumbar vertebrae. On average it took the present technique about 105 s to finish the computation. The errors of reconstructing surface models of all 3 lumbar vertebrae in both studies are shown in Table 1. A more detailed box-plot description of the reconstruction errors in both studies is shown in Fig. 8. When the surface-based anisotropically-scaled rigid registration was used to recover the unknown scale factors, an average mean reconstruction error of 0.77 mm (range: from 0.7 to 0.9 mm) was found for the leave-all-in study and an average mean reconstruction error of 0.83 mm (range: from 0.8 to 0.9 mm) was found for the leave-four-out study. In contrast, when the surface-based isotropically-scaled rigid registration was used, the average mean reconstruction error of the leave-all-in study was changed to 1.03 mm (range: 0.9–1.3 mm) and the average mean reconstruction error of the leave-four-out study was changed to 1.17 mm (range: 1.0–1.4 mm).

Figure 9 shows the surface model reconstruction accuracies of all three vertebrae in the leave-four-out study, where the ground truth models with a color-coded error distribution (middle column) are displayed together with the reconstructed surface models (right column) after surface-based anisotropically-scaled rigid registrations were used to recover the associated unknown scale factors of all three vertebrae.

5 Discussions and Conclusions

In this chapter, we presented a single image based 2D/3D reconstruction technique and showed its application to reconstruct a scaled, patient-specific 3D surface model of the lumbar vertebra from a single lateral fluoroscopic image. This single

Table 1 Errors of reconstructing the surface models of the 3 lumbar vertebrae

Vertebra	Cadaver_1_L1	Cadaver_1_L2	Cadaver_1_L3
Leave-all-in study, when the anisotropically-scaled rigid registration was used to recover the scale			
Errors (mm)	0.9 ± 0.7	0.7 ± 0.7	0.7 ± 0.6
Leave-one-out study, when the anisotropically-scaled rigid registration was used to recover the scale			
Errors (mm)	0.9 ± 0.8	0.8 ± 0.8	0.8 ± 0.7
Leave-all-in study, when the isotropically-scaled rigid registration was used to recover the scale			
Errors (mm)	1.3 ± 1.0	0.9 ± 0.8	0.9 ± 0.7
Leave-one-out study, when the isotropically-scaled rigid registration was used to recover the scale			
Errors (mm)	1.4 ± 1.1	1.0 ± 0.9	1.1 ± 0.8

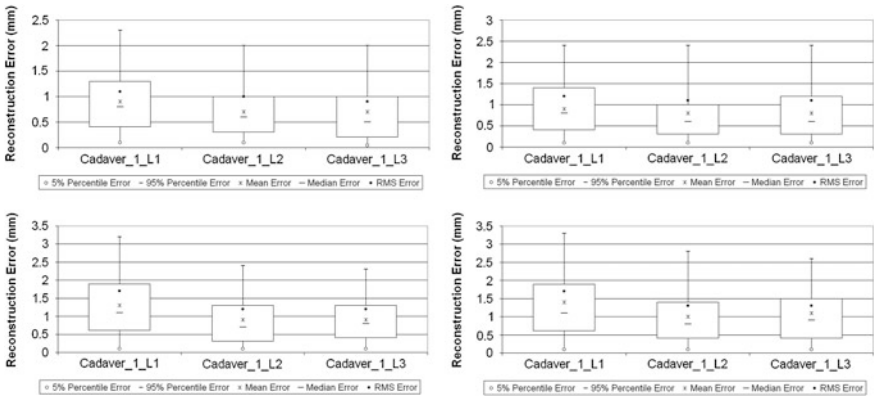


Fig. 8 Errors of reconstructing surface models of all 3 lumbar vertebrae when different surface-based matching techniques were used to recover the unknown scale factors between the reconstructed surface models and the associated ground truths. *Top row* when a surface-based anisotropically-scaled rigid registration was used; and *bottom row* when a surface-based isotropically-scaled rigid registration was used. In both rows, the *left column* shows results of the leave-all-in study while the *right column* shows the results of the leave-four-out study

image based 2D/3D reconstruction technique is based on a hybrid 2D/3D deformable registration process combining a landmark-to-ray registration with a SSM-based 2D/3D reconstruction. We validated the present 2D/3D reconstruction technique by designing and conducting two studies and in both studies the present technique could successfully reconstruct scaled surface models of all test lumbar vertebrae. To evaluate the overall reconstruction accuracy, we investigated two different surface-based registration techniques to recover the unknown scale factors between the reconstructed surface models and their associated ground truths: the surface-based anisotropically-scaled rigid registration and the surface-based isotropically-scaled rigid registration. Our experimental results demonstrated that the present technique can reconstruct scaled surface models of all 3 test lumbar vertebrae in a reasonably good accuracy, i.e., the mean reconstruction errors were found to be in the range of 0.7–1.4 mm. The overall reconstruction accuracy was slightly different when different surface-based registration techniques were used to estimate the unknown scale factors. It was also reasonable to observer that the results of the leave-all-in study were better than the other study. Such an observation indicated that the more shape variations that we integrated, the more accurate the present technique was.

The differences between the present technique and other works on reconstructing a patient-specific surface model of the vertebra should be discussed. Most of existing works [6–11, 14], except those introduced by Benameur et al. [12, 13], focused on the reconstruction of a surface models of the complete spine from two or more X-ray radiographs, while in this investigation we were only interested in reconstructing a surface model of the lumbar vertebra due to our targeted application, i.e., the spine kinematics analysis. The main difference between the present technique and the

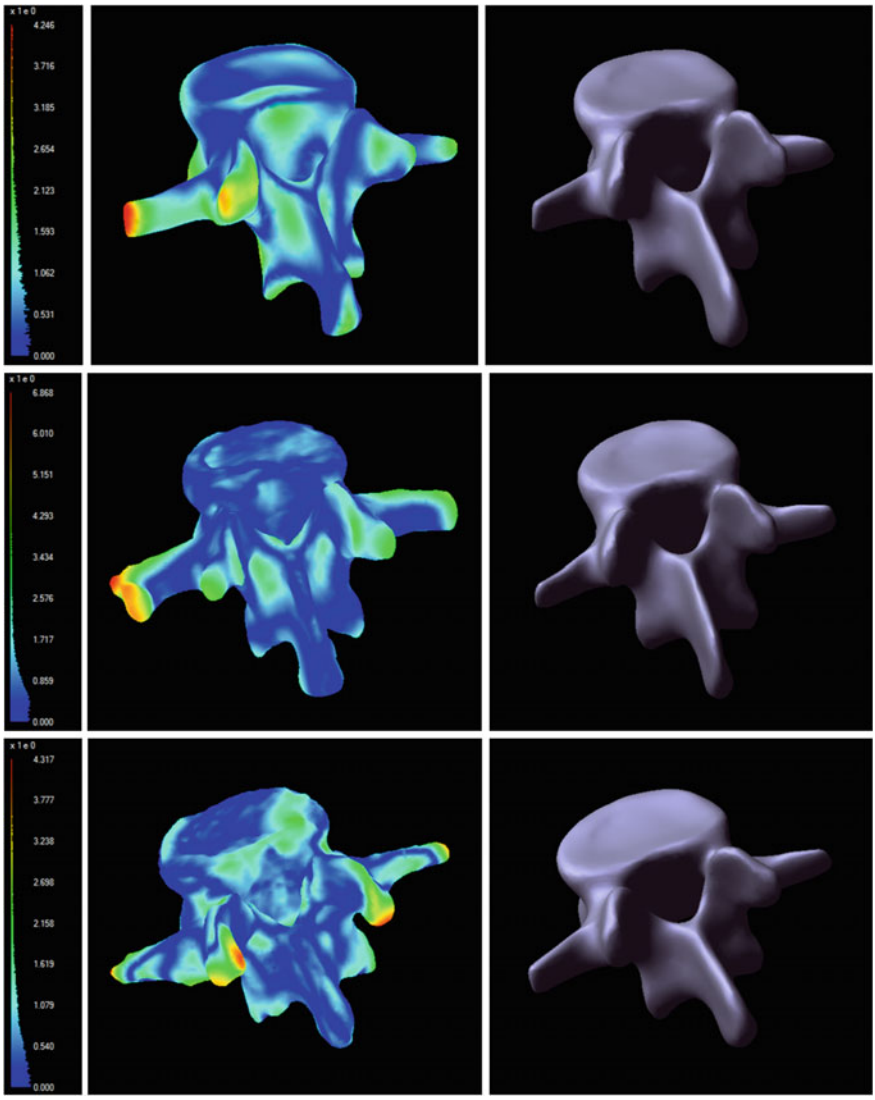


Fig. 9 Color-coded reconstruction error distribution. *Left column* error bar; *middle column* the ground truth models with the color-coded error distributions; *right column* the reconstructed models after surface-based iterative affine registrations were used to recover the unknown scale factors of the associated vertebrae

methods introduced by Benameur et al. [12, 13] lies in the optimization techniques that were used to reconstruct a patient-specific surface model. More specifically, in Benameur et al. [12, 13] a PDM of scoliotic vertebrae was fitted to two calibrated X-ray radiographs by simultaneously optimizing both the shape and the pose parameters while in this investigation we sequentially optimized the shape and the pose

parameters. Furthermore, in this investigation the surface model obtained after the statistical instantiation stage was further refined by the regularized shape deformation algorithm. The advantages of integrating this additional stage into the present technique over other existing attempts to instantiate a patient-specific surface model from a statistical shape model were explained in details in our previous work [28]. Briefly speaking, such integration enables the present technique to handle more complicated shape variation of any future instance [28].

While accurate, the present approach has limitations related with the number of training models used to construct the statistical shape models and the number of validation cases. The accuracy of the present approach depends not only upon how accurate the image-to-model correspondences can be established but also upon how well the unknown, patient-specific shape variation can be covered by the statistical shape model that is constructed from a fixed number of training models. Although the image-to-model correspondence establishing process has been thoroughly validated in our previous works [20–22, 28] as well as in this investigation, the statistical shape models used in the present study were constructed from a limited number of training lumbar vertebral models (39 for the leave-all-in study and 35 for the leave-four-out study). Furthermore, the validation of the present approach, though successful, was only conducted on datasets of 3 lumbar vertebrae. Thus, the results reported in this paper are regarded still preliminary and more thorough validation study is needed before it can be transferred to a routine usage. Nonetheless, the experiment results from the present study demonstrate the efficacy of the present approach and the prediction power of the present approach can be enhanced in the future by incorporating more training models into the statistical shape model and/or by constructing a patient-oriented statistical shape model.

Acknowledgments The authors gratefully acknowledge the financial support from the Swiss National Science Foundation through the National Centers of Competence in Research CO-ME. The test data are provided by Prof. Dr. S.J. Ferguson.

References

1. Muggleton JM, Allen R (1997) Automatic location of vertebrae in digitized videofluoroscopic images of the lumbar spine. Med Eng Phys 19:77–89
2. Zheng Y, Nixon MS, Allen R (2003) Lumbar spine visualization based on kinematic analysis from videofluoroscopic imaging. Med Eng Phys 25:171–179
3. Wang S, Passias P, Li G, Li G, Wood K (2009) Measurement of vertebral kinematics using noninvasive image matching method—validation and application. Spine 33:E355–E361
4. Bifulco P, Sansone M, Cesarelli M, Allen R, Bracale M (2002) Estimation of out-of-plane vertebra rotations on radiographic projections using ct data: a simulation study. Med Eng Phys 24:295–300
5. Hanson GR, Suggs JS, Freiberg AA, Durbhakula S, Li G (2006) Investigation of in vivo 6DOF total knee arthroplasty kinematics using a dual orthogonal fluoroscopic system. J Orthop Res 24:974–981

6. Aubin CE, Descrimes JL, Dansereau J, Skalli W, Lavaste F, Labelle H (1995) Geometrical modeling of the spine and thorax for biomechanical analysis of scoliotic deformities using finite element method. Ann Chir 49:749–761
7. Mitton D, Landry C, Vron S, Skalli W, Lavaste F, De Guise JA (2000) 3D Reconstruction method from biplanar radiography using non-stereocorresponding points and elastic deformable meshes. Med Biol Eng Comput 38:133–139
8. Pomero V, Mitton D, Laporte S, De Guise JA, Skalli W (2004) Fast accurate stereoradiographic 3D-reconstruction of the spine using a combined geometric and statistic model. Clin Biomech 19:240–247
9. Humbert L, De Guise JA, Aubert B, Godbout B, Skalli W (2009) 3D Reconstruction of the spine from biplanar X-rays using parametric models based on transversal and longitudinal inferences. Med Eng Phys 31:681–687
10. Dumas R, Blanchard B, Carlier R, de Loubresse CG, Le Huec J-C, Marty C, Moinard M, Vital J-M (2009) A semi-automated method using interpolation and optimization for the 3D reconstruction of the spine from bi-planar radiography: a precision and accuracy study. Med Biol Eng Comput 46:85–92
11. Kadoury S, Cheriet F, Labelle H (2009) Personalized X-ray 3D reconstruction of the scoliotic spine from hybrid statistical and image-models. IEEE Trans Med Imaging 28:1422–1435
12. Benameur S, Mignotte M, Parent S et al (2003) 3D/2D registration and segmentation of scoliotic vertebra using statistical models. Comput Med Imaging Grap 27:321–337
13. Benameur S, Mignotte M, Labelle H, De Guise JA (2005) A hierarchical statistical modeling approach for the unsupervised 3-D biplanar reconstruction of the scoliotic spine. IEEE Trans Biomed Eng 12:2041–2057
14. Boisvert J, Cheriet F, Pennec X, Labelle H, Ayache N (2008) Articulated spine models for 3-D reconstruction from partial radiographic data. IEEE Trans Biomed Eng 55:2565–2574
15. Kendall D (1989) A survey of the statistical theory of shape. Stat Sci 4:87–120
16. Small CG (1996) The statistical theory of shape., Series in StatisticsSpringer, New York
17. Dryden I, Mardia K (1998) Statistical shape analysis., Series in Probability and StatisticsWiley, New York
18. Turk M, Pentland A (1991) Eigenfaces for recognition. J Cogn Neurosci 3:71–86
19. Cootes TF, Hill A, Taylor CJ, Haslam J (1994) The use of active shape models for locating structures in medical images. Image Vis Comput 12:355–366
20. Zheng G, Gonzalez Ballester MA, Styner M, Nolte L-P (2006) Reconstruction of patient-specific 3D bone surface from 2D calibrated fluoroscopic images and point distribution model. In: MICCAI'06, vol part I, LNCS 4190, pp 25–32
21. Zheng G, Schumann S (2009) 3D reconstruction of a patient-specific surface model of the proximal femur from calibrated X-ray radiographs: a validation study. Med Phys 36:1155–1166
22. Zheng G (2009) Statistically deformable 2D/3D registration for accurate determination of post-operative cup orientation from single standard X-ray radiograph. In: MICCAI'09, part I, LNCS 5761, pp 820–827
23. Toussaint N, Souplet J-C, Fillard P (2007) Image matching as a diffusion process: an analogy with maxwell's demons. In: MICCAI'07 workshop on interaction in medical image and visualization
24. Jolliffe IT (2002) Principal component analysis, 2nd edn., Series in StatisticsSpringer, Berlin
25. Hofstetter R, Slomczykowski M, Sati M, Nolte L-P (1999) Fluroroscopy as an imaging means for computer-assisted surgical navigation. Comput Aided Surg 4:65–76
26. Barrett WA, Mortensen EN (1997) Interactive live-wire boundary extraction. Med Image Anal 1:331–341
27. Canny J (1986) A computational approach to edge detection. IEEE Trans Pattern Anal Mach Intell 8:679–714

28. Zheng G, Gollmer S, Schumann S, Dong X, Feilkas T, Gonzalez Ballester MA (2009) A 2D/3D correspondence building method for reconstruction of a patient-specific 3D bone surface model using point distribution models and calibrated X-ray images. Med Image Anal 13 (6):883–899
29. Hertzmann A, Zorin D (2000) Illustrating smooth surfaces. In: Proceedings of the 27th annual conference on computer graphics and interactive techniques (SIGGRAPH'00). ACM Press/Addison-Wesley Publishing Cp., New York, pp 517–526
30. Besl P, McKay ND (1992) A method for registration of 3D shapes. IEEE Trans Pattern Anal Mach Intell 14:239–256
31. Aspert N, Santa-Cruz D, Ebrahimi T (2002) MESH: measuring errors between surfaces using the Hausdorff distance. In: Proceedings of 2002 IEEE conference on multimedia and expo (ICME 2002), August 2002, Lausanne, Switzerland, vol 1, pp 705–708
32. Davies RH (2002) Learning shape: optimal models for analyzing natural variability. Doctoral dissertation, University of Manchester
33. Styner MA, Rajamani KT, Nolte L-P, Zsemlye G, Székely G, Taylor CJ, Davies RH (2003) Evaluation of 3D correspondences methods for model building. In: Proceeding of IPMI 2003, LNCS 2732, pp 63–75

Graphical Model-Based Vertebra Identification from X-Ray Image(s)

Xiao Dong and Guoyan Zheng

Abstract Automated identification of vertebrae from X-ray image(s) is an important step for various medical image computing tasks such as 2D/3D rigid and non-rigid registration. In this chapter we present a graphical model-based solution for automated vertebra identification from X-ray image(s). Our solution does not ask for a training process using training data and has the capability to automatically determine the number of vertebrae visible in the image(s). This is achieved by combining a graphical model-based maximum a posterior probability (MAP) estimate with a mean-shift based clustering. Experiments conducted on simulated X-ray images as well as on a low-dose low quality X-ray spinal image of a scoliotic patient verified its performance.

1 Introduction

Several studies have shown automated identification of vertebral bodies from medical image(s) is important for medical image processing tasks such as segmentation, registration, reconstruction and intervertebral disc identification. Due to the complexity of the spinal structure, simple feature (for example, landmarks or edges) based solutions are not reliable and researchers are paying more attention to graphical model-based solutions [1–5]. The reason to use graphical models lies in the following observations:

1. Human spine is a multi-component object with a stable anatomical structure. It is preferable to explore those structural constraints among components to achieve a joint identification of vertebral bodies or intervertebral discs rather than dealing with them independently.

X. Dong · G. Zheng (✉)
Institute for Surgical Technology and Biomechanics, University of Bern,
Stauffacherstrasse 78, 3014 Bern, Switzerland
e-mail: guoyan.zheng@ieee.org; guoyan.zheng@istb.unibe.ch

© Springer International Publishing Switzerland 2015

S. Li and J. Yao (eds.), *Spinal Imaging and Image Analysis*,
Lecture Notes in Computational Vision and Biomechanics 18,
DOI 10.1007/978-3-319-12508-4_12

2. Probabilistic graphical model is a general tool to model a multi-component structure like spine such that both the local feature information of each individual component and the constraints among components can be encoded in a model parameter space.
3. Probabilistic graphical model also enables various inference methods to find the optimal model parameters that can fit the model to the observation.

The current vertebral body or intervertebral disc identification approaches usually face the following challenges:

- *Unknown object number*. Detecting an unknown number of vertebrae or intervertebral discs invokes a model selection problem. In [1, 2], either the lumbar or the whole spine is investigated such that the number of intervertebral discs is taken as fixed. This is the reason why the authors can build their graphical models with a fixed number of nodes and avoid the model selection problem. In [6], the number of vertebrae is detected by a Generalized Hough Transform (GHT) along the detected spinal cord. The robustness of the exact number determination is highly dependent on the image quality.
- *Off-line training*. Due to the complexity of the spinal structure, most of the existing work on spine area asks for the involvement of prior knowledge which is usually obtained by off-line training. In [1, 2], both the low-level image observation models and the high-level disc context potentials need to be trained using training data. In [6], statistical surface models for each vertebra, the sacrum, the vertebra coordinate system and GHT models are obtained from the training data. Besides the fact that the model training and model building are complex problems themselves, the dependency on training data makes these approaches only applicable to the data with similar characteristics to the training data.

Our contributions in this chapter are: (1) firstly we designed a graphical model to model a spinal structure, which can adaptively determine the number of visible vertebrae during the inference procedure; (2) secondly, in the graphical model, both the low-level image observation model and the high-level vertebral context potentials need not to be learned from training data. Instead they are designed such that they can be learned from the target image data during the inference procedure.

2 Method

2.1 Graphical Model

Similar to [2], we build a graphical model $G = \{V, E\}$ with N nodes for the spinal structure as shown in Fig. 1. Each node $V_i, i = 0, 1, \ldots, N-1$ represents a connected disc-vertebra-disc component of the spinal structure, in which both the discs and the vertebral body are modelled as rectangular shapes. We assign a parameter

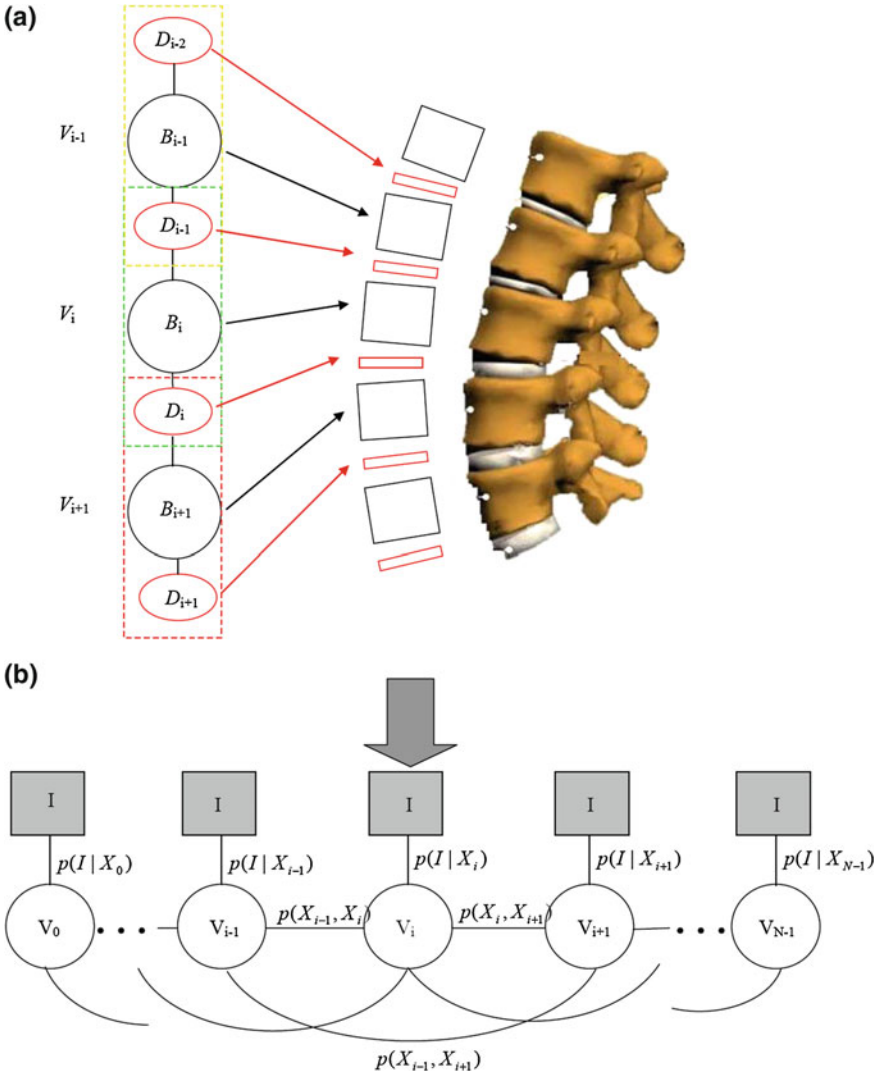


Fig. 1 A schematic view of the graphical model based representation of a spinal structure

set $\mathbf{X}_i = \{x_i, y_i, r_i, h_i, a_i(\theta_i), h_i^u, \theta_i^u, h_i^l, \theta_i^l\}$ to V_i to describe the positions, sizes and orientations of the vertebral body and the upper/lower intervertebral discs of V_i as shown in Fig. 2. $E = \{e_{i,j}\}, i,j = 0,1,2,\ldots,N-1$ define a connection matrix of the graph G. On this graphical model, the observation model of a single component V_i is defined as $p(\mathbf{I}|\mathbf{X}_i), i = 0,1,\ldots,N-1$ and the potential among neighboring components V_i and V_j with $e_{i,j} = 1$ is defined as $p(\mathbf{X}_i, \mathbf{X}_j), i,j = 0,1,\ldots,N-1$, $e_{i,j} = 1$. From a probabilistic point of view, $p(\mathbf{I}|\mathbf{X}_i)$ represents the probability that

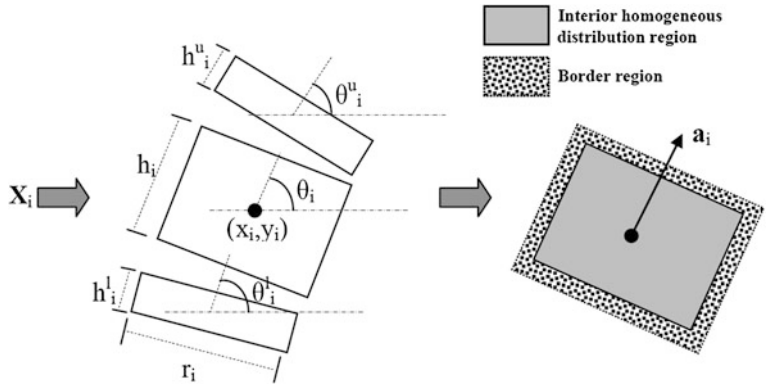


Fig. 2 A schematic view of the vertebral body template for the component observation model

the configuration $\mathbf{X}_i$ of the node $\mathbf{V}_i$ match the observed image(s) $\mathbf{I}$ and the potential $p(\mathbf{X}_i, \mathbf{X}_j)$ encodes the geometrical constraint between components. The identification of the spinal structure is then to find the configuration $\mathbf{X} = \{\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_i, \ldots, \mathbf{X}_{N-1}\}$, that maximizes.

$$p(\mathbf{X}|\mathbf{I}) \propto \prod_i p(\mathbf{I}|\mathbf{X}_i) \prod_{e_{i,j}=1} p(\mathbf{X}_i, \mathbf{X}_j). \tag{1}$$

2.2 Component Observation Model

The component observation model $p(\mathbf{I}|\mathbf{X}_i)$ is to match a template, which is determine by $\mathbf{X}_i$, with the observed image(s) **I.** We define our component observation model as:

$$p(\mathbf{I}|\mathbf{X}_i) = p_I(\mathbf{I}|\mathbf{X}_i) p_G(\mathbf{I}|\mathbf{X}_i) p_V(\mathbf{I}|\mathbf{X}_i). \tag{2}$$

The three items in Eq. (2) come from the intensity, gradient and local variance of the template as detailed below:

- *Intensity observation model* $p_I(\mathbf{I}|\mathbf{X}_i)$: The intensity observation model represents the probability that the parameterized model of V_i with the correspondent parameter set X_i fits the appearance of the observed image(s) $\mathbf{I}$. Each $\mathbf{X}_i$ determines a disc-vertebra-disc template as shown in Fig. 2. We assume that the interior area of the vertebral body has a homogeneous intensity distribution modeled as a Gaussian distribution $\mathcal{N}(\mu_i, \sigma_i)$. While the border region, which is defined as a small neighborhood outside the vertebral body as shown in Fig. 2, is assumed to obey a different intensity distribution from the interior area of the vertebral body. For each pixel s that falls in the interior and the border region of

the template with an intensity value $\mathbf{I}(s)$, the image appearance value of s is computed as

$$p(s|\mathbf{X}_i) = e^{-\frac{(\mathbf{I}(s)-\mu_i)^2}{2\sigma_i^2}}. \tag{3}$$

We further define $p_I(\mathbf{I}|\mathbf{X}_i) = e^{\omega_I c_I^i}$, where c_I^i is the cross-correlation between the image appearance values $p(s|\mathbf{X}_i)$ and a binary template which sets value 1 to the interior area of the template and 0 to the border region. $\omega_I > 0$ is a weighting factor. Intuitively this means that we assume that the interior region of the template should obey the Gaussian distribution and the border area should have a different intensity distribution. The Gaussian model $\mathcal{N}(\mu_i, \sigma_i)$ can be learned from the observed image(s) $\mathbf{I}$ once $\mathbf{X}_i$ is given, i.e., to fit a Gaussian distribution with the intensity values of the interior region of the vertebral body determined by $\mathbf{X}_i$.

- *Gradient observation model* $p_G(\mathbf{I}|\mathbf{X}_i)$: Similar to $p_I(\mathbf{I}|\mathbf{X}_i)$, we can define $p_G(\mathbf{I}|\mathbf{X}_i) = e^{\omega_G c_G^i}$, where c_G^i is the cross-correlation between the gradient image values of the observed image(s) in the template area and a binary gradient template, which sets 0 in the interior area and 1 in the border region. This means that the interior region of the vertebral body is homogeneous and high gradient values should only happen on the border of the vertebral template.
- *Local variance observation model* $p_V(\mathbf{I}|\mathbf{X}_i)$: We define the local variance image I_V of a pixel in the image(s) $\mathbf{I}$ as the intensity variance in a small window centered at this pixel. We set $p_V(\mathbf{I}|\mathbf{X}_i) = e^{\omega_V c_V^i}$, where c_V^i is the cross-correlation between the local variance values and a binary template identical to the gradient template. Similar to the gradient observation model, this item is used to model the observation that intensities of the interior area of a vertebral body should be more homogeneous than those of the border region.

We only consider the image observation model of the vertebral bodies but ignore the observation model of the discs. This is due to the fact that for X-ray image(s) with different view direction(s), the above mentioned observation model is more reliable for the vertebral bodies than for the discs. A unified observation model for the discs is more difficult to be designed.

It can also be observed that the three components in the observation model do not need to be trained with training data as done in [1, 2, 5]. Instead their parameters can be directly learned from the target X-ray image(s) **I.**

2.3 Potentials Between Components

We define inter-node potentials to apply geometric constraints between neighboring nodes such that all the nodes will be assembled to a meaningful spinal structure. More specifically, we have:

$$p(\mathbf{X}_i, \mathbf{X}_j) = p_S(\mathbf{X}_i, \mathbf{X}_j) p_O(\mathbf{X}_i, \mathbf{X}_j) p_D(\mathbf{X}_i, \mathbf{X}_j) \tag{4}$$

The three items in (4) specify the size, the orientation and the distance constraints as detailed below:

- *Size constraint* $p_S(\mathbf{X}_i, \mathbf{X}_j)$: $p_S(\mathbf{X}_i, \mathbf{X}_j)$ is used to set constraint on the sizes (radius and height of the vertebral body) of the neighboring components and is defined as

$$p_S(\mathbf{X}_i, \mathbf{X}_j) = e^{-\frac{(\omega_r \frac{|r_i - r_j|}{r_i + r_j} + \omega_h \frac{|h_i - h_j|}{h_i + h_j})}{|i-j|}} \tag{5}$$

  which means that neighboring components should have similar sizes and that the strength of the constraint should decay with the order distance between these two components.
- *Orientation constraint* $p_O(\mathbf{X}_i, \mathbf{X}_j)$: we define

$$p_O(\mathbf{X}_i, \mathbf{X}_j) = e^{-\frac{\omega_0 (a_i(\theta_i) \cdot a_j(\theta_j))}{|i-j|}} \tag{6}$$

  to ensure that the neighboring vertebral bodies should have similar orientations, in which $a_i(\theta_i)$ defines the orientation of a vertebral body template as shown in Fig. 2.
- *Distance constraint* $p_D(\mathbf{X}_i, \mathbf{X}_j)$: for direct neighboring nodes V_i, V_j, i.e., $|i-j| = 1$, we also define constraints on the spatial distance between their vertebral body centers. Without losing any generality, for the case $j = i + 1$ we define $p_D(\mathbf{X}_i, \mathbf{X}_j)$ as

$$p_D(\mathbf{X}_i, \mathbf{X}_j) = \begin{cases} e^{-\omega_D \frac{d_{C,ij}}{d_{h,ij}}}, & d_{C,ij} < \frac{d_{h,ij}}{4} \\ e^{-\omega_D \frac{d_{C,ij} - \frac{\left(d_{h,ij} + h_i^l + h_j^u\right)}{2}}{d_{h,ij}}}, & \frac{5}{4} d_{h,ij} > d_{C,ij} > \frac{3}{4} d_{h,ij} \\ 0, & \text{elsewhere} \end{cases} \tag{7}$$

  where $d_{C,ij} = |C_i - C_j|$, $d_{h,ij} = h_i + h_j$; C_i and C_j are the centers of the neighboring nodes V_i and V_j, respectively. The value of $p_D(\mathbf{X}_i, \mathbf{X}_j)$ with respect to the value of $d_{C,ij}/d_{h,ij}$ is shown in Fig. 3. It can be observed that $p_D(\mathbf{X}_i, \mathbf{X}_j)$ obtains a high value in the regions 0–0.25 and 0.75–1.25. Intuitively this constraint means that direct neighboring components should either be closely connected side-by-side (around the region when $d_{C,ij}/d_{h,ij}$ is in the range of 0.75–1.25 as shown in Fig. 3) or merge to one object (around the region when $d_{C,ij}/d_{h,ij}$ is in the range of 0.0–0.25 as shown in Fig. 3). This makes our graphical model capable of automatically adjusting the configuration of the

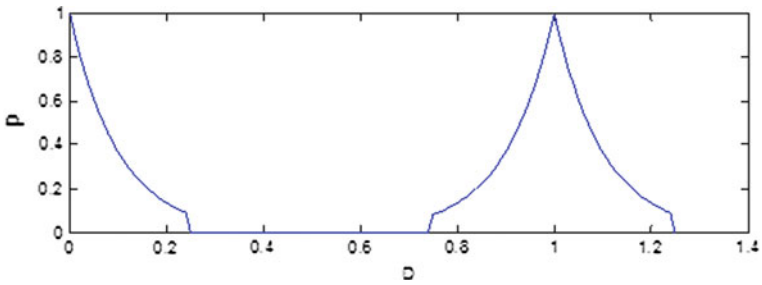


Fig. 3 Distance constraints $p_D(\mathbf{X}_i, \mathbf{X}_j)$ w.r.t. $d_{C,ij}/d_{h,ij}$, where a higher value of p_D indicates a configuration $(\mathbf{X}_i, \mathbf{X}_j)$ with a higher probability

nodes in the component chain to adaptively determine the number of vertebrae during the inference procedure. The details about the vertebra number determination will be explained in details in part 2.5.

2.4 Optimization

The optimization procedure aims to find the configuration $\mathbf{X} = \{\mathbf{X}_0, \mathbf{X}_1, \ldots, \mathbf{X}_i, \ldots, \mathbf{X}_{N-1}\}$ that maximizes

$$p(\mathbf{X}|\mathbf{I}) \propto \prod_i p(\mathbf{I}|\mathbf{X}_i) \prod_{e_{i,j}=1} p(\mathbf{X}_i, \mathbf{X}_j) \tag{8}$$

i.e., to obtain the MAP estimation of the model configuration that can fit the observed data.

In [1], the optimization is achieved by a generalized Expectation-Maximization (EM) algorithm given the known disc number and a proper initialization. In [2], the candidate configuration for each object can be detected by searching the whole data volume using trained random classification trees and the inference is achieved by the A* algorithm. Both of their optimization methods are not suitable for our graphical model. This can be explained briefly as follows. Firstly, we do not have a proper initialization as in [1]. Secondly, the configuration of each object in our case is high dimensional so that the complete search for candidate configurations of each object as presented in [2] is computationally expensive. In [5], the optimization is achieved by a joint application of a decision forest to detect the vertebral centers and a graphical model to refine the detection results.

Our optimization procedure to find the solution of Eq. (1) consists of two levels:

1. An inner iteration to find the configuration $\mathbf{X}_i$ of each individual component $\mathbf{V}_i$ by a particle filtering.

2. An outer iteration to find the joint configuration of the component set $\{\mathbf{X}_i, i = 0, 1, \ldots, N-1\}$ by a belief propagation (BP) based inference [7].

The inner iteration and outer iteration are described as follows:

Algorithm I Inner iteration to find the configuration $\mathbf{X}_i$ of $\mathbf{V}_i$
For an object V_i and its configuration parameters $\mathbf{X}_i$

1. Randomly generate a set of K configurations of V_i, $\mathbf{X}_i^k, k = 0, 1, \ldots, K-1$.
2. Compute the *belief* of each configuration as $\omega_i^k \propto p(\mathbf{I}|\mathbf{X}_i^k)$ based on the component observation model as defined in Eq. (1), (2) and (3). Obviously the K configurations of V_i, $\mathbf{X}_i^k, k = 0, 1, \ldots, K-1$ with their correspondent beliefs $\omega_i^k, k = 0, 1, \ldots, K-1$ can be regarded as a particle based non-parametric representation of the distribution $p(\mathbf{I}|X_i)$
3. Resample from the distribution $p(\mathbf{I}|X_i)$ to obtain new samples of the configuration of V_i, $\mathbf{X}_{i,new}^k, k = 0, 1, \ldots, K-1.$, which can be approximated by drawing samples from the distribution density $\{\omega_i^k\}$ with the correspondent configuration $\mathbf{X}_i^k$ followed by a random perturbation of the configuration $\mathbf{X}_i^k$.
4. Repeat 2–3 until the procedure converges.

Algorithm II Outer iteration (BP) to compute the joint distribution $p(\mathbf{X}|\mathbf{I})$
Given all the samples $X_i^k, i = 0, 1, \ldots, N-1, k = 0, 1, \ldots, K-1\}$ of $\{V_i, i = 0, 1, \ldots, N-1\}$ and the correspondent beliefs $\{\omega_i^k, i = 0, 1, \ldots, N-1, k = 0, 1, \ldots, K-1\}$

1. Taking the randomly generated configurations $\{\mathbf{X}_i^k\}$ as candidate configurations of each component and the beliefs $\{\omega_i^k\}$ as local beliefs, run a (loopy) belief propagation on the graphical model as shown in Fig. 1 to approximate the joint distribution $p(\mathbf{X}|\mathbf{I})$.
2. Compute the marginal distribution of each component as $\{\bar{\omega}_i^k\}$, which can be obtained from the distribution $p(\mathbf{X}|\mathbf{I})$.

The basic concept of our optimization algorithm is to combine the inner and outer iterations as shown in **Algorithm III**.

Algorithm III Optimization of the joint configuration

1. For each component $\mathbf{V}_i$, run the step 1 and 2 of the inner iteration to compute $\{\mathbf{X}_i^k, k = 0, 1, \ldots, K-1\}$ with their correspondent beliefs $\{\omega_i^k, k = 0, 1, \ldots, K-1\}$.
2. Run the outer iteration to compute the joint distribution $p(\mathbf{X}|\mathbf{I})$ and update marginal distribution $\{\bar{\omega}_i^k\}$.
3. For each component $\mathbf{V}_i$, run the step 3 of the inner iteration to update the configuration of each component.
4. Repeat 1–3 until the procedure converges.

The particle filter based inner iteration and the BP based outer iteration are combined such that

- The particle filtering in the inner iteration is used to find probable candidates for each individual component.
- The belief propagation (BP) in the outer iteration is used to set regularization on multiple components so that only the candidates that can fulfill the inter-component constraint will be selected.

In our algorithm, the stop criteria is set to run **Algorithm III** for a fixed number of iterations and the optimal configuration of $\mathbf{V}_i$ is set as the sample $\mathbf{X}_i^k$ with the highest probability ω_i^k.

2.5 The Determination of the Number of Vertebrae

Vertebral number determination is a key factor for correct vertebral body detection. In our approach we handle this problem by a semi-automatic method as described below:

- User clicks two landmarks on the X-ray image to indicate the centers of the first and the last visible vertebral bodies.
- From the locations of these two landmarks and the projection parameters of the X-ray image(s), we can estimate a minimal number of vertebrae N^{upper}, by dividing the distance between the two landmarks by an empirical estimation of the average vertebral height. Usually N^{upper} is greater than the true number N of the visible vertebrae in the image. We then construct a graph model with N^{upper} nodes.
- Carry out the inference procedure described in **Algorithm III**. Due to the distance potential between neighboring components as shown in Fig. 3, neighboring components will either be located side-by-side or overlap with each

other. Therefore, some vertebrae will merge with their neighbors, i.e., multiple nodes may be located very close to each other.

- After the optimization, a mean-shift based clustering on the central positions of the components using the mean height of the vertebral bodies as its bandwidth is used to merge overlapping components and therefore find the number of vertebrae [8].

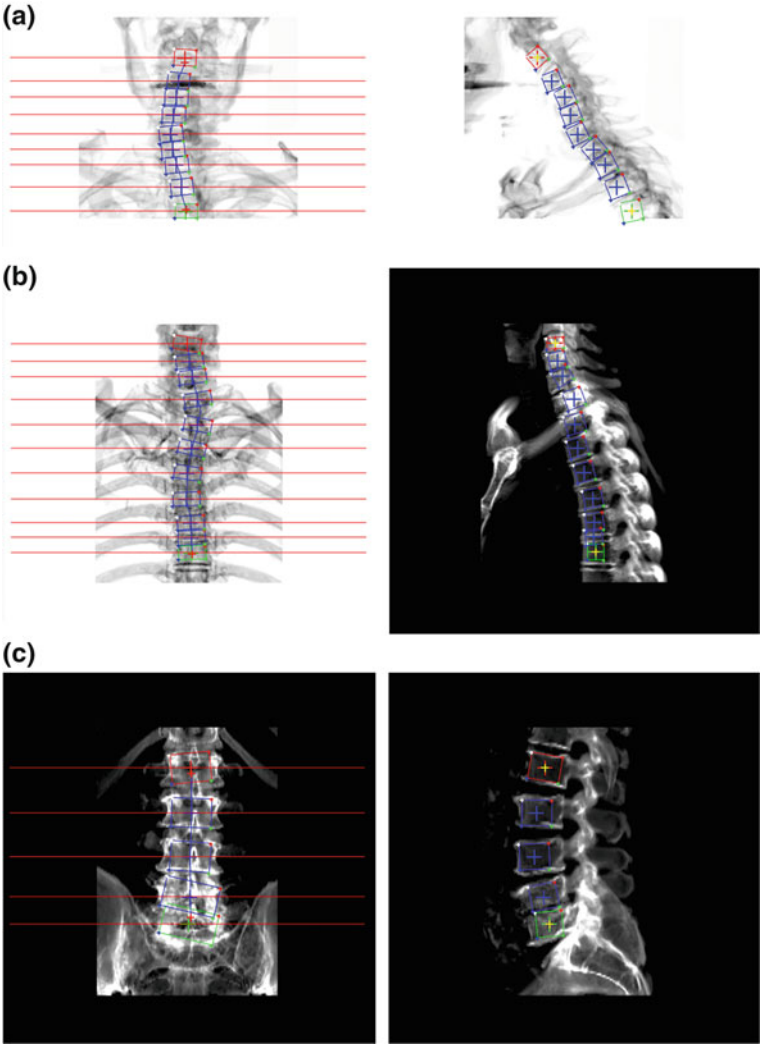


Fig. 4 Examples of detection vertebrae in different spinal regions. **a** An example of cervical vertebra detection. **b** An example of thoracic vertebra detection. **c** An example of lumbar vertebra detection

Table 1 Automated vertebral body detection results

Spine regions	Detection results	Image number	Vertebra number
Cervical vertebrae	Correct	6	38
	False/miss	2	7
Thoracic vertebrae	Correct	4	53
	False/miss	2	3
Lumbar vertebrae	Correct	7	31
	False/miss	0	0

- Using the detected vertebral number to construct a new graphical model and run the inference procedure as shown in **Algorithm III** to detect all vertebrae visible in the image(s).

3 Experimental Results

We validated the present approach on digitally reconstructed radiographs (DRRs) of twenty-one cadaver spinal segments, where eight of them were from cervical region, six of them were from thoracic region and the rest were from lumbar region, as well as one low-dose X-ray radiography of a scoliotic patient. The DRRs were constructed from the CT volumes of the associated spinal segments. For each CT volume, a pair of DRRs consisting of an anterior-posterior (AP) image and a lateral-medial (LM) image were generated. There are in total 132 vertebrae in the DRRs (45 cervical vertebrae, 56 thoracic vertebrae, and 31 lumbar vertebrae) and there are 13 vertebrae visible in the low-dose X-ray radiography.

For each pairs of DRRs, we started the detection from the LM image due to the observation that the vertebral bodies in the LM image were more homogeneous than those in the AP image. As soon as all the vertebrae were detected from the LM image, we could apply the same approach to the AP image but with a fixed number of the vertebrae that is determined from the LM image. For each detection, the user interactively specified two points as the input to our approach with one picked around the center of the top vertebra and the other around the center of the bottom vertebra. Our approach was then used to detect all vertebrae from the input image pair. The outputs from our approach include the number of vertebrae in the image, as well as the 3D location and orientation of each vertebra, which are reconstructed from the associated 2D detection results in both images. Figure 4 shows three examples of the automated detection of vertebrae in three different anatomical regions.

The automated vertebral body detection results are presented in Table 1. Although our approach had false/miss detection on four pairs of images, the false/miss vertebra detection rate was low. From the totally 132 vertebrae, our approach could successfully detect 122 vertebrae, which results in a 92.4 % success rate.

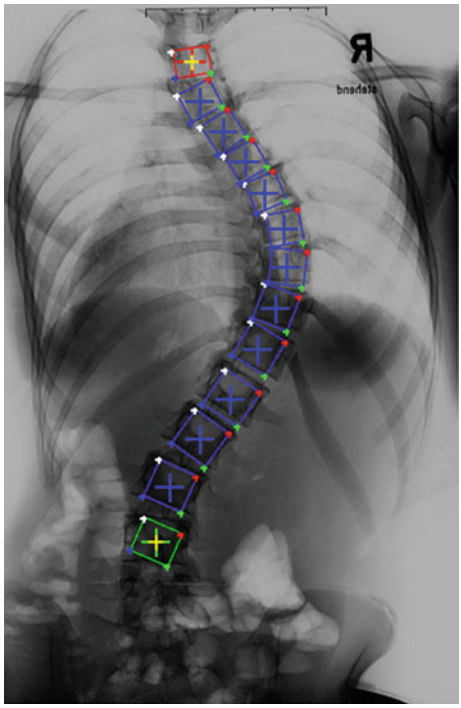


Fig. 5 Results of automatic vertebra detection from a low-dose X-ray radiograph of a scoliotic patient. All 13 visible vertebrae have been successfully detected by our algorithm

For the low-dose X-ray radiography of the scoliotic patient, our algorithm can successfully detect all 13 vertebrae. Figure 5 shows the detection results.

4 Discussions and Conclusions

In this chapter, different from previous work [9, 10], we proposed a graphical model-based method for automated detection of vertebral bodies from X-ray image (s). We validated our method on DRRs of twenty-one cadaver spinal segments of different regions as well as on one low-dose X-ray radiography of a scoliotic patient. Compared to previously introduced approach, our approach has following advantages: (1) it does not need to be trained using training data, (2) it does not ask for the prior information of the examined anatomical region and (3) it can automatically identify the number of vertebrae visible in the image(s) and therefore does not ask for a prior information about the number of vertebrae to be identified. Our future work focuses on investigating the performance of the proposed approach on more clinical X-ray images.

Acknowledgments The authors gratefully acknowledge the financial support from the Swiss National Science Foundation through the National Centers of Competence in Research CO-ME.

References

1. Corso J, Raja'S A, Chaudhary V (2008) Lumbar disc localization and labeling with a probabilistic model on both pixel and object features. In: MICCAI, pp 202–210
2. Schmidt S, Kappes J, Bergtholdt M, Pekar V, Dries S, Bystrov D, Schnoerr C (2007) Spine detection and labeling using a parts-based graphical model. IPMI 4584(2007):122–133
3. Oktay A, Akgul Y (2011) Localization of the lumbar discs using machine learning and exact probabilistic inference. In: Fichtinger G, Martel A, Peters T (eds) MICCAI. vol 6893 of LNCS. Springer, Berlin, pp 159–165
4. Kappes J (2011) Inference on highly-connected discrete graphical models with applications to visual object recognition. PhD thesis
5. Glocker B, Feulner J, Criminisi A, Haynor DR, Konukoglu E (2012) automatic localization and identification of vertebrae in arbitrary field-of-view CT scans. In: 15th international conference on medical image computing and computer assisted intervention (MICCAI)
6. Klinder T, Ostermann J, Ehm M, Franz A, Kneser R, Lorenz C (2009) Automated model-based vertebra detection, identification, and segmentation in CT images. Med Image Anal 13:471–482
7. Murphy KP, Weiss Y, Jordan MI (1999) Loopy belief propagation for approximate inference: an empirical study. In: Proceedings of uncertainty in AI, pp 467–475
8. Comaniciu D, Meer P (2002) Mean shift: a robust approach toward feature space analysis. IEEE Trans Pattern Anal Mach Intell 24:603–619
9. Rasoulian A, Rohling RN, Abolmaesumi P (2014) Automatic labeling and segmentation of vertebrae in CT images. In: Medical imaging 2014: image-guided procedures, robotic interventions, and modeling, San Diego, California, USA,15 Feb 2014
10. Aslan et al (2011) A new shape based segmentation framework using statistical and variational methods. In: Proceedings of 2011 IEEE international conference on image processing (ICIP), pp 717–720

Model-Based Segmentation, Reconstruction and Analysis of the Vertebral Body from Spinal CT

Melih Aslan, Ahmed Shalaby, Asem Ali and Aly A. Farag

Abstract In this chapter, we present novel vertebral body segmentation methods in computed tomography (CT) images. Three pieces of information (intensity, spatial interaction, and shape) are modeled to optimize new probabilistic energy functions; and hence to obtain the optimum segmentation. The information of the intensity and spatial interaction are modeled using the Gaussian and Gibbs distribution, respectively. A shape model is proposed using new probabilistic functions to enhance the segmentation results. The models are generic shape information which is obtained using the cervical, lumbar, and thoracic spinal regions. The proposed methods are validated with clinical CT images and on a phantom with various Gaussian noise levels. This study reveals that the proposed methods are robust under various noise levels, less variant to the initialization, and quite faster than alternative methods. Applications on bone mineral density (BMD) measurements of vertebral body are given to illustrate the accuracy of the proposed segmentation approach.

1 Introduction

Isolating an organ from its surrounding anatomical structures is a crucial step in many unsupervised frameworks. Examples of these frameworks are those that assess the organ functions and those that are proposed for automatic classification of normal organ and acute rejection transplants. In this work, we propose segmentation frameworks for spine bone [more specifically the Vertebral Body (VB)].

M. Aslan · A. Shalaby · A.A. Farag (✉)
Computer Vision and Image Processing Laboratory, University of Louisville, Louisville, KY, USA
e-mail: aly.farag@louisville.edu

M. Aslan
Computer Science Department, Wayne State University, Detroit, MI, USA

A. Ali
Electrical Engineering Department, Assiut University, Assiut, Egypt

© Springer International Publishing Switzerland 2015

S. Li and J. Yao (eds.), *Spinal Imaging and Image Analysis*,
Lecture Notes in Computational Vision and Biomechanics 18,
DOI 10.1007/978-3-319-12508-4_13

Segmentation can be defined as partitioning the image into the meaningful areas using the existing (low level) information in the image and prior (high level) information which can be obtained using a number of features of an object. The human vision system aims to extract and use as much as possible information in the image. The possible information includes the intensity, possible motion of the object (in sequential images), spatial relations (interaction) as the existing information, and the shape of the object which is learnt from the experience as the prior information. The machine visual system cannot predict the prior information unless it is supplemented. Hence, any prior cue can be specified beforehand to enhance the segmentation or to obtain the desired segmentation. If the prior information of the object is not given beforehand to the machine vision task, the segmentation method may not give desired results due to noise, occlusion, and missing information in the image.

One of the bone diseases, which is characterized by a reduction in bone mass, is Osteoporosis. This disease results in an increased risk of fractures. Bone Mineral Density (BMD) measurements and Fracture Analysis (FA) of the VBs should be obtained to accurately diagnose the osteoporosis. To obtain these measurements and analysis, VBs should be accurately segmented, which is our main objective in this work.

Since BMD measurements and fracture analysis are restricted to the vertebral bodies, segmentation approaches should successfully isolate VB from processes, which constitute spine bone as shown in Fig. 1. However, due to region inhomogeneities existing in CT images, isolating a VB form its background is not an easy task as shown in Fig. 2. To overcome these inhomogeneities and accurately segment VBs, we use both shape and appearance information.

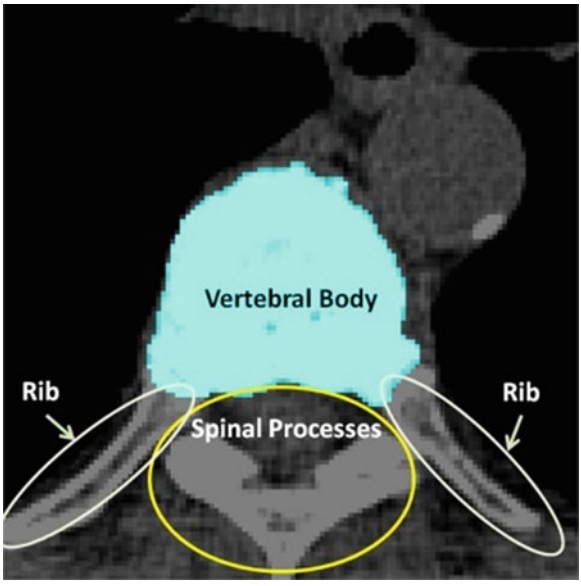


Fig. 1 Vertebral Body: A CT slice of vertebral body illustrates its regions: the VB region (in *blue color*), which is our region of interest to be segmented. Spinal processes and ribs, which should not be included in the BMD measurements, are shown as well

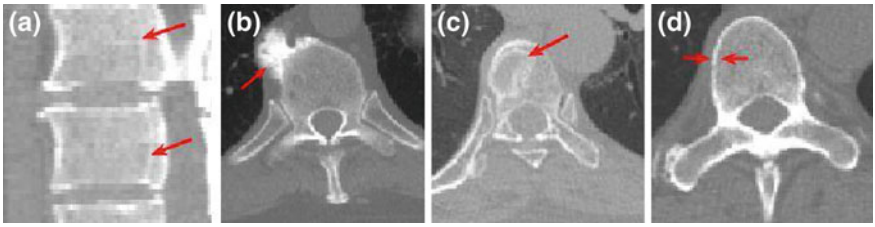


Fig. 2 Typical challenges for vertebrae segmentation. **a** Inner boundaries. **b** Osteophytes. **c** Bone degenerative disease. **d** Double boundary

The literature is rich with organ segmentation techniques. However, we will discuss only some of these techniques whose basics depend on shape modeling and whose application is VB segmentation. To tackle the problem of segmenting a spine bones, various approaches have been introduced. For instance, Klinder et al. [1] developed an automated model-based vertebra detection, identification and segmentation approach. Kang et al. [2] developed a 3D segmentation method for skeletal structures from CT images. Their method starts with a three dimensional region growing step using local adaptive thresholds. Then a closing of boundary discontinuities and an anatomically-oriented boundary adjustment steps are done. They presented various anatomical bony structures as applications. They evaluated their segmentation accuracy using the European Spine Phantom (ESP) [3]. In order to measure bone mineral density, Mastmeyer et al. [4] presented a hierarchical segmentation approach for the lumbar spine. They reported that it takes less than 10 min to analyze three vertebrae, which is a huge improvement compared to what is reported in [5]: 1–2 h. However, this timing is far from the real time required for clinical applications. To analyze the fracture of VBs, Roberts et al. [6] used the active appearance model. Other techniques have been developed to segment skeletal structures and can be found for instance in [7–9].

Actually, there are a huge number of segmentation techniques in the literature: simple techniques (e.g. region growing or thresholding), parametric deformable models and geometrical deformable models. However, all these methods tend to fail in the case of noise, gray level inhomogeneities, and diffused boundaries. Organs have well-constrained forms within a family of shapes. Therefore segmentation algorithms have to exploit the prior knowledge of shapes and other properties of the structures to be segmented. Leventon et al. [10] combined the shape and deformable model by attracting the level set function to the likely shapes from a training set specified by principal component analysis (PCA). To make the shape guides the segmentation process, Chen et al. [11] defined an energy functional, which basically minimizes an Euclidean distance between a given point and its shape prior. Huang et al. [12], combined registration with segmentation in an energy minimization problem. The evolving curve is registered iteratively with a shape model

using the level sets. They minimized a certain function to estimate the transformation parameters.

In this chapter, we present universal and probabilistic shape based segmentation methods that are less variant to the initialization. Contribution of this chapter can be generalized as follows:

- This chapter solves problems caused by noise, occlusion, and missing information of the object by integrating the prior shape information.
- In this chapter, the conventional shape based segmentation results are enhanced by proposing a new probabilistic shape models and a new energy functional to be minimized. The shape variations are modelled using a probabilistic functions.
- The proposed shape based segmentation method is less variant to the initialization.
- To optimize the energy functional, the original ICM method, which was originally proposed by Besag [13], is extended by integrating the shape prior. With integrating the shape model to the original ICM method, possible local minimums of the energy functional are eliminated as much as possible, and enhance the results.
- One of the most important contributions of this study is to offer a segmentation framework which can be suitable to the clinical works with acceptable results. If the proposed method in this study is compared most published bone segmentation methods, the large execution time is reduced effectively.
- Many works are restricted to the specific regions of spine bone column as such lumbar, thoracic, and others. In this study, there is no any region restriction, and the proposed framework is processed on different regions.
- The proposed framework and the new probabilistic shape model extract the spinal processes and ribs which should not be included in the bone mineral density measurements.
- This work is not dependent on any identification step thanks to the new universal shape model and its embedding step.

Next section details the proposed methods.

2 Proposed Framework

In this section, we describe the proposed methods. First, the general theoretical idea of the proposed frameworks is described. Then, two pre-processing steps, the spinal cord extraction and VB separation, are described. We present three methods which differ mostly in the shape modeling and optimization parts.

2.1 Overview

In this chapter, three pieces of information (intensity, spatial interaction, and shape) are modelled to obtain the optimum segmentation. The data is assumed to have two classes: background and object regions which are represented as "*b*" or "0" and "*o*" or "1", respectively. So, let $\mathcal{L} = \{0, 1\}$ denotes the set of labels. In this work, the given VB's volume, the shape model and the desired map (labeled volume) are described by a joint Markov-Gibbs random field (MGRF) model. We can define the gray level volume $\mathbf{I}$ by the mapping $\mathcal{P} \rightarrow \mathbf{G}$ and its desired map $\mathbf{f}$ by the mapping $\mathcal{P} \rightarrow \mathcal{L}$, where $\mathcal{P}$ is the set of voxels and $\mathbf{G}$ is the set of gray levels. Shape information is represented by the set of distances of variability region's voxels $\mathbf{d}$ (more details explained later). Since $\mathbf{I}$ and $\mathbf{d}$ are independent, a conditional distribution model of input volume, its desired map, and the shape constraints can be written by the as follows:

$$P(\mathbf{f}|\mathbf{I}, \mathbf{d}) \approx P(\mathbf{I}|\mathbf{f})P(\mathbf{d}|\mathbf{f})P(\mathbf{f}), \tag{1}$$

where $P(\mathbf{I}|\mathbf{f})$ and $P(\mathbf{f})$ represents appearance models, and the conditional distribution $P(\mathbf{d}|\mathbf{f})$ is the shape model. Given $\mathbf{I}$ and $\mathbf{d}$, the map $\mathbf{f}$ can be obtained using Bayesian *maximum-a posteriori* estimate as follows:

$$\mathbf{f}^* = \arg\max_{\mathbf{f} \in \mathcal{F}} L(\mathbf{I}, \mathbf{d}, \mathbf{f}), \tag{2}$$

where $\mathcal{F}$ is the set of all possible $\mathbf{f}'\mathbf{s}$, and $L(\mathbf{I}, \mathbf{d}, \mathbf{f})$ is the log-likelihood function, which can be written as follows:

$$L(\mathbf{I}, \mathbf{d}, \mathbf{f}) \propto \log P(\mathbf{d}|\mathbf{f}) + \log P(\mathbf{I}|\mathbf{f}) + \log P(\mathbf{f}). \tag{3}$$

The parameters of the shape model $P(\mathbf{d}|\mathbf{f})$ and the volume appearance models should be identified, to completely define this log-likelihood function.

Intensity and interaction models may not be enough to obtain optimum segmentation. To segment the VB, a new shape based methods which integrate the models of the intensity, spatial interaction, and shape prior information is proposed. The proposed method presents several advantages which can be written as: (i) the probabilistic shape model is automatically registered to the testing image, hence manual interaction is eliminated, (ii) the registration benefits from the segmented region to be used in the shape representation, and (iii) the probabilistic shape model refines the initial segmentation result using the registered variability volume.

The segmentation part has following steps: (1) initial segmentation using only intensity and spatial interaction information (this step is needed to obtain the feature correspondence between the image domain and shape model), (2) shape model registration, and (3) the final segmentation using three models.

Next section, we describe the spinal cord extraction stage as a preprocessing step.

2.2 Spine Cord Extraction

As a pre-processing step, the spinal cord is extracted using the Matched filter. This process helps to remove the spinal processes roughly; hence the shape model will be registered to the image domain easily. In the first step, the Matched filter (MF) [14, 15] is employed to detect the VB automatically. This procedure eliminates the user interaction and improves the segmentation accuracy. Let $f(x, y)$ and $g(x, y)$ be template and test images, respectively. To compare the two images for various possible shifts τ_x and τ_y, one can compute the cross-correlation $c(\tau_x, \tau_y)$ as

$$c(\tau_x, \tau_y) = \int \int g(x, y) f(x - \tau_x, y - \tau_y) dxdy, \tag{4}$$

where the limits of integration are dependent on $g(x, y)$. The Eq. (4) can also be written as

$$\begin{aligned} c(\tau_x, \tau_y) &= \int \int G(f_x, f_y) F^*(f_x, f_y) \exp(j2\pi(f_x\tau_x + f_y\tau_y)) df_x df_y \\ &= FT^{-1}(G(f_x, f_y) F^*(f_x, f_y)), \end{aligned} \tag{5}$$

where $G(f_x, f_y)$ and $F(f_x, f_y)$ are the 2-D FTs of $g(x, y)$ and $f(x, y)$, respectively with f_x and f_y denoting the spatial frequencies. The test image $g(x, y)$ is filtered by $H(f_x, f_y) = F^*(f_x, f_y)$ to produce the output $c(\tau_x, \tau_y)$. Hence, $H(f_x, f_y)$ is the correlation filter which is the complex conjugate of the 2-D FT of the reference image $f(x, y)$. Figure 3a shows the reference image used in the Matched filter. Some examples of the VB detection are shown in Figs. 3b–d. The Matched filter is tested using 4,000 clinical CT images. The detection accuracy for the VB region is 97.2 %. The detection accuracy is increased to around 100 % by smoothing all detected points of a dataset in the z-axis.

To extract the spinal processes and ribs roughly, some simple steps are followed as shown in Fig. 4. These steps are required to: (i) extract the spinal processes and ribs roughly, (ii) crop the ROI minimize the execution time. Figures 5 and 6 show different examples of this stage in the sagittal view.

2.3 Vertebrae Separation

This process is required in the proposed framework hence the shape model is registered to each VB in an easy way. In this process, two methods are suggested to separate each vertebrae. The first suggestion is the manual separation, and the second one is the previously proposed automatic framework [16] as shown in Fig. 7c.

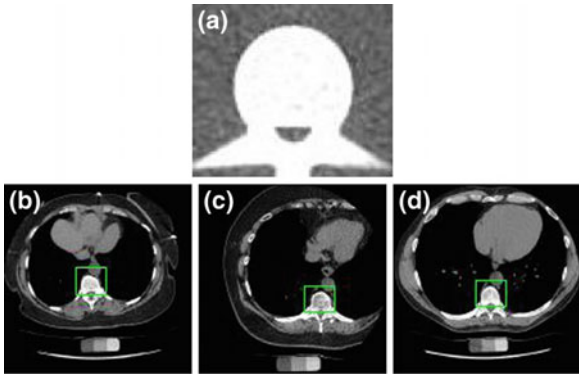


Fig. 3 **a** The template used for the Matched filter. **b–d** A few images of automatic VB detection. The *green line* shows the detection of VB region

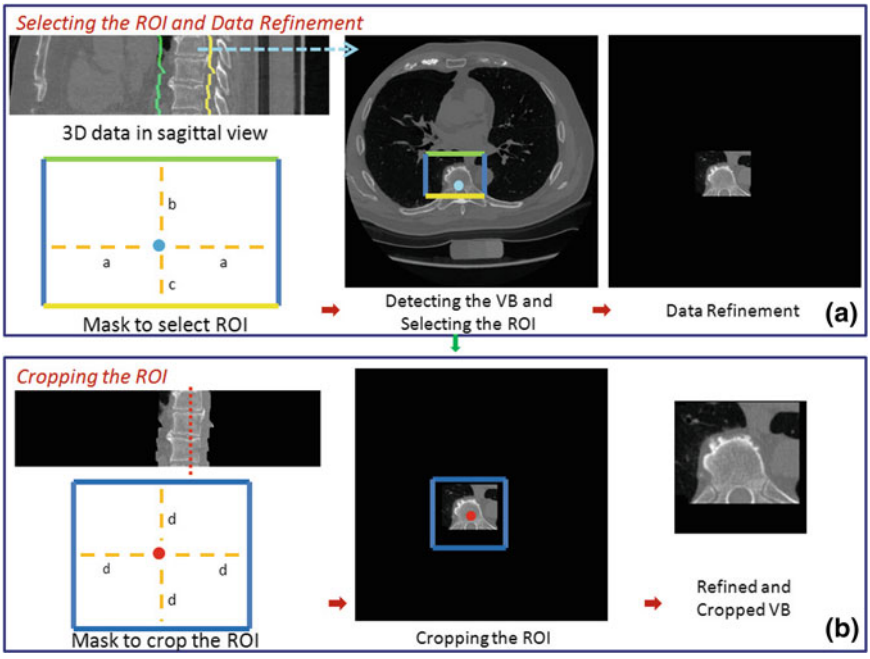


Fig. 4 In the first step, the MF is run on each slice of the 3D data. The output of this process is the detected points of each CT slices as shown with the *blue dot*. After the center points are detected, a mask is used to refine the data to specify the region of interest (*ROI*). In the mask, it is accepted that $\boldsymbol{a} = 50$, $\boldsymbol{b} = 60$, and $\boldsymbol{c} = 20$. Any user can change these values. But the user should be careful to extract the spinal processes and ribs roughly. Then, another mask can be used to crop the ROI using the average center points (the *red color*) of all slices of 3D data. $\boldsymbol{d} = 60$ is accepted to capture the VB region

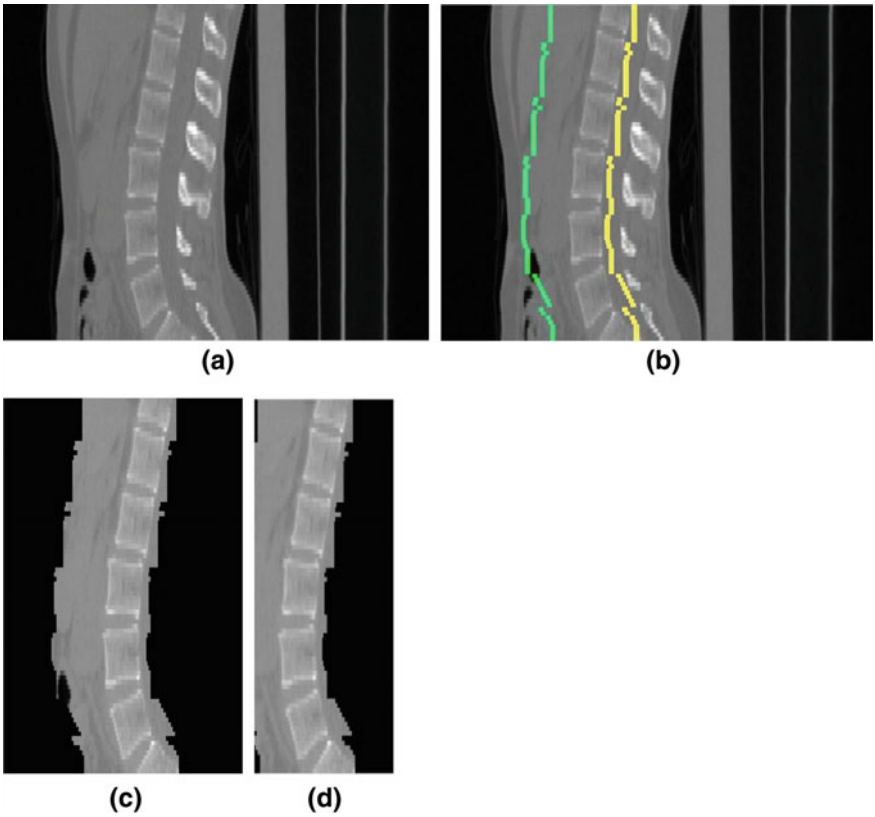


Fig. 5 The extraction of the spinal cord on a data set (Example-1). **a** Sagittal view of each data. **b** The detected VB region. **c** The refined data to extract the spinal processes and ribs. **d** The cropped data to reduce the size of the image

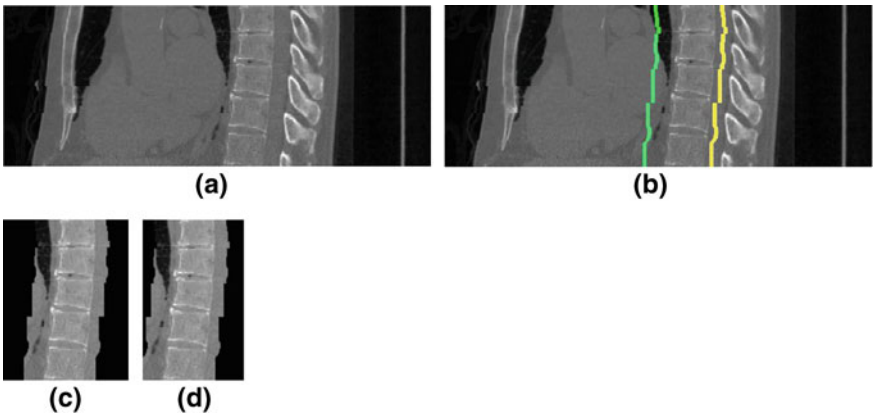


Fig. 6 The extraction of the spinal cord on a data set (Example-2). **a** Sagittal view of each data. **b** The detected VB region. **c** The refined data to extract the spinal processes and ribs. **d** The cropped data to reduce the size of the image

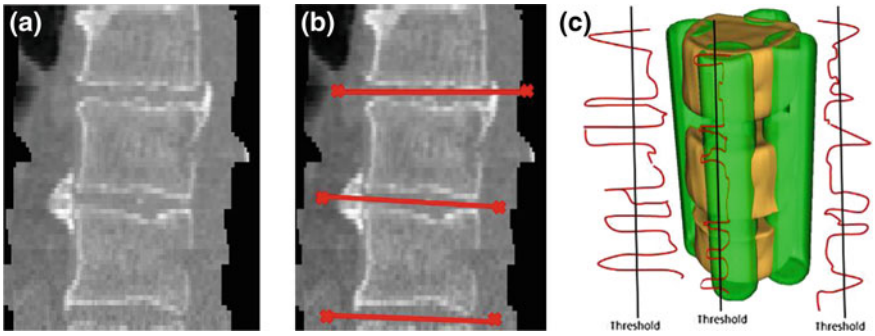


Fig. 7 The separation of each VB in a data set. Two choices are given to the user: The Manual and automatic options. Each option has its own advantages and disadvantages which are described the section. **a** An image which has 3 adjacent VBs. **b** The manual separation process with 6 points selected by a user. **c** The automatic separation method which was proposed by Aslan et al. in [16]

The advantage of automatic separation is to eliminate user interaction(s). However, there are two disadvantages: (i) increased error, (ii) current methods in the literature have higher execution time respect to the manual methods. To give the user his own choice, two methods are described.

2.3.1 Manual

After the spinal cord, processes, and ribs are extracted roughly, we need to separate adjacent VBs in order to embed the shape model to the image domain. In the manual separation, simple manual annotations are needed to specify the cut-points of VBs. For instance, if there are three VBs in the dataset, six points are annotated on the image. In the experiments, the average execution time to separate 12 adjacent VBs is 18 s. This timing may still not be optimum one, however, with manual annotations there should not be any possible data loss. In the next section, the automated separation process, which Aslan et al. previously published in [16], is described. It should be noted that segmentation accuracy is measured when the VBs are separated manually.

2.3.2 Automatic

In this section, a 3D framework to separate vertebral bones in CT images without any user intervention [16] is used. To separate the VBs, the previously developed approach based on 4 points automatically placed on cortical shell is used. An example of separation and segmentation of a VB is shown in Fig. 8c. After the spinal cord is extracted; the approximate centerline of VB column is obtained. These seeds are placed using the relatively higher gray level intensity values of the cortex region.

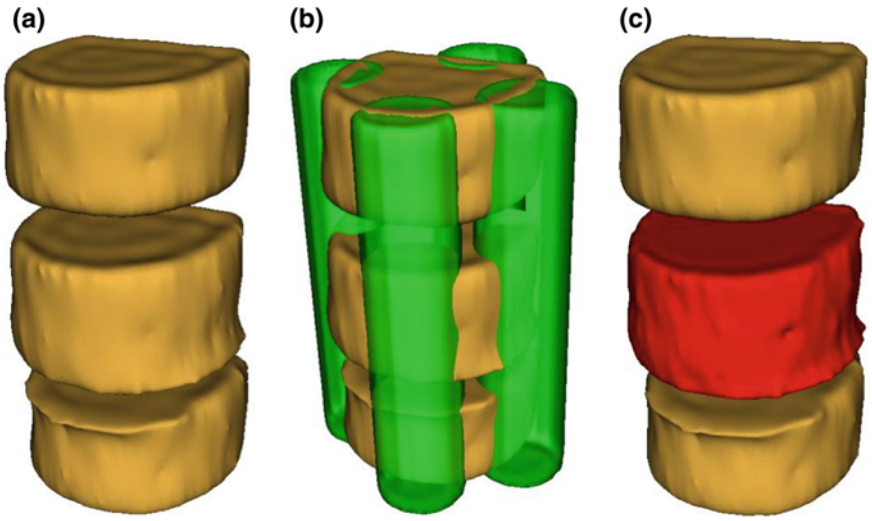


Fig. 8 The separation of the VB region. **a** 3D view of three adjacent VB. **b** Automated placement of 4 seeds on cortical bone and disc. **c** Separation of VB shown with *red color*

Next, the histogram for a neighborhood around each seed is obtained. The histogram represents the number of voxels whose intensity values are above 200 Hounsfield Unit (HU). This value is obtained empirically. Vertical boundaries of a VB show higher gray level intensity than inner region of the VB and disks. Figure 9 shows histograms (the red line), and thresholds (the black line). To search vertical limits of the VB, the following adaptive threshold equation is used as follows:

$$TH = \mu(A) + \kappa * [\max(A) - \mu(A)], \tag{6}$$

where $\kappa = 0.3$ which is derived from experiments by trial-and-error, where A represents each histogram vector with the red line as shown in Fig. 9, $\max(A)$ and $\mu(A)$ are the maximum and average values in the histogram vector.

In the separation step, 30 patients which totals to 117 VBs are used. The results can be classified as in [17]. There are five respective categories as described below.

- Excellent: The VB is successfully separated without any misclassification. Vertical limits are correctly obtained.
- Good: The VB is separated with small parts of adjacent disk or VB. Around 90 % or vertical limits are correctly obtained.
- Bad: The VB is separated, however noticeable parts of it are missed. Around 70–90 % of vertical limits are correctly obtained.
- Poor: Large portions of anatomical structure of VB are missed. Around 50–70 % of vertical limits are correctly obtained.
- Fail: The VB is not separated due to challenges.

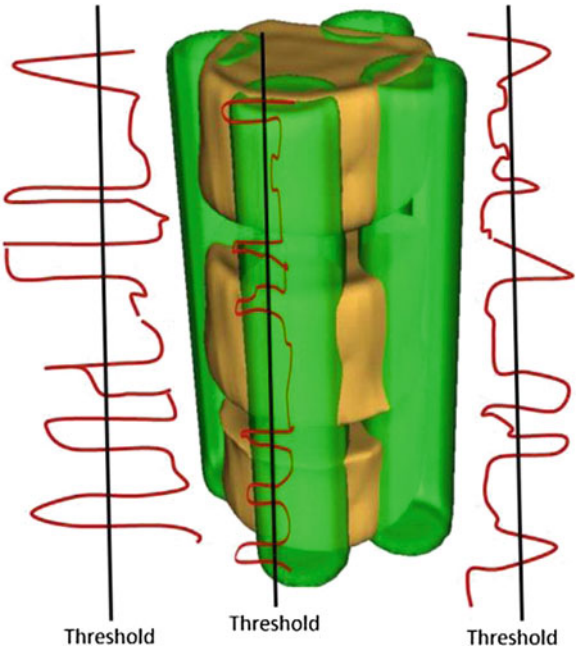


Fig. 9 The VB separation: The *green* volume shows the 4 four seeds. The *red lines* correspond to the number of voxels whose *gray* level are bigger than 200 HU. The *black lines* correspond to the threshold written in Eq. (6)

The proposed method produced about 85 % successful separation results, if excellent and good grades are considered. Hence, 15 % separation results were considered as fair, bad, or fail.

2.4 Segmentation Using Sign Distance-based Shape Model, Gaussian-based Intensity Model, and Symmetric Gibbs Potential-based Spatial Interaction Model

In the following subsections, we describe three proposed methods developed for the VB segmentation problem. We, first, describe each method, then show its experimental results.

The overall segmentation framework is shown in Fig. 10. The proposed framework steps are described in Algorithm 1 as follows:

Algorithm 1:

A(Training)): 80 training VB shapes are used to obtain the new probabilistic shape model. In this step, manually segmented VB shape which are obtained from 20 different patients and different regions (such as cervical, thoracic, and lumbar spine bone sections) are used.

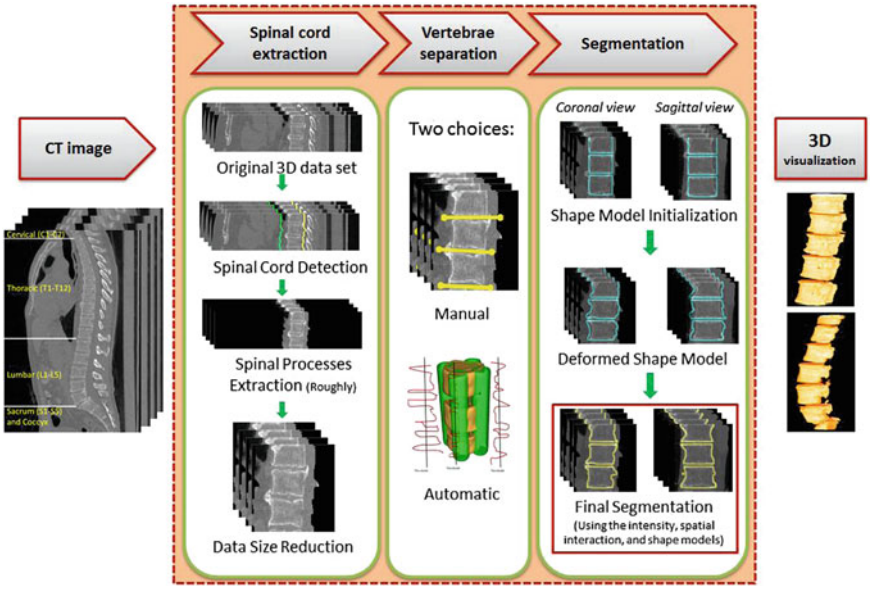


Fig. 10 The general segmentation framework. (Prior to this framework, it is required to obtain the shape model). In the first phase, the spinal cord is extracted, processes, and ribs roughly using the Matched filter. Also, the data size is reduced to minimize the execution time. In the second phase, the VBs are separated with two choices: (i) manual, (ii) automatic. In the third phase, a new shape based ICM method is proposed to segment the VBs

B) Spinal cord extraction (Pre-processing): The Matched filter is used to detect the spinal cord. This step roughly extracts the ribs and spinal processes. Also, the data size is reduced to 120x120xZ from 512x512xZ where Z is the number of slices. This step reduces the execution time of the segmentation process. The output of this phase is used in the following steps.

C) Separation of VBs (Pre-processing): Two choices are given for the user(s)-i) manual selection of disk to obtain each VB in a datasets, ii) fully automatic VB separation using the histogram based information. It should be noted all steps of the framework are fully automated except this step.

D) Segmentation: Three models are used to segment VBs: The intensity, spatial interaction, and shape models. The following '**While**' loop is processed for the segmentation.

While $j < N_{slices}$ **do** (N_{slices}: the number of slices.)

1) The initial segmentation using the ICM method which integrates the intensity and spatial interaction information. Using the EM algorithm, $p(\mathbf{I}|\mathbf{f}=0)$ and $p(\mathbf{I}|\mathbf{f}=1)$ are estimated; and $p(\mathbf{f}=0)$ and $p(\mathbf{f}=1)$ are estimated using the MGRF modeling. Then ICM method is used to select the optimum labeling which maximizes $\log p(\mathbf{I}|\mathbf{f}) + \log p(\mathbf{f})$.
2) The shape model is registered to the initially segmented region.
3) Final segmentation is carried out using the ICM which maximizes $\log p(\mathbf{I}|\mathbf{f}) + \log p(\mathbf{f}) + \log p(\mathbf{d}|\mathbf{f}, \mathbf{T})$.

End While

To obtain a good intensity model, the conditional probability distribution, $p(\mathbf{I}|\mathbf{f})$, of the original image is estimated. The intensity information is modelled using the Gaussian distribution. The Gaussian function can be written as

$$p(\mathbf{I}|\mathbf{f}=i) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(\mathbf{I}-\mu_i)^2}{2\sigma_i^2}\right) \tag{7}$$

The parameters of distributions (μ_i, σ_i) are estimated using the expectation-maximization (EM) method in [18, 19] where i = "0" and i = "1" represent 'background' and 'object' classes, respectively.

Spatial interaction helps correcting errors and recovering missing information in the image labeling problem [20]. Stochastic process on a random field is used to realize the image [21]. In this study, the unconditional probability distribution of the desired map (labeling), $p(\mathbf{f})$, is obtained. To estimate $p(\mathbf{f})$, the Gibbs distribution is used. The Gibbs distribution takes the following form

$$p(\mathbf{f}) = \frac{1}{Z} \exp\left(-\frac{U(\mathbf{f})}{T}\right) \tag{8}$$

where

$$Z = \sum_{f \in \mathcal{F}} \exp\left(-\frac{U(\mathbf{f})}{T}\right) \tag{9}$$

is a normalizing constant called the partition function, T is a control parameter called the temperature which is assumed to be 1 unless otherwise stated, and $U(\mathbf{f})$ is the Gibbs energy function. The energy is a sum of clique functions $V_c(\mathbf{f})$ over all possible cliques $\mathcal{C}$ as

$$U(\mathbf{f}) = \sum_{c \in \mathcal{C}} V_c(\mathbf{f}). \tag{10}$$

A clique is a set of sites in which all pairs of sites are neighbors. The clique potentials can be defined by

$$V_c(f) = \begin{cases} \gamma_c & \textit{if all sites on c have the same label} \\ -\gamma_c & \textit{otherwise}, \end{cases} \tag{11}$$

where γ_c is the potential for type-c cliques. In this proposed method, the Potts model [22] which is similar to Derin-Elliot model [23] is used. This model uses the potentials of the Potts model describing the spatial pairwise interaction between two neighboring pixels. The MGRF with the second order (8-pixel) neighborhood depends only on the whether the nearest pairs of pixel labels are equal or not. In this method, γ_c is estimated using the method proposed by Ali et al. in [24].

Human anatomical structures such as spine bones, kidneys, livers, hearts, and eyes may have similar shapes. These shapes usually do not differ greatly from one individual to another. There are many works which analyze the shape variability. Cootes et al. [25] proposed effective approach using principle component analysis (PCA). Abdelmumin [26] proposed another shape based segmentation method using the Level sets algorithm. Tsai et al. [27] proposed a shape model which is obtained using a signed distance function of the training data. Eigenmodes of implicit shape representations are used to model the shape variability. Their method does not require point correspondences. Their shape model is obtained using a coefficient of each training shape. Cremers et al. [28] proposed a simultaneous kernel shape based segmentation algorithm with a dissimilarity measure and statistical shape priors. This method is validated using various image sets in which objects are tracked successfully. Most published works are theoretically valuable. However, parameter optimization of the shape priors may take high execution time if the training set is large. Also, the optimization methods used in shape registration, such as the gradient descent, takes high execution time.

For the shape definition, mathematician and statistician D.G. Kendall writes: "All the geometrical information that remains when location, scale, and rotational effects are filtered out from an object." Hence, the shape information is modeled after the sample shapes are transformed into the reference space. Finally, the shape variability is modeled using the occurrences of the transformed shapes. In the proposed work, the vertebral body shape variability is analyzed using a probabilistic model.

In the next sections, each step is described in detail.

2.4.1 Shape Model Construction (Training)

Registration is the important method for shape-based segmentation, shape recognition, tracking, feature extraction, image measurements, and image display. Shape registration can be defined as the process of aligning two images of a scene. Image registration requires transformations, which are mappings of points from the source (reference) image to the target (sensed) image. The registration problem is formulated such that a transformation that moves a point from a given source image to

another target image according to some dissimilarity measure, needs to be estimated. The dissimilarity measure can be defined according to either the curve or to the entire region enclosed by the curve. The source and target images and transformation can be defined as follows:

- Source (I_s): Image which is kept unchanged and is used as a reference. This image can be written as a function $I_s{:}R^2 \rightarrow R$ for $\forall \mathbf{x} \in \Omega_s$.
- Target (I_t): Image which is geometrically transformed to the source image. This image can be written as a function $I_t{:}R^2 \rightarrow R$ for $\forall \mathbf{y} \in \Omega_t$.
- Transformation ($\mathbf{T}$): The function is used to warp the target image to take the geometry of the reference image. The transformation can be written as a function $\mathbf{T}{:}R^2 \rightarrow R^2$ which is applied to a point $\mathbf{x}$ in I_s to produce a transformed point which is calculated as $\mathbf{X} = \mathbf{T}(\mathbf{x})$. The registration error is calculated as $\mathbf{T}(\mathbf{x}) - \mathbf{y}$ for each transformed pixel.

Steps in the registration can be categorized in 5 different ways such as:

(i) Preprocessing: Image smoothing, deblurring, edge sharpening, edge detection, and etc.
(ii) Feature selection: Points, lines, regions and etc. from an the source and target image.
(iii) Feature correspondence: The correspondence between two images.
(iv) The transformation functions: Affine, rigid, projective, curved and etc.
(v) Resampling: Transformed image should be resampled in the new image domain.

In general, there are three categories of the registration methods: rigid, affine, and elastic transformation. In literature the rigid and affine transformations are classified as global transformations and elastic transformations are as local transformation. A transformation is global if it is applied to the entire image. A transformation is local if it is a composition of two or more transformations determined on different domains (sub-images) of the image.

- A rigid body transformation is the most fundamental transformation and is useful especially when correcting misalignment in the scanner. This transformation allows only translation and rotations, and preserves all lengths and angles in an image.
- An affine transformation allows translation, rotation, and scaling. Some authors defined the affine transformation as the rigid transformation plus scaling. Affine transformations involving shearing (projection) are called projective transformation. An affine transformations will map lines and planes into lines and planes but does not preserve length and angles.
- An elastic transformation allows local translation, rotation, and scaling, and it has more number of parameters than affine transformations. It can map straight lines into curves. An elastic registration is also called as a non-linear or curved transformation. This transformation allows different regions to be transformed independently.

2.4.2 The Registration of Training Shapes

For the training stage, 80 VB cross-sections (34 thoracic, 34 lumbar, 12 cervical) are used. These VBs are selected form 10 healthy and 10 with low bone mass patients. The more information about the testing CT data sets will be given in the experimental section.

In this subsection, we overview the shape representation approach used in this work. The training set consists of VB shapes, $\{C_1, \ldots, C_N\}$, as shown in Fig. 11; with the signed distance functions $\{\phi_1, \ldots, \phi_N\}$. Any pixel in this shape representation is shown as **x**. The registration of all training shapes is done using the similar approach described in [28] and used in [29, 30] as follows:

1. First, the average of the position factor (μ) and scale factor (σ) are obtained using the following equations

$$\mu = \begin{bmatrix} \mu_x & \mu_y \end{bmatrix}^T = \left[\frac{\sum_{i=1}^{N} \sum_{\Omega} xH(-\phi_i(\mathbf{x}))}{\sum_{i=1}^{N} \sum_{\Omega} H(-\phi_i(\mathbf{x}))} \quad \frac{\sum_{i=1}^{N} \sum_{\Omega} yH(-\phi_i(\mathbf{x}))}{\sum_{i=1}^{N} \sum_{\Omega} H(-\phi_i(\mathbf{x}))} \right]^T, \tag{12}$$

$$\sigma = \begin{bmatrix} \sigma_x^2 & \sigma_y^2 \end{bmatrix}^T = \left[\frac{\sum_{i=1}^{N} \sum_{\Omega} (x-\mu_x)^2 H(-\phi_i(\mathbf{x}))}{\sum_{i=1}^{N} \sum_{\Omega} H(-\phi_i(\mathbf{x}))} \quad \frac{\sum_{i=1}^{N} \sum_{\Omega} (y-\mu_y)^2 H(-\phi_i(\mathbf{x}))}{\sum_{i=1}^{N} \sum_{\Omega} H(-\phi_i(\mathbf{x}))} \right]^T, \tag{13}$$

   where $H(.)$ is the Heaviside step function.
2. A global transformation is used to register training shapes with scale and translation parameters. The transformation has scaling, **S**, and translation

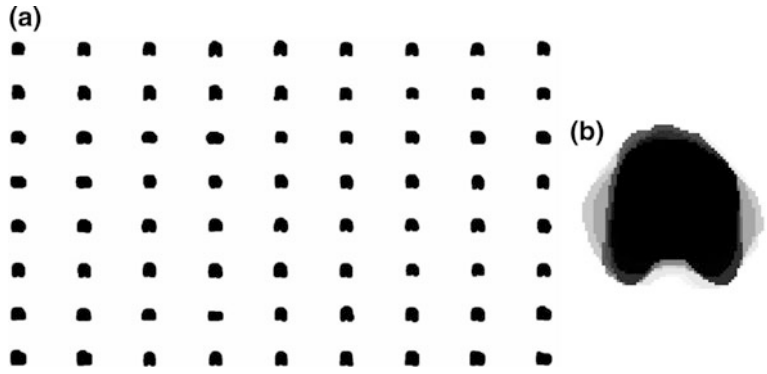


Fig. 11 **a** The training VB images. In this experiment, 80 VB shapes which are obtained from 20 different patients and different regions (such as cervical, thoracic, and lumbar spine bone sections) are proposed. **b** The average shape of all training VB images. The *darker color* represents the higher object probability

components **Tr**. Obtain the transformation parameters (t_x, t_y, s_x, s_y) for each training shape, ϕ, as

$$\mathbf{Tr} = \begin{bmatrix} t_x & t_y \end{bmatrix}^T = \left[\mu_x - \frac{\sum_\Omega xH(-\phi(\mathbf{x}))}{\sum_\Omega H(-\phi(\mathbf{x}))} \quad \mu_y - \frac{\sum_\Omega yH(-\phi(\mathbf{x}))}{\sum_\Omega H(-\phi(\mathbf{x}))} \right]^T, \tag{14}$$

$$\mathbf{S} = \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix} = \begin{bmatrix} \frac{\sigma_x}{\sqrt{\frac{\sum_\Omega (x-\mu_x)^2 H(-\phi(\mathbf{x}))}{\sum_\Omega (H-\phi(\mathbf{x}))}}} & 0 \\ 0 & \frac{\sigma_y}{\sqrt{\frac{\sum_\Omega (y-\mu_y)^2 H(-\phi(\mathbf{x}))}{\sum_\Omega H(-\phi(\mathbf{x}))}}} \end{bmatrix}^T \tag{15}$$

3. The transformation will be in the form $\mathbf{T}(\mathbf{x}) = \mathbf{X} = \mathbf{Sx} + \mathbf{Tr}$, where $\mathbf{X}$ is the transformed point of $\mathbf{x}$.

Note: In 2D case, the rotation parameter for the VB shape registration is not necessary since VB shape does not show important variation in different rotation.

2.4.3 Training stage (Obtaining Probabilistic Shape Model)

1. Segment training images manually.
2. Align segmented images.
3. Generate shape variation. Intersection of training shape is accepted as an object volume. The rest of the volume is accepted as variability volume except the background region.
4. Obtain the probabilities of the object and background in the variability volume of the shape model.

A new probabilistic shape model is formed using the training shapes as shown in Fig. 11a. All registered training shapes are combined as shown in Fig. 11b. The shape prior represented as $\mathcal{R} = \mathcal{O} \cup \mathcal{B} \cup \mathcal{V}$ is generated. The proposed shape model functions are defined as follows:

$$\mathcal{O} = \bigcap_{i=1}^{N} \mathrm{H}(-\phi_i), \tag{16}$$

$$\mathcal{B} = \bigcap_{i=1}^{N} \mathrm{H}(\phi_i), \tag{17}$$

$$\mathcal{V} = \bigcup_{i=1}^{N} H(-\phi_i) - \bigcap_{i=1}^{N} H(-\phi_i), \tag{18}$$

where ϕ_i^t represents any training shape.

Figure 12 shows the detailed description of the shape models. The green color shows the background region ($\mathcal{B}$) which does not have any intersection with any training shape. The blue color shows the object region ($\mathcal{O}$) which is the intersection of all training shapes. In (a), the gray color represents the variability region ($\mathcal{V}$) that can be described as the union of all projected training shapes subtracted by the intersection of those shapes. In this variability region, the object and background probabilistic shapes are modeled. The red color, in (b), shows the outer contour of the variability region, and it is represented as ($\mathbf{J}$). In the registration step, the shape model is embedded to the initially segmented region. $\mathbf{J}$ is used to estimate the registration parameters. The object ($p(\mathbf{d}|\mathbf{f}=1)$) and background ($p(\mathbf{d}|\mathbf{f}=0)$) probabilistic models are defined in the variability region. The probabilistic shape models are defined as follows:

- If $\mathbf{x} \in \mathcal{O}$, then $p(d_\mathbf{x}|f_\mathbf{x}=1)=1$ and $p(d_\mathbf{x}|f_\mathbf{x}=0)=0$
- if $\mathbf{x} \in \mathcal{B}$, then $p(d_\mathbf{x}|f_\mathbf{x}=1)=0$ and $p(d_\mathbf{x}|f_\mathbf{x}=0)=1$
- if $\mathbf{x} \in \mathcal{V}$, then

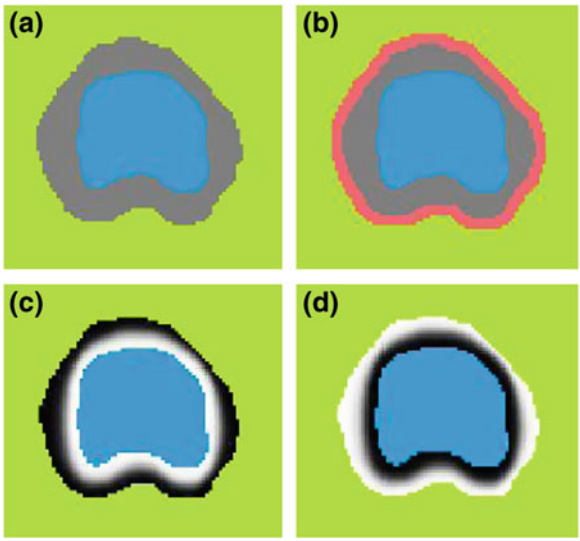


Fig. 12 The shape model. The *green color* shows the background region which does not have any intersection with any training shape. The *blue color* shows the object region which is the intersection of all training shapes. **a** The *gray color* represents the variability region that can be described as the union of all projected training shapes subtracted by the intersection of those shapes. In this variability region, the object and background probabilistic shapes are defined. **b** The *red color* shows the outer contour of the variability region. **c** The object ($\boldsymbol{p}(\mathbf{d}|\mathbf{f}=1)$). **d** background ($\boldsymbol{p}(\mathbf{d}|\mathbf{f}=0)$) shapes are modelled in the variability region in which the pixel values are defined in (0:1)

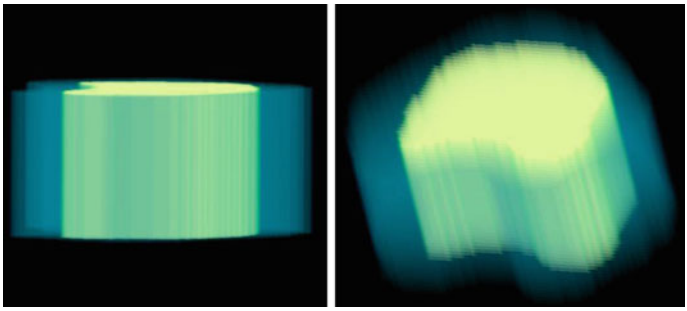


Fig. 13 The shape model is shown in 3D (when propagating 2D shape model into 3D). The outer volume represents the variability region, the inner volume represents the object region

$$p(d_{\mathbf{x}}|f_{\mathbf{x}}=1)=\frac{\sum_{i=1}^{N}H(-\phi_i(\mathbf{x}))}{N}, \tag{19}$$

$$p(d_{\mathbf{x}}|f_{\mathbf{x}}=0)=\frac{\sum_{i=1}^{N}H(\phi_i(\mathbf{x}))}{N}. \tag{20}$$

3D representation of the shape model is shown in Fig. 13. It should be noted that Eqs. (19) and (20) represents the probability value at each pixel, $\mathbf{x}$.

2.4.4 Initial Segmentation

To estimate the initial labeling $\mathbf{f}^*$, the ICM method which integrates the intensity and spatial interaction information is used. It should be noted that the shape model has not been used in this process. The initial segmented region is used to obtain the SDF representation which is required in the registration process. An example of the initial labeling is shown in Fig. 14. The method has acceptable results, because a relatively large amount of pedicles and ribs are separated from the vertebral body. It should be noted that there may still some portion of pedicles and ribs which could not be separated. Between Fig. 14d, e, there is a shape registration process which is shown in Figs. 15 and 16.

2.4.5 Embedding the Shape Model

To use the shape prior in the segmentation process, $\mathbf{f}^*$ and the shape prior are required to be registered. The shape model has a variability region as shown in Fig. 12a. The outer contour is represented as $\mathbf{J}$. In the registration process, $\mathbf{J}$ and $\mathbf{f}^*$ will be the source and target information, respectively. The registration step is done in 2D slice by slice since the shape model can deform locally independently from

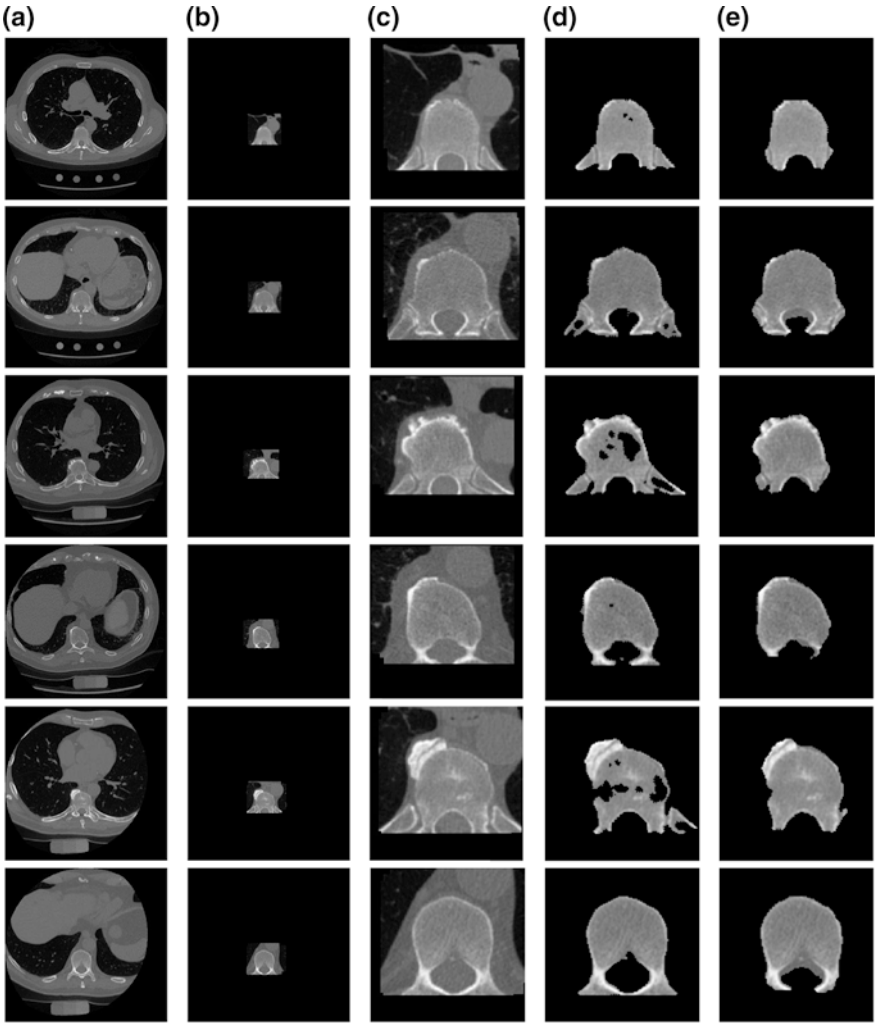


Fig. 14 An example of the initial labeling. **a** Original CT images. **b** Detection of the VB region and refinement. **c** The cropped VB region. **d** The initial labeling, $\mathbf{f}^*$ using only the intensity and spatial interaction models. **e** The final segmentation using three models

other slices. This approach gives deformation flexibility between each slices which stocks in z-axis. The transformation has 4 parameters such as s_x, s_y (for scale, **S**), and t_x, t_y (for translation, **Tr**). It should be noted that the rotation is not necessary in the method since the registration is done slice by slice; and the VB does not show important rotational variation in the axial axes. Let us define the transformation result by β that is obtained by applying a transformation **T** to a given contour/surface

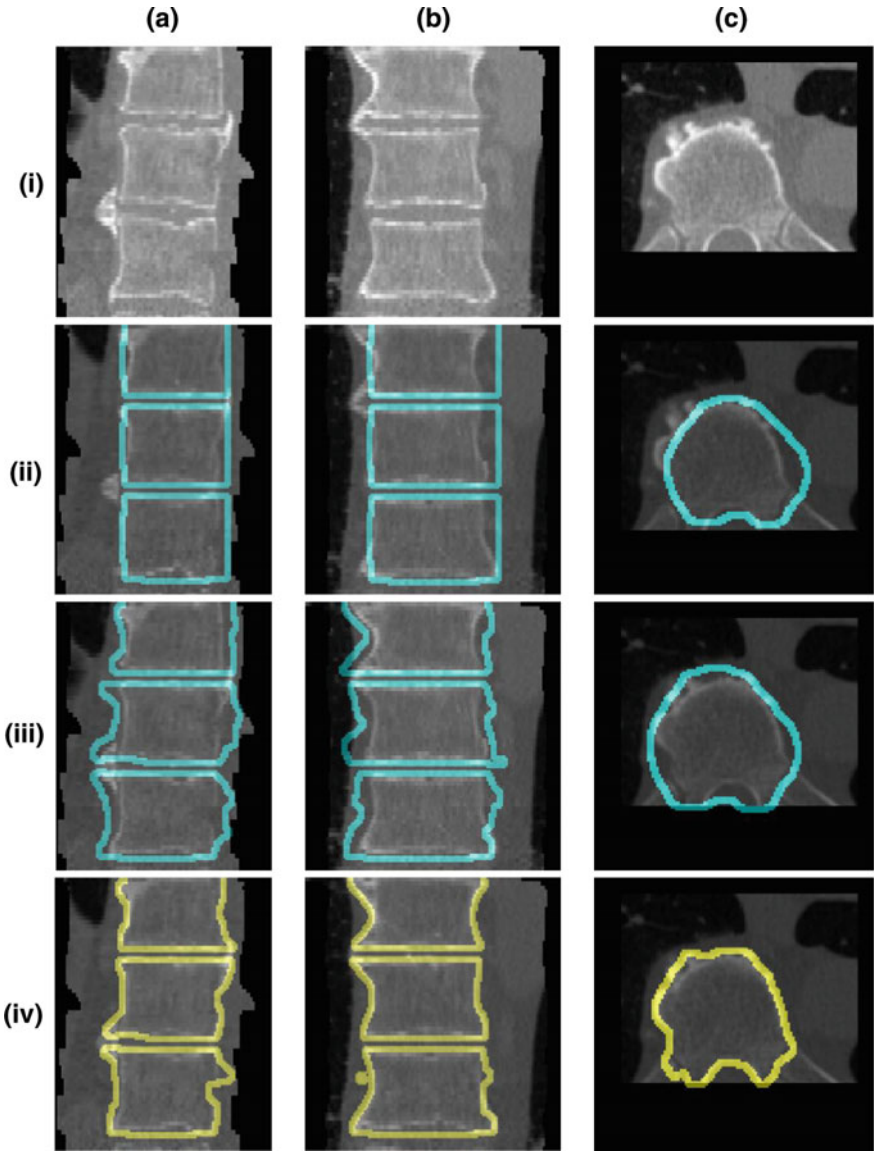


Fig. 15 Embedding the shape model to the image domain and the final segmentation. (*i*) A CT data after the extraction of spinal cord. (*ii*) Shape model initialization (the *blue color* show the outer surface of the variability region **J**). The contour **J** is placed equally in every slice using the obtained ROI. (*iii*) Embedding the shape model to the image domain. (*iv*) Final segmentation using three models: The intensity, spatial interaction, and shape information. Images and results are shown in the (*a*) sagittal, (*b*) coronal, and (*c*) axial views

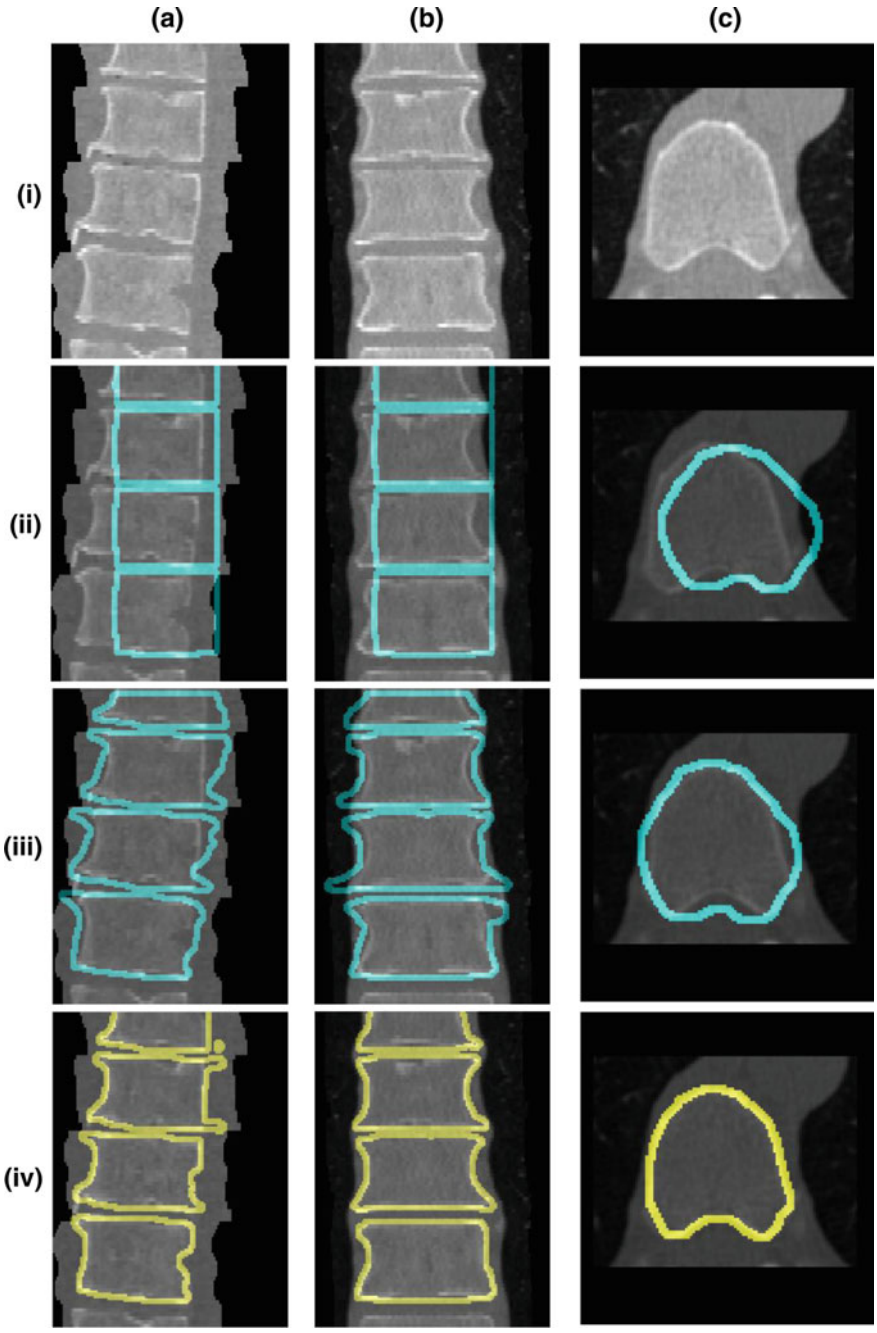


Fig. 16 Embedding the shape model to the image domain and the final segmentation. (*i*) A CT data after the extraction of spinal cord. (*ii*) Shape model initialization (the *blue color* show the outer surface of the variability region **J**). The contour **J** is placed equally in every slice using the obtained ROI. (*iii*) Embedding the shape model to the image domain. (*iv*) Final segmentation using three models: The intensity, spatial interaction, and shape information. Images and results are shown in the (*a*) sagittal, (*b*) coronal, and (*c*) axial views

α. In this case, β and α correspond to $\mathbf{f}^*$ and $\mathbf{J}$, respectively. The transformation can be written for any point $\mathbf{X}$ in the space as $\mathbf{T}(\mathbf{x}) = \mathbf{X} = \mathbf{S}\mathbf{x} + \mathbf{T}\mathbf{r}$. Now consider $\mathbf{x} \in \phi_{\mathbf{J}}$ and $\mathbf{X} \in \phi_{\mathbf{f}^*}$.

The registration of the shape model and testing image is done as follows:

(i) First, the average of the position factor ($\mu^{\mathbf{f}^*}$) and scale factor ($\sigma^{\mathbf{f}^*}$) are obtained using the following equations

$$\mu^{\mathbf{f}^*} = \begin{bmatrix} \mu_x^{\mathbf{f}^*} & \mu_y^{\mathbf{f}^*} \end{bmatrix} = \begin{bmatrix} \frac{\sum_\Omega xH(-\phi_{\mathbf{f}^*}(\mathbf{x}))}{\sum_\Omega H(-\phi_{\mathbf{f}^*}(\mathbf{x}))} & \frac{\sum_\Omega yH(-\phi_{\mathbf{f}^*}(\mathbf{x}))}{\sum_\Omega H(-\phi_{\mathbf{f}^*}(\mathbf{x}))} \end{bmatrix}^T, \tag{21}$$

$$\sigma^{\mathbf{f}^*} = \begin{bmatrix} (\sigma_x^{\mathbf{f}^*})^2 & (\sigma_y^{\mathbf{f}^*})^2 \end{bmatrix} = \begin{bmatrix} \frac{\sum_\Omega (x-\mu_x^{\mathbf{f}^*})^2 H(-\phi_{\mathbf{f}^*}(\mathbf{x}))}{\sum_\Omega H(-\phi_{\mathbf{f}^*}(\mathbf{x}))} & \frac{\sum_\Omega (y-\mu_y^{\mathbf{f}^*})^2 H(-\phi_{\mathbf{f}^*}(\mathbf{x}))}{\sum_\Omega H(-\phi_{\mathbf{f}^*}(\mathbf{x}))} \end{bmatrix}^T. \tag{22}$$

(ii) Obtain the transformation parameters (t_x, t_y, s_x, s_y) for the shape model, $\phi_{\mathbf{J}}$, as

$$\mathbf{Tr} = \begin{bmatrix} t_x & t_y \end{bmatrix}^T = \begin{bmatrix} \mu_x^{\mathbf{f}^*} - \frac{\sum_\Omega xH(-\phi_{\mathbf{J}}(\mathbf{x}))}{\sum_\Omega H(-\phi_{\mathbf{J}}(\mathbf{x}))} & \mu_y^{\mathbf{f}^*} - \frac{\sum_\Omega yH(-\phi_{\mathbf{J}}(\mathbf{x}))}{\sum_\Omega H(-\phi_{\mathbf{J}}(\mathbf{x}))} \end{bmatrix}^T, \tag{23}$$

$$\mathbf{S} = \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix} = \begin{bmatrix} \frac{\sigma_x^{\mathbf{f}^*}}{\sqrt{\frac{\sum_\Omega (x-\mu_x^{\mathbf{f}^*})^2 H(-\phi_{\mathbf{J}}(\mathbf{x}))}{\sum_\Omega H(-\phi_{\mathbf{J}}(\mathbf{x}))}}} & 0 \\ 0 & s_y = \frac{\sigma_x^{\mathbf{f}^*}}{\sqrt{\frac{\sum_\Omega (y-\mu_x^{\mathbf{f}^*})^2 H(-\phi_{\mathbf{J}}(\mathbf{x}))}{\sum_\Omega H(-\phi_{\mathbf{J}}(\mathbf{x}))}}} \end{bmatrix}^T \tag{24}$$

(iii) Transform each point $\mathbf{x} \in \Omega$ to the new point $\mathbf{X}$. Hence, the shape model is registered to the image domain.

(iv) The new probabilistic function at each pixel is $p(d_{\mathbf{X}}|f_{\mathbf{X}}) = p(d_{\mathbf{x}}|f_{\mathbf{x}}, \mathbf{T})$. Hence, the new transformed pixels will have the same probabilistic value with corresponding pixels. An example of the registration and final segmentation results are shown in Figs. 15 and 16.

2.4.6 Final Energy Minimization Using Three Models: Intensity, Spatial Interaction, and Shape

Three probabilistic models are used. Before this step, the followings are obtained already (i) the initial labeling $\mathbf{f}^*$ that maximizes $p(\mathbf{I}|\mathbf{f}^*)$, (ii) the MGRF model for $p(\mathbf{f}^*)$, and (iii) the transformed shape prior to maximize $p(\mathbf{d}|\mathbf{f}^*, \mathbf{T})$. It should be noted that the transformation step is not an iterative process, and there is a unique

solution for a given initial segmentation and shape model. Now, the objective is to optimize the following equation to maximize the likelihood energy functional. Algorithm 2 shows the proposed segmentation process using a new ICM method.

Algorithm 2: Optimization of Three Models

1. While $i < N_{iter}$ do
2. For all $\mathbf{X} \in \Omega$ do
3. Update $f_{\mathbf{X}}^{*}$ by the value of $f_{\mathbf{X}}$ which maximizes

$$\log p(I_{\mathbf{X}}|f_{\mathbf{X}}) + \log p(f_{\mathbf{X}}^{*}) + \log p(d_{\mathbf{X}}|f_{\mathbf{X}}^{*})$$

4. End for
5. Increase i
6. End while

Note: It should be noted that $\mathbf{X} = \mathbf{Sx} + \mathbf{Tr}$ is any transformed pixel,and Ω is the pixel domain in the image.

$$L(\mathbf{I}, \mathbf{d}, \mathbf{f}, \mathbf{T}) = \log p(\mathbf{I}|\mathbf{f}) + \log p(\mathbf{f}) + \log p(\mathbf{d}|\mathbf{f}, \mathbf{T}). \tag{25}$$

2.4.7 Experimental Results-CT Data Population

The training and testing images were acquired from GE LightSpeed VCT, Toshiba Aquilion, and Imatron C-150 CT scanners with an in-plane resolution range of 0.63–0.98 mm and a slice thickness of 0.63–3.00 mm. For the testing stage, 18 patient data sets, of which 10 are from female ('F') and 8 are from male ('M'), and a phantom are examined in this study. There are 16–96 axial slices with 512×512 voxels. The proposed algorithm is tested on 932 CT slices/66 VBs which are obtained from different spine bone regions (i.e. lumbar, thoracic, and etc.). In the datasets, the number of visible VBs changes from 2 to 8. The data sets are also categorized as 'healthy' (H) and 'with low bone mass' (L) with respect to their calcium absorbtion. The experiments are run on 7 'H' and 11 'L' data sets. The ages of the test subjects varies between 38 and 76 years with an average age of 61.3 years with 12.2 years standard deviation.

2.4.8 Experimental Results-Segmentation Measurements

Figure 17 shows the region of true positive (TP), true negative (TN), false positive (FP), and false negative (FN). In this figure, the reference region represents the ground truth which is verified by a radiologist. The test region represents the automated segmented region. For the ESP, the segmentation quality is measured

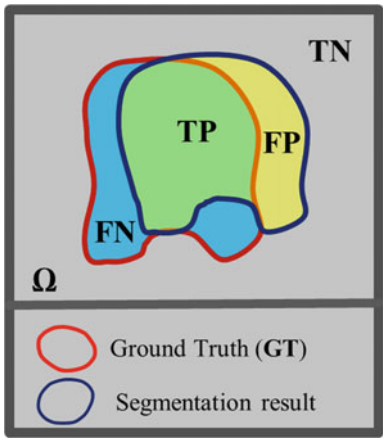


Fig. 17 In the segmentation quality measurements, there are 4 regions to be considered as: True positive (*TP*), false positive (*FP*), true negative (*TN*), and false negative (*FN*). The reference and test regions represent the ground truth and automatic segmented regions

using the Jaccard distance whereas for the clinical data sets, the segmentation quality is measured using 4 difference formulations. The measurements can be defined as follows:

$$Accuracy\ (\%) = 100 * \frac{TP + TN}{TP + FP + FN + TN} \tag{26}$$

$$Precision\ (\%) = 100 * \frac{TP}{TP + FP} \tag{27}$$

$$Jarrardcoefficient\ (\%) = 100 * \frac{TP}{TP + FN + FP} \tag{28}$$

$$Dice'scoefficient\ (\%) = 100 * \frac{2\mathrm{TP}}{2\mathrm{TP} + \mathrm{FP} + \mathrm{FN}} \tag{29}$$

2.4.9 Experimental Results-Validation Using the Phantom

In the experiments, the ESP, which is an accepted standard for quality control [3] in bone densitometry, is used to validate the segmentation algorithms. Because clinical CT images have gray level inhomogeneity, noise, and weak edges in some slices, the ESP was scanned with the same problems to validate the robustness of any method. CT images may have various noise and image uncertainties. Image noise is related to the numbers of X-ray photons absorbed by each small area of the image [31]. The higher exposure levels result in a better image, and less image noise, but more radiation is absorbed by a patient. Hence, segmentation methods should be robust to various image conditions. It is assumed that CT images may have random noise. To assess the proposed method under various challenges, Gaussian noise with a zero mean and different variance σ_n^2 values (from 0 to 0.5) is added to the CT

images. The segmentation accuracy is measured for each method using the ground truths. The proposed method is compared with other 3 alternatives which can be represented as follows: $A1$: Active appearance method described in [32], $A2$: Level sets method described in [33], and $A3$: Shape based level sets method described in [29].

The segmentation results and the average accuracy on the ESP (when the initialization is optimal) are shown in Figs. 18 and 19, respectively. In this test, the initial point is chosen at the center of the object of interest. The elapsed time[1] to segment 12 ESP images is 136.2 s for $A1$, 194 s for $A2$ (with 30-pixel radius seed size), 248 s for $A3$ (with 30-pixel radius seed size) and 12.8 s for the proposed method (without the detection and VB separation parts). It should be noted that the all results are obtained until each method reaches its possible convergence stage.

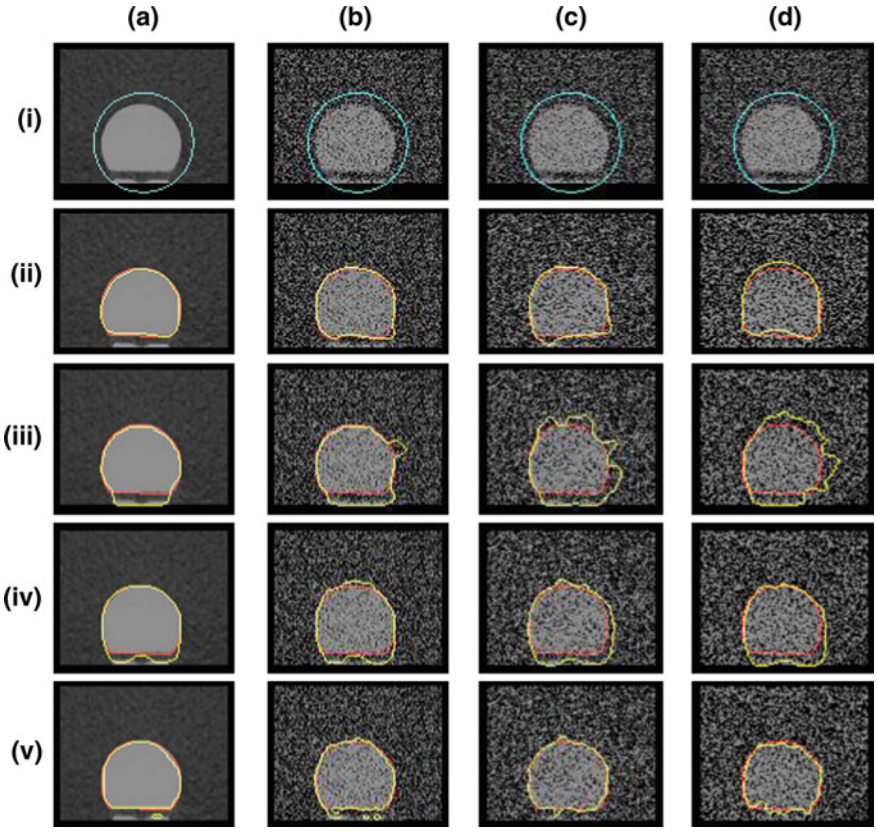


Fig. 18 Good Initialization: Segmentation comparison under (*a*) no noise, noise variances (*b*) $\boldsymbol{\sigma_n^2} = 0.1$, (*c*) $\boldsymbol{\sigma_n^2} = 0.25$, and (*d*) $\boldsymbol{\sigma_n^2} = 0.5$. (*i*) Initialization. The results of (*ii*) $\boldsymbol{A}1$, (*iii*) $\boldsymbol{A}2$, (*iv*) $\boldsymbol{A}3$ and (*v*) the proposed method. (The *red* and *yellow colors* show the contour of the ground truths and segmented regions, respectively.)

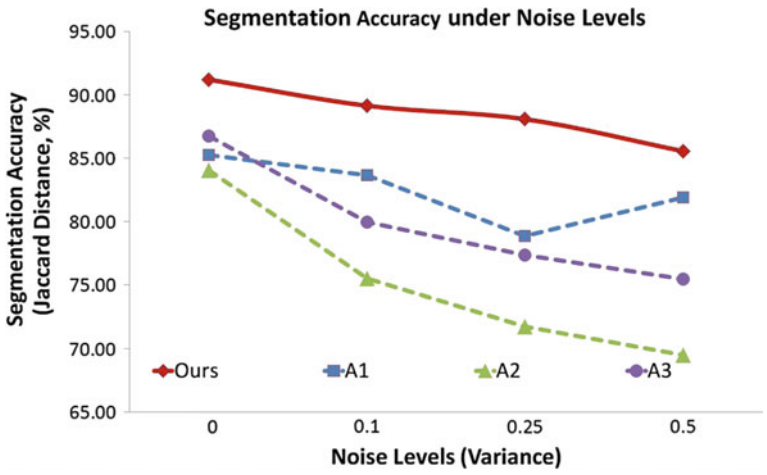


Fig. 19 Average segmentation accuracy of the VB segmentation on 12 CT images (ESP) with respect to the various noise levels

The results show that the proposed method is robust under various noise levels as well as faster than other famous alternatives.

The initialization effect is also validated in this experiments. It should be noted that *A*1–*A*3 need perfect manual initializations. However, the method is almost independent of the initialization (which is usually required in the registration step). The segmentation results and the accuracy on the ESP (when the initialization is not optimal) are shown in Figs. 20 and 21, respectively. In this figures, the initial point is chosen not close to the center point of the object of interest. It's clear that the proposed method performance is almost constant with different initial points. On the contrary, the alternative methods are severely suffering from performance degradation.

The effect of each model is validated as shown in Fig. 22. In the figure, (i) shows the initialization place for each method. The results which are based on (ii) only the intensity model, (iii) intensity and spatial interaction, (iv) intensity, spatial interaction, and shape models are shown. The intensity based approach is not robust under various noise levels. After the spatial interaction model is used, the segmentation is getting better and most of the noise is eliminated. However, there are still missing information and some noise using the two models. With the proposed approach much better results are obtained compared with other models. The segmentation accuracy with respect to the various noise levels is shown in Fig. 23.

2.4.10 Experimental Results-Results on Clinical CT Images

In this study, different type of data sets are used. Classification of data sets are categorized into 3 groups as shown in Table 1. Classification is based on 6 features.

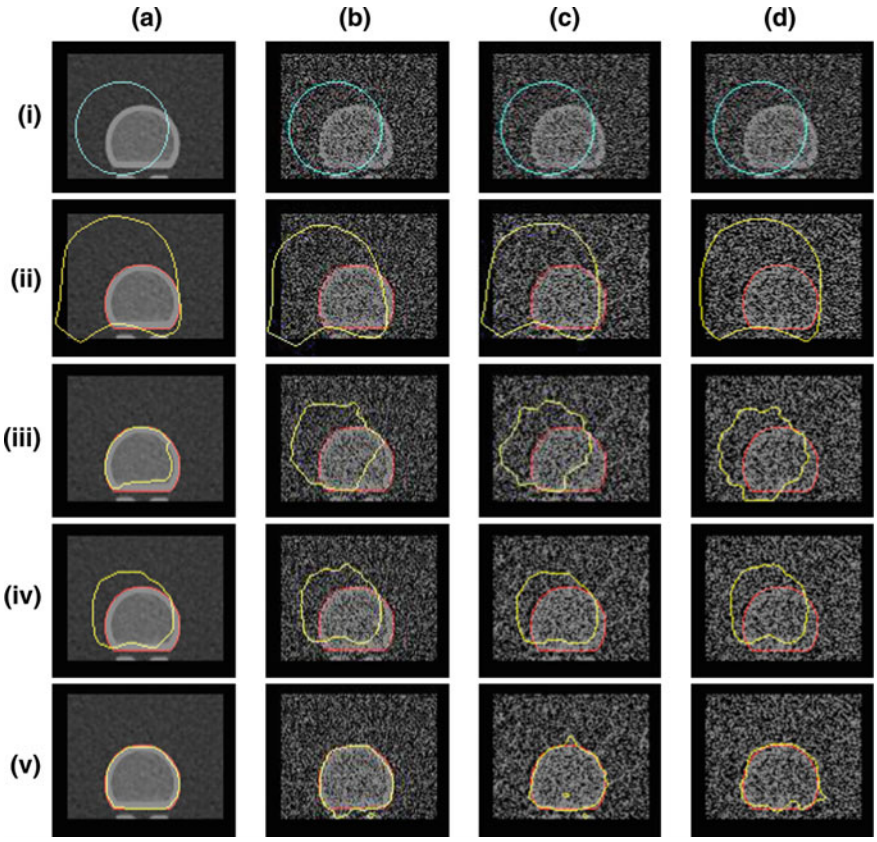


Fig. 20 Worse Initialization: Segmentation comparison under (*a*) no noise, noise variances (*b*) $\sigma_n^2 = 0.1$, (*c*) $\sigma_n^2 = 0.25$, and (*d*) $\sigma_n^2 = 0.5$. (*i*) Initialization. The results of (*ii*) $\boldsymbol{A}1$, (*iii*) $\boldsymbol{A}2$, (*iv*) $\boldsymbol{A}3$ and (*v*) the proposed method. (The *red* and *yellow* colors show the contour of the ground truths and segmented regions, respectively.)

Slice thickness, resolution, spine column region (shape), fractures, diseases, and spine bone edges are main factors of the classification. Class A is the best data sets which can be segmented and analyzed easily. Data sets which are classified in class C have serious problems such as diseases, fractures, weak spine edges, and low resolution. Data sets in class B have some problems but they are better than data sets in class C. Categorization could help to analyze the results separately.

As mentioned above, the proposed algorithm is tested on 932 CT slices/66 VBs which are obtained from 18 (7 H and 11 L) different patients and different spine bone regions (i.e. lumbar, thoracic, and etc.). The segmentation accuracy is given with respect to the health condition of bone ('H', 'L', and 'H + L'), and the classification criteria (Class A, B, and C). Table 2 shows the quality measurement results of the proposed segmentation method. The four different measurements are given to be judged fairly. As can be interpreted from the results in the table, the

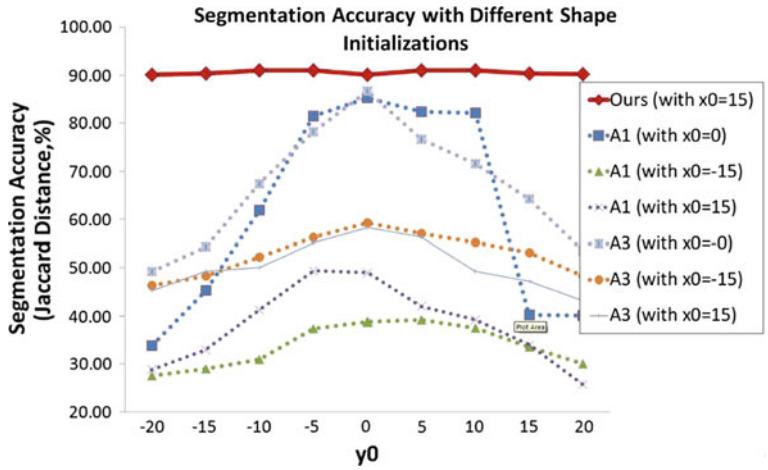


Fig. 21 The effect of the initialization on the segmentation accuracy of 12 CT images (ESP) using $A1$, $A3$, and ours. x_0 and y_0 represent the initial point in the X-direction and Y-direction respectively w.r.t. the center of the object

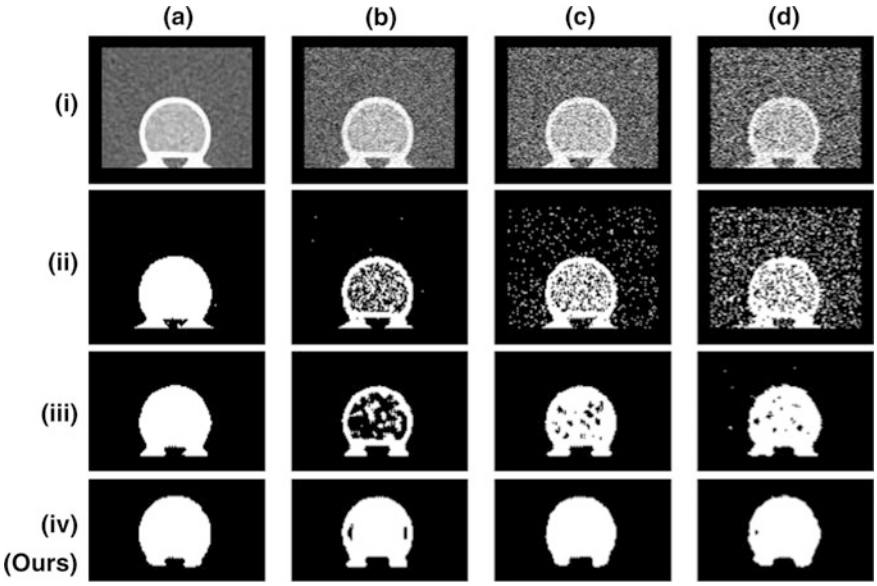


Fig. 22 Segmentation results of an ESP CT slice with (*a*) no noise, noise variances (*b*) $\sigma_n^2 = 0.1$, (*c*) $\sigma_n^2 = 0.25$, and (*d*) $\sigma_n^2 = 0.5$. (*i*) The original *gray* level image with various noise levels. The results of (*ii*) only the intensity based segmentation, (*iii*) the initial segmentation $\mathbf{f}^*$ based on the intensity and spatial interaction models, (*iv*) proposed method integrating three models (intensity, interaction, and shape)

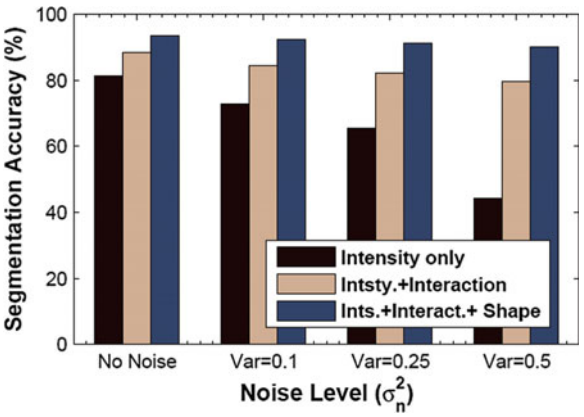


Fig. 23 Average segmentation accuracy of the VB segmentation method on 12 ESP CT images. After each VB is detected, the size is reduced to 120 × 120 that the segmentation measurements are calculated

Table 1 Classification of clinical data sets used in the experiments: There are totally 18 data sets in the data base

	Class A	Class B	Class C
Slice thickness	Usually <2.5 mm	Usually ≥2.5 mm, but may be <2.5 mm	Usually ≥3.00 mm, but can be smaller if disease exists
Resolution	High	Usually low	Lower
Shape	Straight	Straight/Curvy	Usually curvy but it can be straight
Bone degeneration or osteophyte	No	May have disease	May have disease
Fracture	No	No serious fracture	May have serious fractures
Edge	Strong	Strong/Weak	Usually weak
Note	Optimum data	This class has some problems	This class has very serious problems

Class A, B, and C have 7, 5, and 6 data sets, respectively

Table 2 Segmentation results of each data class based on different measures

	'H'	'L'	'H + L'	Class A	Class B	Class C
Accuracy, %	98.2	97.9	97.6	99.0	98.7	98.0
Precision, %	91.1	86.6	88.6	90.9	89.9	84.4
Jaccard coefficient, %	87.6	83.0	85.0	87.7	86.9	80.3
Dice's coefficient, %	93.1	90.4	91.5	93.8	92.9	89.0

The measurements are based on data sets with 120 × 120 × ***Z*** size where ***Z*** represents number of slices

Jaccard coefficient gives the lowest quality score respect to the others. Also, the accuracy gives the highest quality score respect to the other measurements. By using this information, the proposed segmentation reaches the scores of 'Jaccard

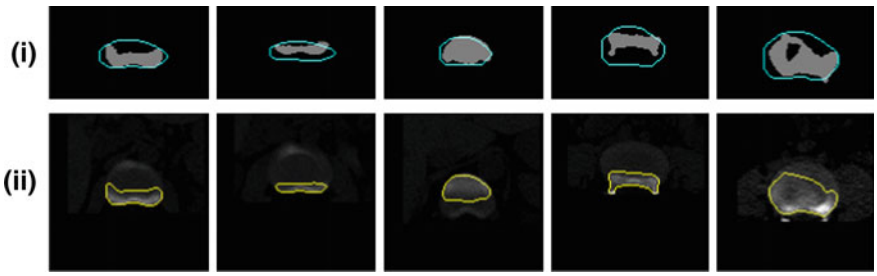


Fig. 24 The shape registration process and segmentation results of end-plate slices of VBs. (*i*) The shape model is registered to the initial segmented region. The *blue color* shows the contour of the registered variability region, **J**. (*ii*) Final segmentation results. The *yellow* color shows the contour of the segmented region

coefficient' 87.6, 83.0, and 85.0 % for 'H', 'L', and 'H + L', respectively. The same measurements gives 87.7, 86.9, and 80.3 % for classes A, B, and C, respectively. The proposed method reaches the scores of 'accuracy' measurement 98.2, 97.9, and 97.6 % for 'H', 'L', and 'H + L', respectively. The same measurements gives 99.0, 98.7, and 98.0 % for classes A, B, and C, respectively.

Figure 24 shows the shape model registration and final segmentation result on end-plate slices of VBs. The proposed method is able to segment end-plate slices thanks to the shape embedding process although the shape model is obtained using the full view of VB slices. However, further improvements can be searched for more accuracy. The unnecessary regions such as ribs and processes are extracted as much as possible using the shape model. Figure 25 shows some of the segmentation examples in axial view.

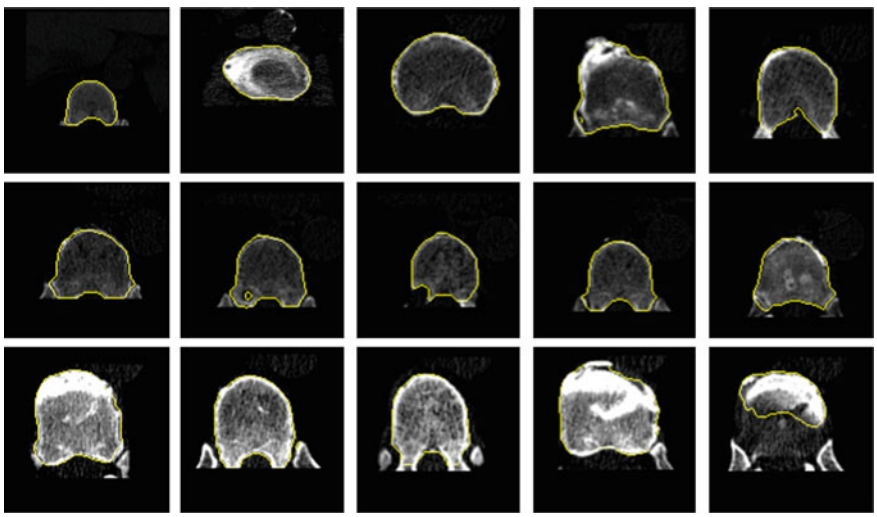


Fig. 25 Some of segmentation results are shown in the axial view. These images are from thoracic and lumbar regions. The *yellow color* shows the contour of the segmented region

Table 3 Relative Comparison

	$F1$	$F2$	$F3$	$F4$	$F5$	$F6$
Klinder et al. [1]	No	>36.5	No	Yes	All	No
Mestmayer et al. [4]	Yes	>36	Yes	No	Specific	Yes
Proposed	Optional	<3	Yes	No	All	No

Table 4 Average execution time of the framework: The average time calculation is based on 12 VBs/96 CT slices

Framework stages	Execution time (seconds)
Spinal cord extraction	15.7
VB separation	18 (manual)/45 (automatic)
Initial segmentation	54.1
Shape registration	32.6
Final segmentation	46.8
Total	167.2 (*when manual VB separation is considered*)

The proposed framework is compared with two of very important spinal bone related publications using features of each method. The features can be described as follows: $F1$: User interactions, $F2$: Execution time (minutes) to run all steps respect to segment 12VBs, $F3$: Extraction of spinal processes, $F4$: Vertebra identification, $F5$: Suitability to all or specific location of spinal column (such as thorocic, lumbar, and etc.), $F6$: The BMD measurements. Since the direct comparison with these two methods are very difficult, each feature is compared as shown in Table 3. Although the results are obtained using difference computer system for each method, the most important contribution of this work is to segment VBs in very low execution times with the acceptable segmentation accuracy. We maintain that the proposed method can be applied in real time clinical studies.

It should be noted that VBs were manually separated in this test. The framework take 167.2 s (less than 3 min) to segment 12 VBs—see Table 4. The number of slices affects the execution time. For the 2D/3D framework, the execution time is related to the number of slices in the image. Some experimental images of 3D results are shown on coronal and sagittal views in Figs. 26, 27, 28 and 29.

2.5 Segmentation Using Euclidian Distance-based Shape Model, LCG- based Intensity Model, and Asymmetric Gibbs Potential-based Spatial Interaction Model

In this method, image modeling is a unified approach, which is created by integrating several of our previous and ongoing efforts in image modelling techniques. The first step is modeling shape variations using our new distance probabilistic

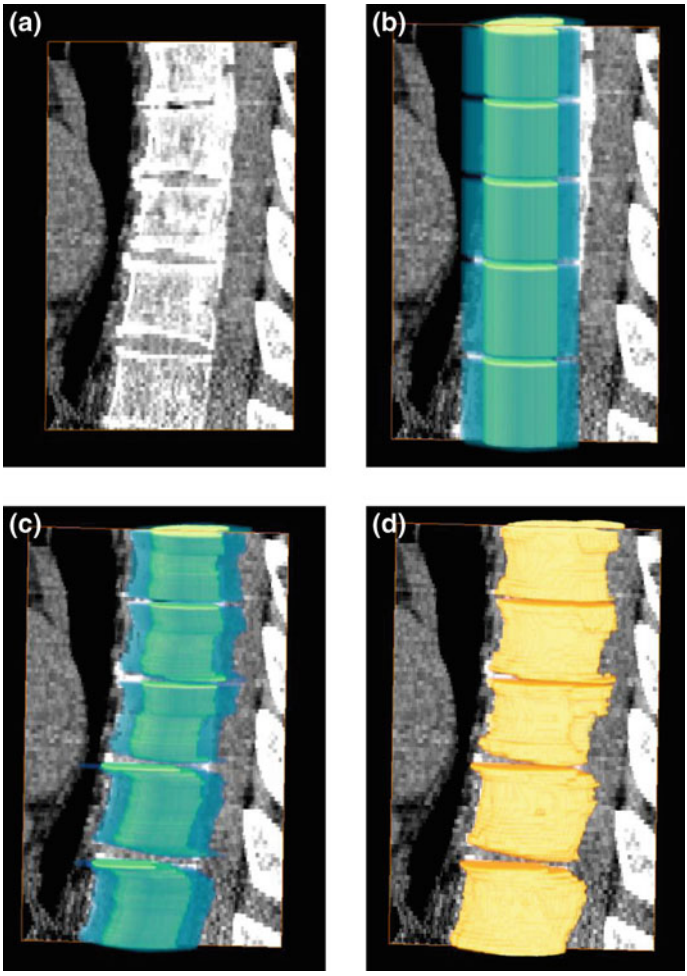


Fig. 26 The shape embedding and final segmentation results are shown in 3D views. **a** A CT data is shown in the sagittal axis (without the refinement). **b** The initial location of the shape models. 2D shape models are propagated in z-axis to form 3D models. The *blue color* (outer volume) shows the variability region, whereas the *yellow color* (inner volume) represents the object region. **c** The shape model after registration. **d** The final segmentation results using the three models

model [34]. Where the distance marginal densities of the VB and its background inside the variability region are approximated using a Poisson distribution, which is refined by positive and negative Gaussian components. In order to use this distance probabilistic model with any given VB set of images, we align this given volume with the training 3D shape. The second step is approximating VB's gray level using our linear combination of Gaussian distributions (LCG) model with positive and negative components [35, 36]. Moreover to model the spatial relationships between

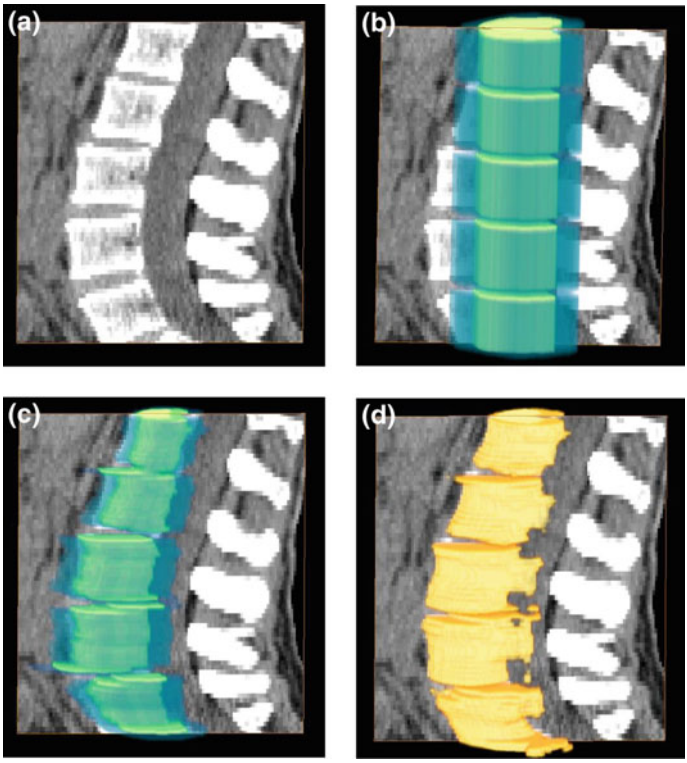


Fig. 27 The shape embedding and final segmentation results are shown in 3D views. **a** A CT data is shown in the sagittal axis (without the refinement). **b** The initial location of the shape models. 2D shape models are propagated in z-axis to form 3D models. The *blue color* (outer volume) shows the variability region, whereas the *yellow color* (inner volume) represents the object region. **c** The shape model after registration. **d** The final segmentation results using the three models

the region labels, Potts model is used. The spatial pairwise interactions between two neighboring voxels, which define the potentials of Potts model, are estimated using our new analytical approach [37]. The last step is integrating these region and boundary properties as well as the shape information using a new energy function, which is globally minimized using s/t graph cuts to get the optimal segmentation.

In this method, only the VB separation process is used as a pre-processing step.

2.5.1 Shape Modeling

We create a 3D shape of vertebral body a subset of VB data sets. This is done as follows: The VBs' volumes, where each VB consists of 8 CT slices, are manually segmented by a medical expert. The segmented VB slices are binary images, as shown in the Fig. 30. These segmented images are aligned to the ESP, which is

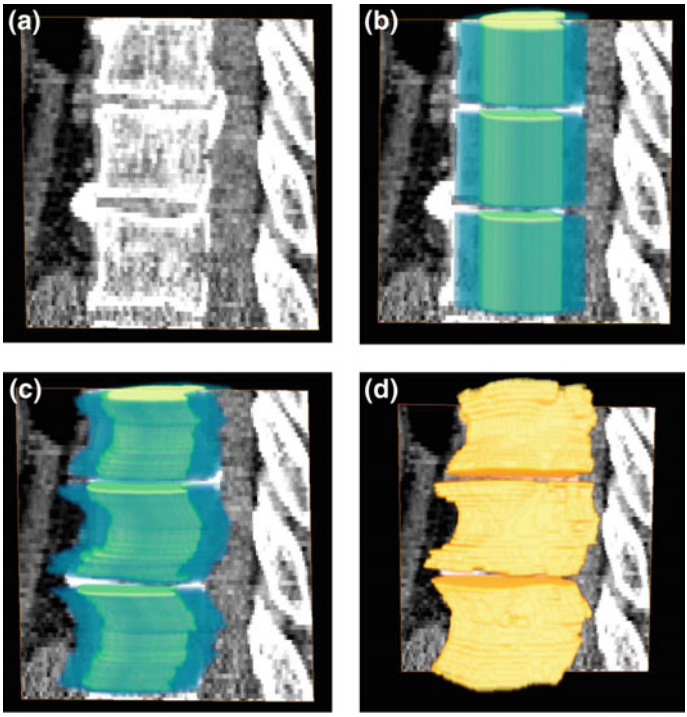


Fig. 28 The shape embedding and final segmentation results are shown in 3D views. **a** A CT data is shown in the sagittal axis (without the refinement). **b** The initial location of the shape models. 2D shape models are propagated in z-axis to form 3D models. The *blue color* (outer volume) shows the variability region, whereas the *yellow color* (inner volume) represents the object region. **c** The shape model after registration. **d** The final segmentation results using the three models

used as a reference to align any given VB later to use the proposed shape model. This alignment approach is similar to the method presented in the previous method. Finally, a "shape volume" $\mathcal{P}_s = \mathcal{O} \cup \mathcal{B} \cup \mathcal{V}$ is generated, which its slices are shown in Fig. 30. Three regions in this shape model: white color represents $\mathcal{O}$ (VB), black represents $\mathcal{B}$ (its background), and gray is the variability region $\mathcal{V}$. Figure 31a illustrates a 3D view of the VB and its variability region.

We use a distance probabilistic model, to model variability region $\mathcal{V}$ i.e., the 3D shape variations. A normal distance is used to describe the VB and its background in the variability region in the distance probabilistic model as follows.

$$d_p = \min_{c \in \mathbf{C}_{\mathcal{OV}}} \| p - c \|, \tag{30}$$

Equation (30) represents the distance from a voxel $p \in \mathcal{V}$ to the organ/variability surface $\mathbf{C}_{\mathcal{OV}}$. Figure 31b, c shows iso-surfaces for $\mathbf{C}_{\mathcal{OV}}$, where an iso-surface $\mathbf{C}_{d_p}$ is a set of voxels located at equal distance d_p from $\mathbf{C}_{\mathcal{OV}}$. Assuming each iso-surface

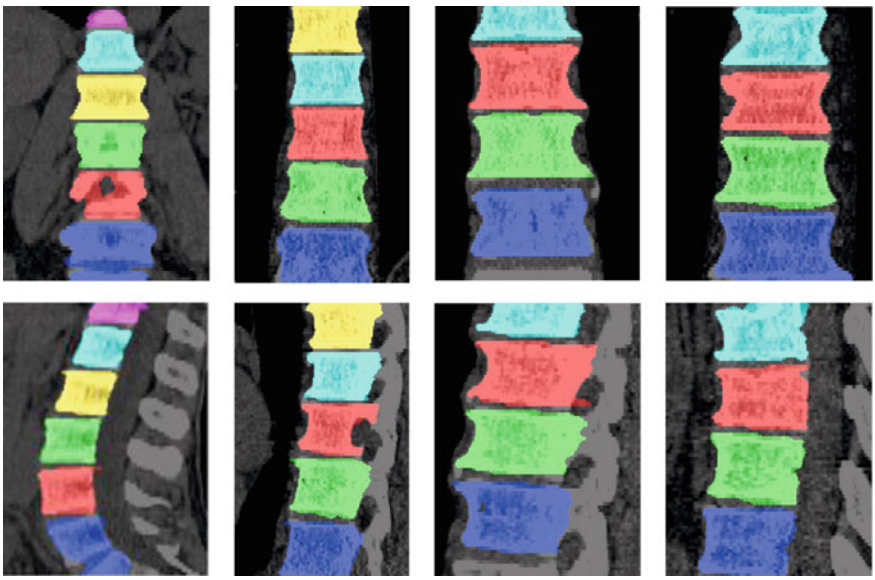


Fig. 29 Some segmentation results examples on coronal (the *first row*) and sagittal (the *second row*) views of 3D segmentation

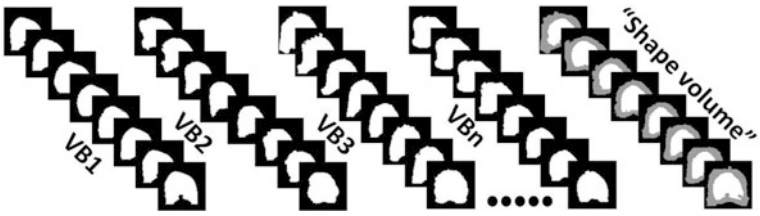


Fig. 30 Constructing the shape prior volume. $\{\boldsymbol{VB}1, \cdots, \boldsymbol{VBn}\}$ training CT slices of different data sets. (n represents the number of training data sets). *Last column* shows the shape prior slices with variability region

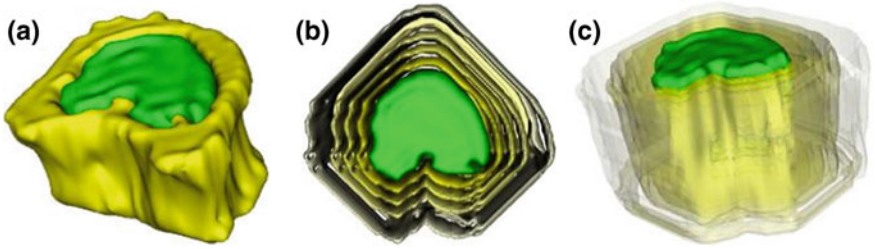


Fig. 31 **a** A 3D view of the 3D shape prior. $\mathcal{O}$ (*green color*), $\mathcal{V}$ (*yellow color*). **b**, **c** Different 3D views for the iso-surfaces $\mathbf{C}_{d_p}, \boldsymbol{p} \in \mathcal{V}$

$\mathbf{C}_{d_p}$ is a normally propagated wave from $\mathbf{C}_{\mathcal{OV}}$, the probability of an iso-surface to be VB decays exponentially as the discrete index d_p increases. Therefore; we can use a Poisson distribution to model the distance histogram, which is estimated as follows. The value of the histogram at a distance d_p can be calculated as

$$h_{dp} = \sum_{i=1}^{M} \sum_{j=1}^{K} \sum_{p \in \mathbf{C}_{d_p}} \delta(p \in \mathcal{O}_{ij}), \tag{31}$$

where $\delta(\mathrm{A})$ is an indicator function equals 1 when the condition $\boldsymbol{A}$ is true, and zero otherwise, M is the number of training data sets, K is the number of CT slices of each data set, and $\mathcal{O}_{ij}$ is the VB region in the training set i and in the slice j. The domain of the distance d_p is the variability region. The histogram should be multiplied by the VB prior value, which can be estimated as follows:

$$\pi_{\mathcal{O}} = \frac{1}{MK|\mathcal{V}|} \sum_{i=1}^{M} \sum_{j=1}^{K} \sum_{p \in \mathcal{V}} \delta(p \in \mathcal{O}_{ij}). \tag{32}$$

Same computations can be done to estimate the marginal density of VB's background

Distance Probabilistic Model

Assuming the conditional distribution $P(\mathbf{d}|\mathbf{f})$ is an independent random field of distances, then

$$P(\mathbf{d}|\mathbf{f}) = \prod_{p \in \mathcal{V}} (Pd_p|f_p). \tag{33}$$

We model the distance marginal density of each class $P(d_p|f_p)$ as a Poisson distribution refined by $K_{f_p}^{+}$ positive and $K_{f_p}^{-}$ negative discrete Gaussians components. So the distance marginal density of each class can be written as follows:

$$P(d_p|f_p) = \vartheta(d_p|\lambda_{f_p}) + \sum_{r=1}^{K_{fp}^{+}} w_{f_p,r}^{+} \varphi(d_p|\theta_{f_p,r}^{+}) - \sum_{l=1}^{K_{fp}^{-}} w_{f_p,l}^{-} \varphi(d_p|\theta_{f_p,l}^{-}), \tag{34}$$

where $\vartheta(d_p|\lambda_{f_p})$ is a Poisson density with rate λ, $\varphi(.|\theta)$ is a Gaussian density with parameter $\theta \equiv (\mu, \sigma^2)$ with mean μ and variance σ^2. $w_{f_p,r}^{+}$ means the rth positive weight in class f_p and $w_{f_p,l}^{-}$ means the lth negative weight in class f_p. This weights have a restriction $\sum_{r=1}^{K_{f_p}^{+}} w_{f_p,r}^{+} - \sum_{l=1}^{K_{f_p}^{-}} w_{f_p,l}^{-} = 1$.

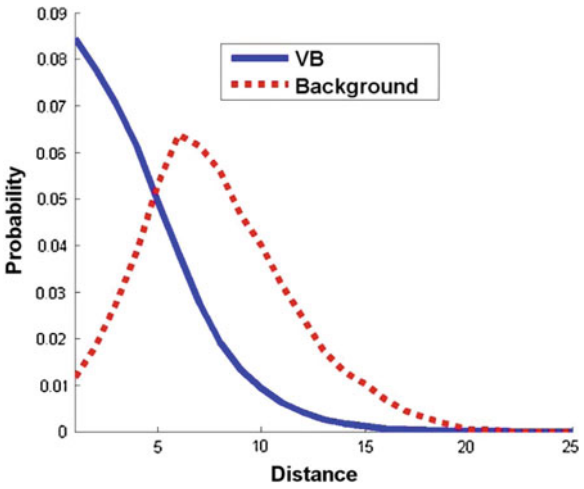


Fig. 32 The distance probabilistic models of VB and its background in the variability region

The maximum likelihood estimator is used to estimate the Poisson distribution parameter. Where the modified EM algorithm [38] is used to estimate the parameters of Gaussians components. Figure 32 illustrates the estimated densities of VB and its background.

2.5.2 The Gray Level Probabilistic Model

Also, assuming the conditional distribution of the original volume given the map is an independent random field of gray levels with different gray value distributions.

$$\mathrm{P}(\mathbf{I}|\mathbf{f}) = \prod_{p \in \mathcal{P}} P(I_p|f_p). \tag{35}$$

To approximate the gray level marginal density of each class $P(I_p|f_p)$, we use our LCG model [35, 36] with C_{p,f_p} positive and C_{n,f_p} negative components. Thus; the gray level marginal density of each class can be written as follows:

$$P(I_p|f_p) = \sum_{r=1}^{C_{\mathrm{p},f_p}} w_{\mathrm{p},r,f_p} \varphi(I_p|\theta_{\mathrm{p,r,f_p}}) - \sum_{s=1}^{C_{\mathrm{n},f_p}} w_{\mathrm{n},s,f_p} \varphi(I_p|\theta_{\mathrm{n},s,f_p}) \tag{36}$$

where, $\varphi(I_p|\theta)$ is a Gaussian density with parameter θ (mean μ and variance σ^2), w_{p,r,f_p} means the rth positive weight in class f_p, w_{n,s,f_p} means the sth negative weight in class f_p. Also the weights should satisfy $\sum_{r=1}^{C_{\mathrm{p},f_p}} w_{\mathrm{p},r,f_p} - \sum_{s=1}^{C_{\mathrm{n},f_p}} w_{\mathrm{n},s,f_p} = 1$.

Also, the modified EM algorithm [38], which deals with the positive and negative components, is used to estimate the parameters of the LCG model. Figure 33

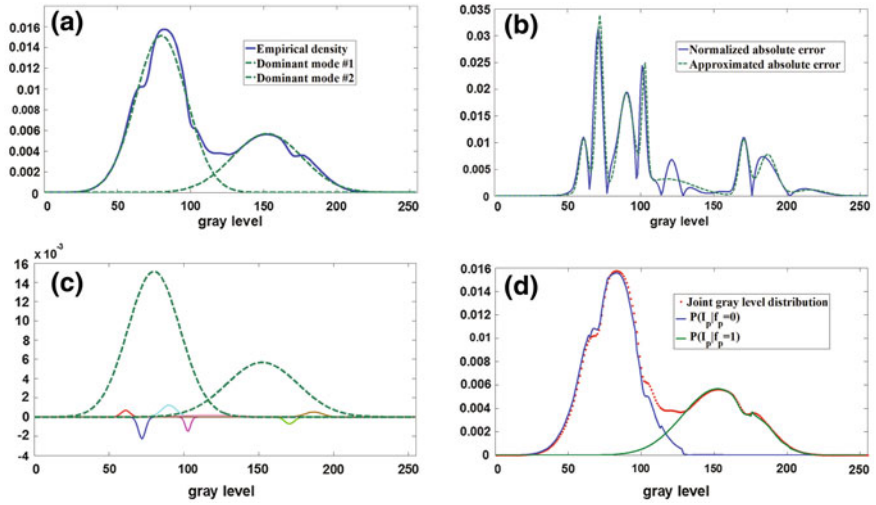


Fig. 33 Synthetic example for the gray level probabilistic model. **a** Empirical densities and the estimated dominant modes. **b** Normalize absolute error. **c** LCG components of the gray level probabilistic models. **d** Final estimated marginals densities

summarizes the estimation of the gray level probabilistic model (More details can be found in [38]).

Figure 33a shows the approximation of the given volume gray levels empirical distribution H with a mixture of two Gaussians P_2 using conventional EM algorithm. Figure 33b illustrates the absolute of the deviations between H and P_2, which is approximated by a mixture of Gaussians P_n using conventional EM algorithm. Figure 33c shows the approximation of the joint distribution P, which consists of P_2 and $+ve$ and $-ve$ components of P_n. The summation of a dominant mode and the closest $+ve$ and $-ve$ components is the marginal distribution of a class $P(I_p|f_p)$ as shown in Fig. 33d.

2.5.3 Spatial Interaction Model

Assuming the region map $\mathbf{f} = \{f_1, \ldots, f_{|\mathcal{P}|}\}$ is a realization of random variables, for which the joint distribution is presented as a Markov-Gibbs Random Field with respect to a neighborhood system $\mathcal{N}$. The Gibbs potential governing asymmetric pairwise co-occurrences of the region labels can be described as follows:

$$V(f_p, f_q) = \gamma \delta(f_p \neq f_q), \tag{37}$$

where γ is the potential value specifying the Gibbs potential. This potential value γ is estimated analytically using our approach [37], which is based on MLE of the MGRF:

$$\gamma^* = \left(2 - 4\tilde{\mathfrak{F}}_{\text{neq}}(\mathbf{f})\right), \tag{38}$$

where $\tilde{\mathfrak{F}}_{\text{neq}}(\mathbf{f})$ denotes the relative frequency of the non-equal labels in the voxel pairs of that family.

$$\tilde{\mathfrak{F}}_{\text{neq}}(\mathbf{f}) = \frac{1}{|\mathcal{T}|}\sum\nolimits_{\{p,q\}\in\mathcal{T}} \delta\left(f_p \neq f_q\right), \tag{39}$$

where $\mathcal{T}$ is the family of the neighboring voxel pairs supporting the Gibbs potentials.

Therefore, the region map unconditional probability distribution $P(\mathbf{f})$ can be specified by the following Gibbs probability distribution:

$$P(\mathbf{f}) = \frac{1}{Z}\exp(-\sum\nolimits_{\{p,q\}\in\mathcal{N}} V(f_p,f_q)). \tag{40}$$

2.5.4 Final Energy Minimization Using Three Models:

After estimating the shape model and the appearance models our goal is to integrate these models to find the best labelling $\mathbf{f}$ i.e., the optimal segmentation. The MAP estimate of $\mathbf{f}$, using Eqs. (33), (35) and (38), is equivalent to minimizing the following function:

$$E(\mathbf{f}) = \sum_{p\in\mathcal{P}} -\log(P(d_p|f_p)) + \sum_{p\in\mathcal{P}} -\log(P(I_p|f_p)) + \sum_{\{p,q\}\in\mathcal{N}} V(f_p,f_q). \tag{41}$$

The first term, in this function, measures disagreement with the shape information for assigning a label f_p to a voxel p. The second term measures the disagreement with with the voxel intensity I_p for assigning a label f_p to a voxel p. Finally, the penalty of the discontinuity between neighbors voxels p and q is represented by $V(f_p,f_q)$.

In order to minimize the function Eq. (41), we construct a 3D graph and define the weight of each edge as shown in Table 5. The minimum cost cut on this graph corresponds to the optimal segmentation boundary between the VB and its background. This minimum cost cut is computed exactly in polynomial time for two terminal graph cuts with positive edges weights via s/t Min-Cut/Max-Flow algorithm [39].

Although each given VB volume should be aligned with the ESP, to which our training images are registered to create the shape model, it is a challenge to directly align a VB from its CT slice (e.g., Figure 34a). Therefore, a preprocessing step [40] is used to detect the VB region (yellow box in Fig. 34a) in the CT slice. After that we can use the shape prior and apply our segmentation frame work on that region. Details are given in the following algorithm.

Table 5 Graph edges weights

Edge	Weight	For
$\{p,q\}$	$V(f_p,f_q)$	$\{p,q\} \in \mathcal{N}$
cm$\{s,p\}$	$-\log[P(I_p\|1)P(d_p\|1)]$	$p \in \mathcal{V}$
	∞	$p \in \mathcal{O}$
	0	$p \in \mathcal{B}$
cm$\{p,t\}$	$-\log[P(I_p\|0)P(d_p\|0)]$	$p \in \mathcal{V}$
	0	$p \in \mathcal{O}$
	∞	$p \in \mathcal{B}$

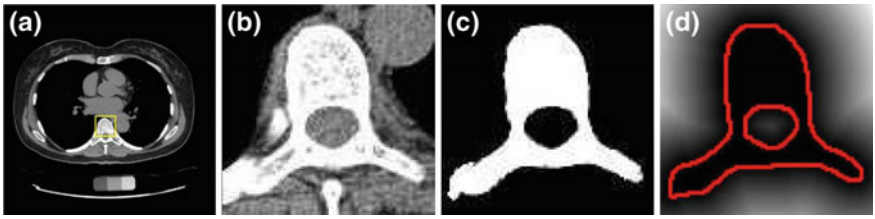


Fig. 34 An example of the initial labeling. **a** A CT slice of vertebral body. (*yellow box* illustrates detected VB region). **b** Detection of the VB region. **c** The initial labeling, $\mathbf{f}^*$. **d** The SDF of the initial segmentation which is used in the registration phase (see Algorithm 3). *Red color* shows the zero level contour

Algorithm 3 *Given*: The input VB set of images, the ESP (**J** as a source information), the probabilistic 3D shape model (**d**).

Objective: To obtain the desired labeling (**f**) using the required transformation matrix (**T**).

1. Detect the VB region using [40]
2. Obtain the initial segmentation ($\mathbf{f}^*$) using graph cuts which integrates the intensity and spatial interaction models only.
3. Register the shape prior to the initially segmented volume. **J** and $\mathbf{f}^*$ will be the source and target information, respectively. After the transformation, the embedded shape model and its features are described as follows:

- After each voxel $p \in \mathcal{P}_s$ is transformed to the new voxel $\widehat{p}$, the shape model is registered to the volume domain. We obtain new $\mathcal{O}^{new}$, $\mathcal{B}^{new}$, and $\mathcal{V}^{new}$.
- The object/variability surface $\mathbf{C}_{\mathcal{OV}}^{new}$ is updated as well.
- Hence, new iso-surfaces at the same distances, they have the same probabilistic distance value with the iso-surfaces which are obtained

before the registration. An example of our registration step is shown in Fig. 35.

4. Compute the final segmentation (**f**) using graph cuts: where the gray level marginal densities of the VB and its background are approximated using the proposed LCG model. Then we use a 3D graph where each vertex in this graph represents a voxel in the VB volume. Then we define the weight of each edge as shown in Table 5. After that, we get the optimal segmentation surface between the VB and its background by finding the minimum cost cut on this graph.

2.5.5 Experimental Results

We test the proposed segmentation framework on CT of human lumbar and thoracic spine data. The experiments are conducted on 30 data sets for which we have the ground truths. The real data sets were scanned at 120 kVp and 2.5 mm slice

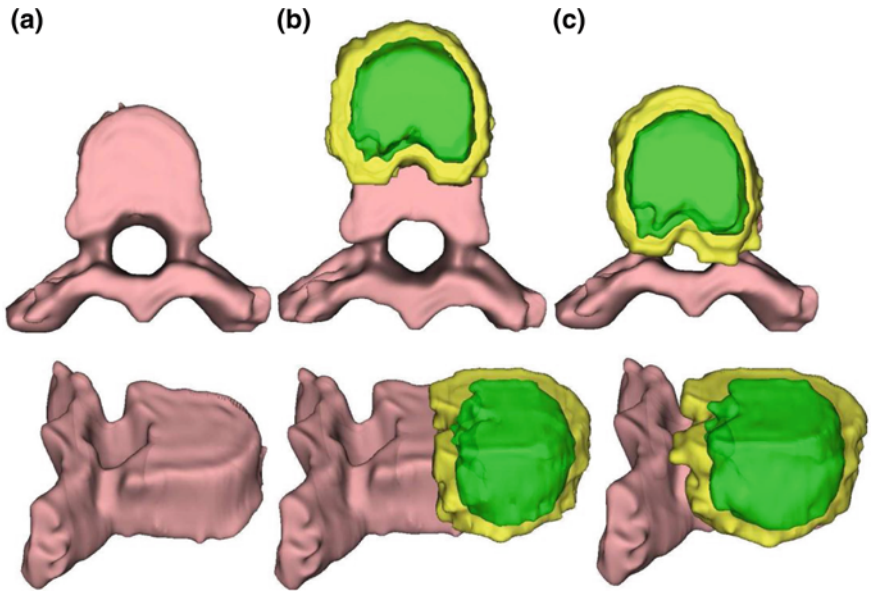


Fig. 35 The registration step. **a** The testing volume (or initially segmented volume) as shown with the *pink* color. **b** The testing and 3D shape prior before the registration. The *green* and *yellow* colors represent the object ($\mathcal{O}$) and variability ($\mathcal{V}$) regions, respectively. **c** The testing and 3D shape prior after the registration. The new object and variability regions are named as $\mathcal{O}^{new}$ and $\mathcal{V}^{new}$, respectively (see Algorithm 3). Each row shows different views

thickness. For sake of comparison, VBs are segmented using other alternatives: the graph cuts [36] without shape constrained ($\mathcal{A}_1$), statistical level sets ($\mathcal{A}_2$) method [41], and the b-spline based interpolation ($\mathcal{A}_3$).

Evaluation

We calculate the percentage segmentation error from the ground truth in order to evaluate the segmentation results as follows:

$$\varepsilon\% = 100(1 - \frac{S_a \cap S_m}{S_a \cup S_m}) \tag{42}$$

where S_m and S_a represent manually and automatically segmented VBs, respectively.

The statistical analysis of our method is shown in the Table 6. In this table the results of the proposed segmentation method and other three alternatives are shown. The average error of the VB segmentation on 30 clinical data sets is 6.8 % for the proposed method. Notice that it is difficult to separate the VB and spine processes because they have very similar gray level information. However, our shape model successfully extracts the spine processes. While other alternative methods fail to completely separate the processes and so they have huge precision error, which may change the BMD measurements. Figure 36 shows an example of 3D segmentation results of all tested methods for a clinical data set. The misclassified voxels, in this figure, are represented by red color.

Validation

ESP, which was scanned at 120 kVp and 0.75 mm slice thickness, is accepted as a standard for quality control in bone densitometry [3]. Therefore, we segment ESP in order to evaluate the proposed segmentation algorithm. The proposed approach accurately segments the VB without its processes and with a segmentation error 2.6 % as shown in Fig. 37.

Table 6 Accuracy and time performance of our VB segmentation on 30 data sets

	Algorithm			
Error %	Our	$\mathcal{A}_1$	$\mathcal{A}_2$	$\mathcal{A}_3$
Min. error, %	1.5	5.7	6.5	15.2
Max. error, %	17.3	83.2	91.2	97.7
Mean error, %	6.8	37.8	43.1	51.6
Stand. dev.,%	4.3	28.6	31.1	33.4
Average time, seconds	9.1	8.3	41.5	8.9

Average volume 512 × 512 × 10

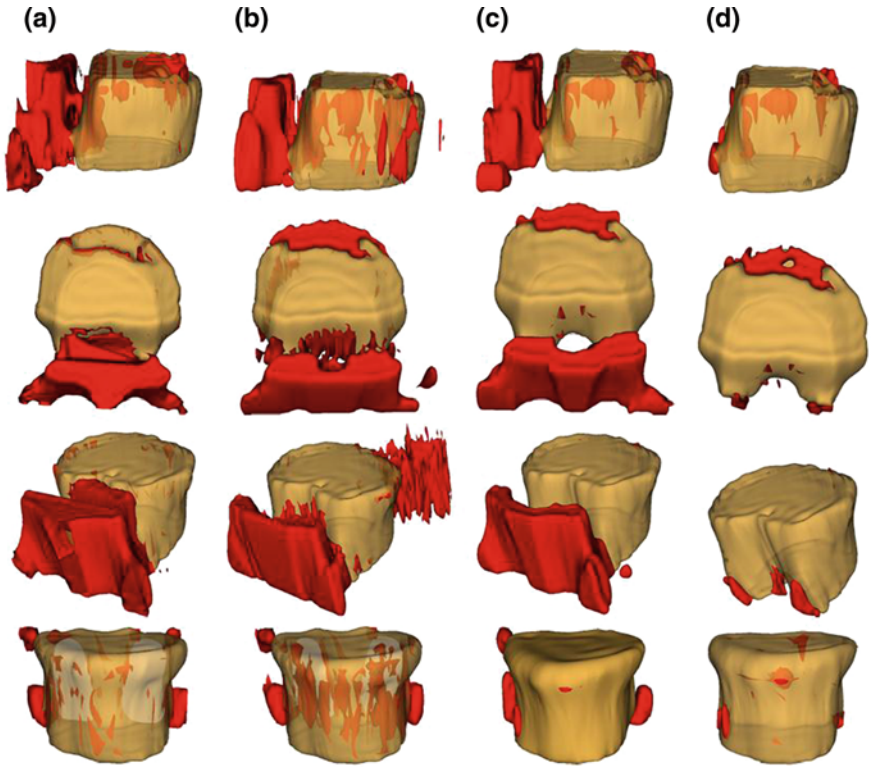


Fig. 36 Examples of 3D segmentation results of clinical data sets using the four different methods. **a** The results of b-spline based interpolation. **b** The results of statistical level sets. **c** The results of Graph cuts without shape constraints. **d** The results of the proposed method. The *red color* shows the segmentation errors

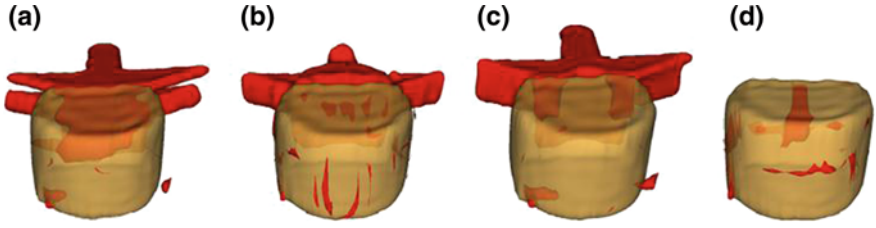
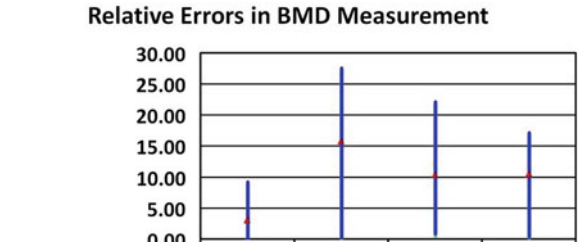


Fig. 37 3D segmentation results of ESP using the four different methods. **a** the result of b-spline based interpolation. **b** The result of statistical level sets. **c** The result of Graph cuts without shape constraints. **d** The result of the proposed method. The *red color* shows the segmentation errors



	Our	A1	A2	A3
max (%, mg/cc)	9.25	27.59	22.18	17.22
min (%, mg/cc)	0.03	0.35	0.89	0.44
mean (%, mg/cc)	3.03	15.66	10.26	10.40

Fig. 38 Precision errors of BMD measurement of each method. The BMD measurement of our method has the lowest error and standard deviation. VBs are subsequently segmented using the graph cuts without shape constrained (*A1*), statistical level sets (*A2*) methods, and the b-spline based interpolation (*A3*)

Bone Mineral Density Measurements

The main goal of this work is to accurately separate VB from around processes in order to obtain the BMD measurements with high trueness and precision from volumetric CT datasets. Thirty volmetric VBs from thoracic and lumbar spine are used in our experiments. For comparison purposes, the BMD measurements for each segmentation method are obtained. The relative errors in BMD measurements for each method are shown in Fig. 38. Our proposed method achieves an average BMD precision error %3.03. This reflects how accurate the proposed segmentation approach is.

2.6 Segmentation Using PCA-based Shape Model, Gaussian-based Intensity Model, and Asymmetric Gibbs potential-based Spatial Interaction Model

In this section, we present another idea to segment vertebral bodies. As pre-processing steps, we use the Matched filter to detect the region of interest and manual VB separation. In the second phase, we obtain initial labeling ($\mathbf{f}^*$) using the graph cuts which integrates the intensity and spatial interaction models. Finally, we register the initial labeled image and the shape priors to obtain the optimum labeling. To obtain the shape priors, we use the 2D-PCA on all training images. Figure 39 summarizes the main components of our framework. The following sections give more details about the shape model construction and the segmentation method. This method generally works in two dimensional space.

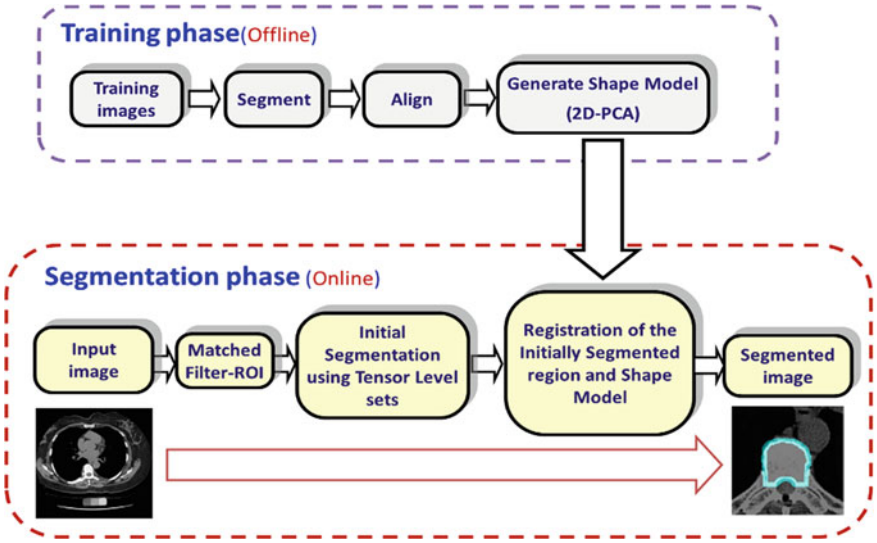


Fig. 39 Our proposed shape-based segmentation framework

2.6.1 Shape Modeling

The shape is represented using the signed distance function. Let $\mathbf{I}: \Omega \rightarrow R$ be an $n-D$ image usually $n=2$ or $n=3$, $\phi: \Omega \rightarrow R$ be a function that refers to a distance function representation for a given shape/contour $\mathcal{S}$ where $\Omega \subset R^n$ be an image domain which is bounded. The shape can be represented as follows:

$$\phi_s(x,y) = \begin{cases} 0, (x,y) \in \mathcal{S} \\ -ED((x,y),\mathcal{S}) > 0, (x,y) \in \mathcal{R}_{\mathcal{S}} \\ +ED((x,y),\mathcal{S}) < 0, (x,y) \in \Omega - [\mathcal{R}_{\mathcal{S}}] \end{cases} \tag{43}$$

where $\mathcal{R}_{\mathcal{S}}$ represents the inside region of the shape $\mathcal{S}$. Let (u,v) represents an pixel location on $\mathcal{S}$. For $\forall(x,y) \in \phi$, the distance between any (x,y) point and its nearest surface point can be calculated as follows:

$$ED((x,y),\mathcal{S}) = \min_{(u,v)\in\mathcal{S}} \sqrt{(u-x)^2 + (v-y)^2}. \tag{44}$$

Figure 40 shows an example of a shape representation using the distance function.

As opposed to conventional PCA, 2D-PCA is based on 2D matrix rather than 1D vector. This means that, the image does not need to be pre-transformed into a vector [42]. In addition, the image covariance matrix ($\mathbf{G}$) can be directly constructed using the original image matrices. As a result, 2D-PCA has two important advantages

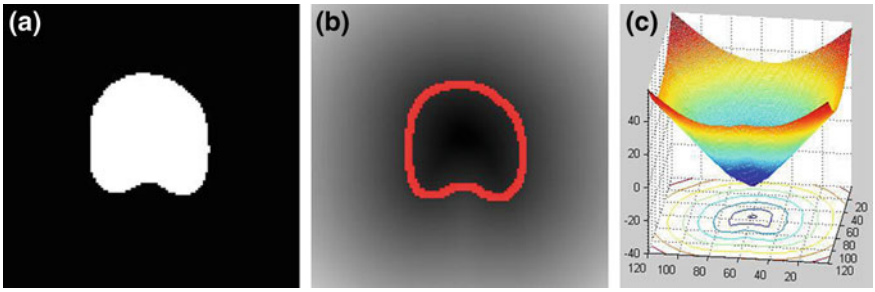


Fig. 40 An example representation of signed distance function. **a** A vertebral body shape. **b** Signed distance function as an image intensity representation. **c** Level set representation of SDF

over PCA. First, it is easier to evaluate $\mathbf{G}$ accurately since its size using 2D-PCA is much smaller. Second, less time is required to determine the corresponding eigenvectors [43].

2D-PCA projects an image matrix $\mathbf{X}$, which is an $m \times n$ matrix onto a vector, $\mathbf{b}$, which is an $n \times 1$ vector, by the linear transformation.

$$\mathbf{y} = \mathbf{X}\mathbf{b}. \tag{45}$$

Suppose that there are M training images, the ith training image is denoted by $\mathbf{X}_i$, (i = 1,2,…,M) and the average image of all training samples is denoted by $\bar{\mathbf{X}} = \frac{1}{M}\sum_{i=1}^{M}\mathbf{X}_i$. Then, let us define the image covariance matrix $\mathbf{G}$ [43]:

$$\mathbf{G} = \frac{1}{M}\sum_{i=1}^{M}(\mathbf{X}_i - \bar{\mathbf{X}})^t(\mathbf{X}_i - \bar{\mathbf{X}}). \tag{46}$$

It is clear that, the matrix $\mathbf{G}$ is $n \times n$ nonnegative definite matrix.

Similar to PCA, the goal of 2D-PCA is to find a projection axis that maximizes $\mathbf{b}^t\mathbf{G}\mathbf{b}$. The optimal K projection axes $\mathbf{b}_k$, where k = 1,2,…,K, that maximize the above criterion are the eigenvectors of $\mathbf{G}$ corresponding to the largest K eigenvalues. For an image $\mathbf{X}$, we can use its reconstruction $\tilde{\mathbf{X}}$ defined below to approximate it.

$$\tilde{\mathbf{X}} = \bar{\mathbf{X}} + \sum_{k=1}^{K}\mathbf{y}_k\mathbf{b}_k^t, \tag{47}$$

where $\mathbf{y}_k = (\mathbf{X} - \bar{\mathbf{X}})\mathbf{b}_k$ is called the kth principal component vector of the sample image $\mathbf{X}$. The principal component vectors obtained are used to form an $m \times K$ matrix $\mathbf{Y}$ = [$\mathbf{y}_1$,$\mathbf{y}_2$,…,$\mathbf{y}_K$] and let $\mathbf{B}$ = [$\mathbf{b}_1$,$\mathbf{b}_2$,…,$\mathbf{b}_K$], then we can rewrite (47) as:

$$\tilde{\mathbf{X}} = \bar{\mathbf{X}} + \mathbf{Y}\mathbf{B}^t. \tag{48}$$

However, one disadvantage of 2D-PCA (compared to PCA) is that more coefficients are needed to represent an image. From (48), it is clear that dimension of the 2D-PCA principal component matrix $\mathbf{Y}$ ($m \times K$) is always much higher than PCA. To reduce the dimension of matrix $\mathbf{Y}$, the conventional PCA is used for further dimensional reduction after 2D-PCA.

Now, let the training set consists of M training images $\{I_1,\ldots, I_M\}$; with SDFs $\{\Phi_1,\ldots,\Phi_M\}$. All images are binary, pre-aligned, and normalized to the same resolution. As in [42], we obtain the mean level set function of the training shapes, $\bar{\Phi}$, as the average of these M signed distance functions. To extract the shape variabilities, $\bar{\Phi}$ is subtracted from each of the training SDFs. The obtained mean-offset functions can be represented as $\{\widehat{\Phi}_1,\ldots,\widehat{\Phi}_M\}$. These new functions are used to measure the variabilities of the training images. We use 80 training VB images with 120 × 120 pixels in our experiment. According to (46), the constructed matrix $\mathbf{G}$ will be:

$$\mathbf{G} = \frac{1}{M}\sum_{i=1}^{M=80} \widehat{\Phi}_i^t \widehat{\Phi}_i. \tag{49}$$

The goal of 2D-PCA is to find the optimal K eigenvectors of $\mathbf{G}$ corresponding to the largest K eigenvalues. The value of "K" helps to capture the necessary shape variation with minimum information. Experimentally, we find that, the minimum suitable value is $K = 10$ [44]. Less than this value, the accuracy of our segmentation algorithm falls drastically below other alternatives. After choosing the eigenvectors corresponding to 10 largest eigenvalues, $\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_{10}$, we obtained the principal component matrix $\mathbf{Y}_i$ ($m = 120 \times K = 10$) for each SDF of our training set ($i = 1, 2, \ldots, 80$). For more dimensional reduction, the conventional PCA is applied on the principal components $\{\bar{\mathbf{Y}}_1,\ldots, \bar{\mathbf{Y}}_M\}$. It should be noted that, $\bar{\mathbf{Y}}$ is the vector representation of $\mathbf{Y}$. The reconstructed components (after retransforming to matrix representation) will be:

$$\tilde{\mathbf{Y}}_{\{l,h\}} = \mathbf{U}\mathbf{e}_{\{l,h\}}, \tag{50}$$

where $\mathbf{U}$ is the matrix which contains L eigenvectors corresponding to L largest eigenvalues λ_l, ($l = 1, 2,\ldots, L$), and $\mathbf{e}_{\{l,h\}}$ is the set of model parameters which can be described as [44]:

$$\mathbf{e}_{\{l,h\}} = h\sqrt{\lambda_l}, \tag{51}$$

where $l = \{1,\ldots, L\}$, $h = \{-\mu,\ldots, \mu\}$, and μ is a constant which can be chosen arbitrarily (in our experiments, we chose $L = 4$, $\mu = 3$). The new principal components of training SDFs are represented as $\{\tilde{\mathbf{Y}}_1,\ldots,\tilde{\mathbf{Y}}_N\}$ instead of $\{\mathbf{Y}_1,\ldots,\mathbf{Y}_M\}$ where N is the multiplication of L and standard deviation in eigenvalues (the number of elements in h), i.e. $N = L(2\mu + 1)$ [42]. Given the set $\{\tilde{\mathbf{Y}}_1,\ldots,\tilde{\mathbf{Y}}_N\}$, the new projected training SDFs are obtained as follows:

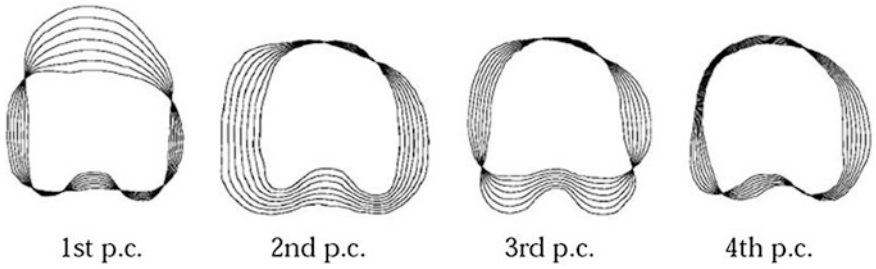


Fig. 41 Sampling up to 3 standard deviations (from $\mathbf{b}_{\{i,-3\}}$ to $\mathbf{b}_{\{i,3\}}$) along the first four principle components (p.c.) (where $i = \{1, 2, 3, 4\}$ for the first four p.c.) from the mean for a set of a 80 training vertebral body shapes

$$\tilde{\Phi}_n = \bar{\Phi} + \tilde{\mathbf{Y}}_n \mathbf{B}^t, \quad n = 1, 2, \ldots, N. \tag{52}$$

The shape variation of the first and fourth p.c.s is shown in Fig. 41. The shape variation decreases from the first to the fourth (or last) p.c. respectively. Hence, selection of L value helps to capture the necessary shape variation with minimum information.

Finally, the shape model is required to capture the variations in the training set. This model is considered to be a weighted sum of the projected SDFs as follows:

$$\Phi_{\mathbf{p}} = \sum_{n=1}^{N} w_n \tilde{\Phi}_n. \tag{53}$$

Let $\mathbf{w} = [w_1, \ldots, w_N]^t$ to be the weighting coefficient vector. By varying these weights, $\Phi_{\mathbf{p}}$ can cover all values of the training distance functions and, hence, the shape model changes according to all of the given images [44].

2.6.2 Segmentation Approach

To estimate the initial labeling **f***, we use the graph cuts which integrates the LCG and MGRF methods as we discussed above. To segment vertebrae, we initially labeled the volume based on its gray level probabilistic model. Then we create a weighted undirected graph with vertices corresponding to the set of volume voxels $\mathcal{P}$, and a set of edges connecting these vertices. Each edge is assigned a nonnegative weight. The graph also contains two special terminal vertices s (source) "vertebrae", and t (sink) "background". Consider a neighborhood system in $\mathcal{P}$, which is represented by a set $\mathcal{N}$ of all unordered pairs $\{p, q\}$ of neighboring voxels in $\mathcal{P}$. Let $\mathcal{L}$ the set of labels {"0", "1"}, correspond to the vertebrae and background regions respectively. Labeling is a mapping from $\mathcal{P}$ to $\mathcal{L}$, and we denote the set of labeling by $\mathbf{f} = \{f_1, \ldots, f_p, \ldots, f_{|\mathcal{P}|}\}$. In other words, the label f_p, which is assigned to the

voxel $p \in \mathcal{P}$, segments it to vertebrae or background region. Now our goal is to find the initial segmentation, **f***, by minimizing the following energy function [45]:

$$E(\mathbf{f}^*) = \sum_{p \in \mathcal{P}} D_p(f_p) + \sum_{\{p,q\} \in \mathcal{N}} V(f_p, f_q). \tag{54}$$

$D(f_p)$ measures how much assigning a label f_p to voxel p disagrees with the voxel intensity, I_p, and $V(f_p, f_q)$ is the pairwise interaction model which represents the penalty for the discontinuity between voxels p and q. For more information see [45]. Initially segmented region is used to obtain the SDF ($\Phi_{\mathbf{f}^*}$) which is required in the next step.

To use the shape prior in the segmentation process, we need to register **f*** and the shape prior **p**. The objective of the shape registration problem is to find the point-wise transformation between any two given shapes α and β minimizing a certain energy function based on some dissimilarity measure.

In this chapter, we follow the similar notation scheme in [42]. Let us define the result by β that is obtained by applying a transformation **A** (with scale, rotation, and translation parameters) to a given contour/surface α (It is clear that β and α correspond to **f*** and **p**). The shape representation used in this work changes the problem from the 2D/3D shape to the higher dimensional representation. Hence, we will look for a transformation **A** that gives pixel-wise correspondences between the two shape representations Φ_α and Φ_β. For the 2D case, we assume that the transformation has scaling components, $\mathbf{S} = \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix}$, rotation angles $\mathbf{R} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$, and translations represented as $\mathbf{Tr} = [\mathcal{T}_x \quad \mathcal{T}_y]^t$. The transformation will be in the form **A(x) = SRx + Tr**. After scaling the components of the $\Phi_{\mathbf{f}^*}$ by **A,** the dissimilarity measure will be:

$$\mathbf{r} = \mathbf{SR}\Phi_{\mathrm{p}} - \Phi_{\mathrm{f}*}(\mathbf{A}). \tag{55}$$

and the squared magnitude of the above measure is summed over the image domain Ω to get an optimization energy function:

$$E(\Phi_{\mathbf{p}}, \Phi_{\mathbf{f}^*}) = \int_{\Omega} \delta_\varepsilon(\Phi_{\mathbf{p}}, \Phi_{\mathbf{f}^*}) \mathbf{r}^t \mathbf{r} d\Omega, \tag{56}$$

where δ_ε is an indicator function defined as:

$$\delta_\varepsilon(\Phi_{\mathbf{p}}, \Phi_{\mathbf{f}^*}) = \begin{cases} 0 & \text{if } min(|\Phi_{\mathbf{p}}|, |\Phi_{\mathbf{f}^*}|) > \varepsilon \\ 1 & \text{if } min(|\Phi_{\mathbf{p}}|, |\Phi_{\mathbf{f}^*}|) \leq \varepsilon \end{cases} \tag{57}$$

Due to δ_ε, all pixels of a distance (measured from the nearest point on the boundary) greater than ε are not considered in the energy optimization problem which reduces the computational time of our problem (Narrow-banding effect).

After applying the gradient descent method [26], it is clear that:

$$\begin{aligned}\frac{d}{dt}s_i &= 2\int_\Omega \delta_\varepsilon(\Phi_{\mathbf{p}}, \Phi_{\mathbf{f}^*})\mathbf{r}^t\left[\nabla_{s_i}\mathbf{S}\Phi_{\mathbf{p}} - \nabla\Phi^t_{f^*}\nabla_{s_i}\mathbf{A}\right]d\Omega,\\ \frac{d}{dt}\theta_i &= 2\int_\Omega \delta_\varepsilon(\Phi_{\mathbf{p}}, \Phi_{\mathbf{f}^*})\mathbf{r}^t\left[\nabla\Phi^t_{\mathbf{p}}\nabla_{\theta_i}\mathbf{A}\right]d\Omega,\\ \frac{d}{dt}\mathcal{T}_i &= 2\int_\Omega \delta_\varepsilon(\Phi_{\mathbf{p}}, \Phi_{\mathbf{f}^*})\mathbf{r}^t\left[\nabla\Phi^t_{\mathbf{f}^*}\nabla_{\mathcal{T}_i}\mathbf{A}\right]d\Omega,\end{aligned} \tag{58}$$

where $s_i \in \{s_x, s_y\}$, $\theta_i \in \{\theta_x, \theta_y\}$ and $\mathcal{T}_i \in \{\mathcal{T}_x, \mathcal{T}_y\}$ of the transformation $\mathbf{A}$. Regarding to the weighting coefficients w_n's, and similar to [26], the energy function is a quadratic function of this weights, which leads to a closed-form when the derivatives with respect to the weights are zeros:

$$\Psi\mathbf{w} = \Lambda, \tag{59}$$

where $\mathbf{\Lambda}$ is a column vector of size N and $\mathbf{\Psi}$ is and $N \times N$ matrix. Their elements are calculated as follows [42]:

$$\Lambda_i = \int_\Omega \delta_\varepsilon(\Phi_{\mathbf{p}}, \Phi_{\mathbf{f}^*})[\mathbf{S}\Phi_{\mathbf{f}^*} - \bar{\Phi}(\mathbf{A})]^t\left[\tilde{\Phi}_i - \bar{\Phi}(\mathbf{A})\right]d\Omega, \tag{60}$$

$$\Psi_{ij} = \int_\Omega \delta_\varepsilon(\Phi_{\mathbf{p}}, \Phi_{\mathbf{f}^*})\left[\tilde{\Phi}_j(\mathbf{A}) - \bar{\Phi}(\mathbf{A})\right]^t\left[\tilde{\Phi}_i(\mathbf{A}) - \bar{\Phi}(\mathbf{A})\right]d\Omega, \tag{61}$$

$\forall(i, j) \in [1, N] \times [1, N]$. Using unique training shapes (with variability not identical) guarantees that $\mathbf{\Psi}$ is a positive definite matrix avoiding singularity.

2.6.3 Experimental Results

We tested our algorithm on 500 CT slices/25 VBs which are obtained from 15 different patients. The goal is to segment the VB region correctly. The segmentation accuracy and robustness of our framework are tested on the phantom named as the ESP as well as the clinical datasets. All algorithms are implemented using Matlab® 7.[1]

[1] All algorithms are run on a PC with a 2 GHz Core i7 Quad processor with 6 GB RAM.

To assess the proposed method under various challenges, we added a zero mean Gaussian noise with different signal-to-noise ratios (SNR)—from 0 to 100 dB—to our CT images. The segmentation accuracy is measured for each method using the ground truths. It should be noted that the ground truths are validated by a radiologist. We calculate the percentage segmentation accuracy (*Acc)* as follows:

$$Acc\% = 100 * (1 - \frac{FP + FN}{The\ total\ number\ of\ slice\ pixels}), \tag{62}$$

where *FP* represents the false positive (i.e. the total number of the misclassified pixels of the background), and *FN* is the false negative (i.e. the total number of the misclassified pixels of the object).

We used a variety of methods to measure the accuracy of our framework. First, we used the visual inspection to evaluate the segmentation quality of our approach. Figure 42 compares the results of different examples for the initial segmentation step using the scalar level set model [46] and the graph cut method [45] which is used in our proposed framework. As shown in this figure, the scalar level sets method fails to segment the whole vertebra in many cases. However, the graph cut approach can segment them well. Additionally, the boundaries detected by scalar level sets are not smooth, and some obvious boundaries are not detected. The graph cut method segments the image accurately. Figure 43 shows various segmentation results of three different methods applied on some clinical datasets. These methods are: (i) The graph cut segmentation (identical to initial labeling in our algorithm), (ii) The PCA based segmentation described in [47], and (iii) Our 2D-PCA based segmentation. The segmentation accuracies of the 2D-PCA based results shown in row (iii) are: 96.8, 92.6, 91.2 and 93.6 % respectively.

For PCA based results in row (ii), the segmentation accuracies are: 89.3, 87.4, 85.6, and 84.5 % respectively. It is clear that our method is more accurate than the method in [47]. Figure 44 shows the segmentation results of the ESP using (i) graph cut method and (ii) our segmentation algorithm (graph cut + shape prior) under different noise level. With our proposed approach, we obtain much better results compared to the graph cut only. Figure 45 studies the effect of the initialization on our proposed framework. Results indicate that the performance of our method is almost constant with different initialization parameters.

To quantitatively demonstrate the accuracy of our approach, we calculate the average segmentation accuracy of our segmentation method on 500 CT images under various signal-to-noise ratios and compare the results with the PCA based segmentation method in [47]. Again, as mentioned before, our 2D-PCA based framework outperforms the conventional PCA described in [47] as shown in Fig. 46a. Additionally, Fig. 46b studies the effect of choosing the number of the projected training shapes *N* on the segmentation accuracy. From this figure, we can conclude that the performance of 2D-PCA is better than the conventional PCA under the same number of training shapes. In another word, to get the same accuracy of PCA framework, the 2D-PCA needs fewer training shapes.

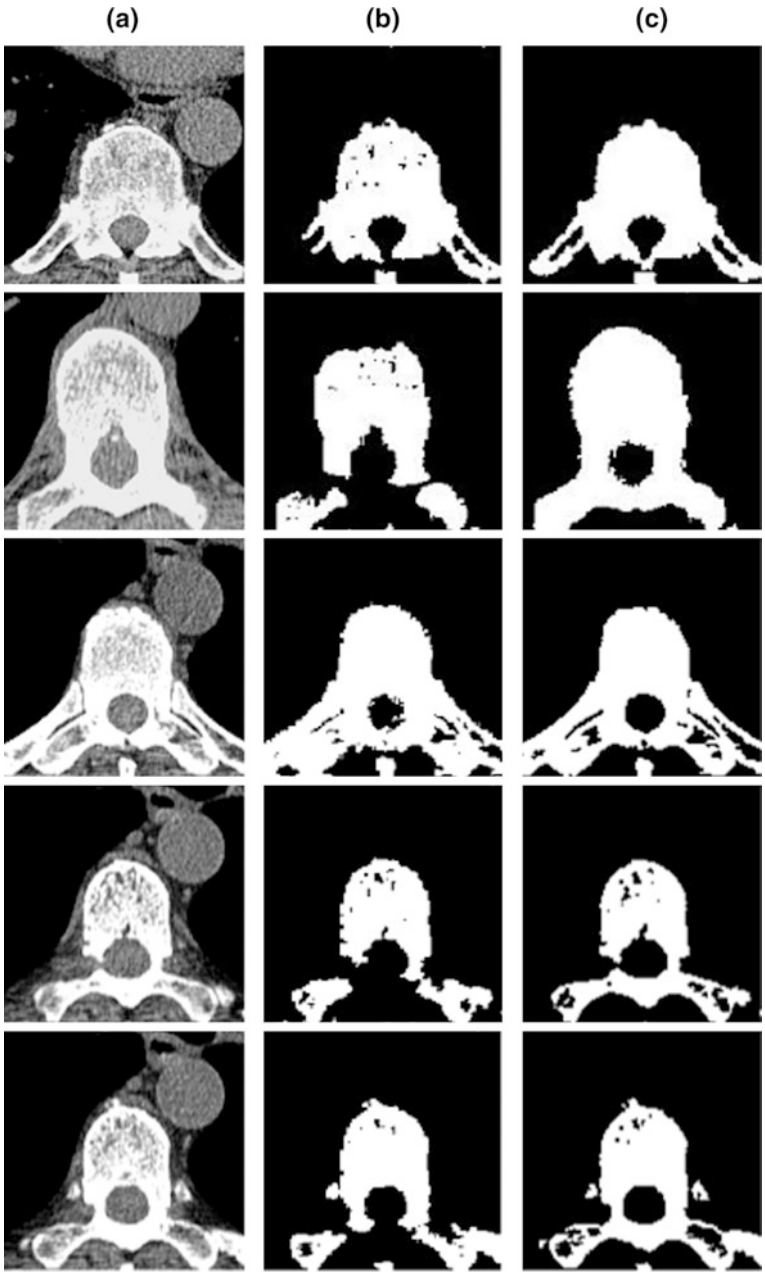


Fig. 42 Comparison between the intensity based segmentation (initial labeling) using: **b** Scalar level sets model [46], and **c** graph cut method [45]

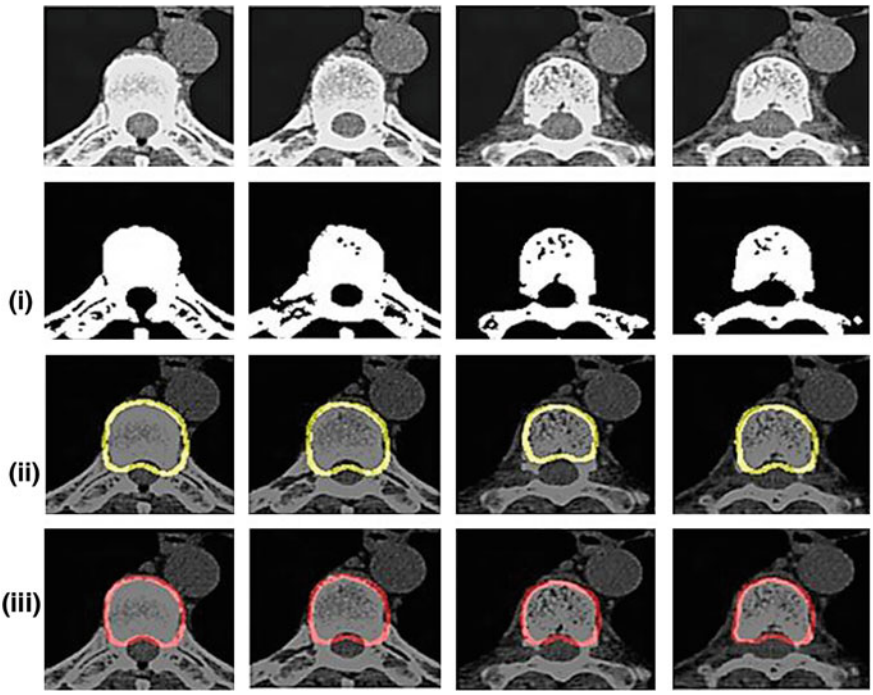


Fig. 43 Segmentation results of three different methods: (*i*) Using graph cuts only, (*ii*) Method described in [47], (*iii*) Our 2D-PCA based segmentation

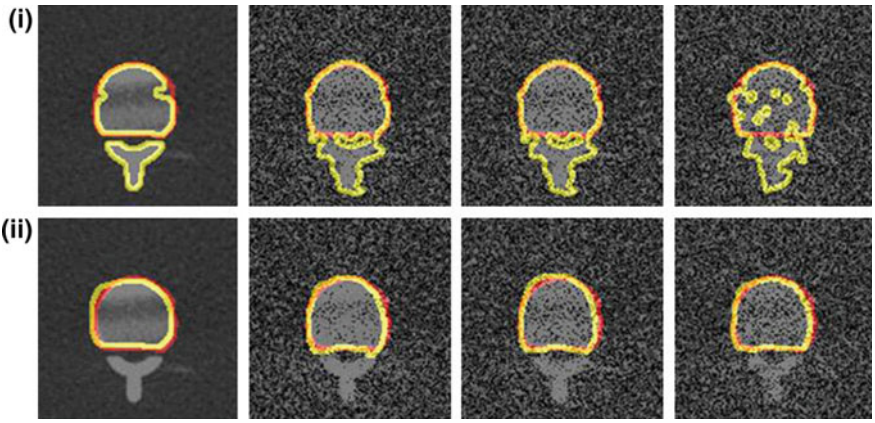


Fig. 44 Segmentation results of the ESP under different noise levels (*i*) using graph cut only. (*ii*) Our algorithm (graph cut + shape prior). The *red* and *yellow colors* show the contour of the gold standards and segmented regions

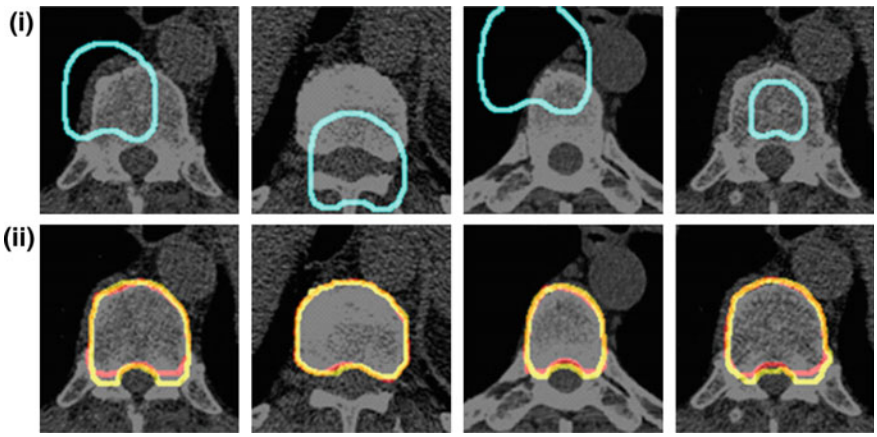


Fig. 45 Segmentation results with various shape initialization. (*i*) the initial shape prior, and (*ii*) is the final results. The *red* and *yellow colors* show the contour of the gold standards and segmented regions

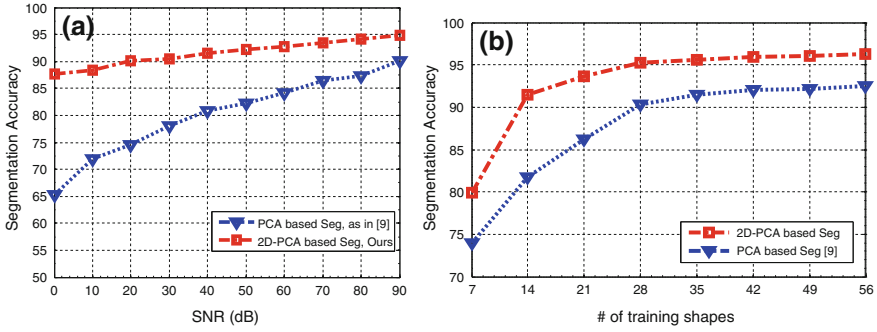


Fig. 46 **a** The average segmentation accuracy of different segmentation methods on 500 CT images under various signal-to-noise ratios. **b** The effect of choosing the number of the projected training shapes N on the segmentation accuracy

3 Discussion and Conclusion

In this chapter, frameworks which are robust under segmentation challenges, appropriate for a clinical workflow, and have theoretical novelty are proposed. This work is validated with various noise levels and compared with several alternative methods. To transfer the developed software into the clinical usage, more experiments on increased number of data sets are necessary.

One of the most important contributions of this study is to offer a segmentation framework which can be suitable to the clinical works with acceptable results. The proposed method completes the VB segmentation in very low execution time. It

should be noted that any parallel programming or graphical acceleration option were not used in this work. However, if needed, the execution time can be reduced using such methods.

The end-plate slices are segmented successfully thanks to the shape registration process although the shape model is obtained using in-plane slices. Further works are suggested to include the rotational transformation in the shape registration process for the sagittal plane to enhance the results and capture more fine details in the end-plate slices. In some cases, small portions of spinal processes and ribs are segmented erroneously. Future work can be including to investigate to reduce the misclassifications.

Possible works to estimate how the segmentation quality affect the BMD measurements and fracture analysis can be analyzed. To assist the VB fracture analysis, an automated point correspondence detection algorithm, such as scale-invariant feature transform (SIFT), can be tested to detect the VB height changes. In this problem, the corresponding points on the same patient, which is scanned at specific time intervals, should be detected successfully.

References

1. Klinder T, Ostermann J, Ehm M, Franz A, Kneser R, Lorenz C (2009) Automated model-based vertebra detection, identification, and segmentation in CT images. Med Image Anal 13:471–482
2. Kang Y, Klaus E, Kalender WA (2003) A new accurate and precise 3-D segmentation method for skeletal structures in volumetric CT data. IEEE Trans Med Imaging 5:586–598
3. Kalender WA, Felsenberg D, Genant HK, Fischer M, Dequeker J, Reeve J (1995) The European Spine Phantom—a tool for standardization and quality control in spinal bone mineral measurements by DXA and QCT. Eur J Radiol 20(2):83–92
4. Mastmeyer A, Engelke K, Fuchs C, Kalender WA (2006) A hierarchical 3D segmentation method and the definition of vertebral body coordinate systems for QCT of the lumbar spine. Med Image Anal 10(4):560–577
5. Kaminsky J, Klinge P, Bokemeyer M, Luedemann W, Samii M (2004) Specially adapted interactive tools for an improved 3D-segmentation of the spine. Comput Med Imaging Graph 28(3):119–127
6. Roberts M, Cootes T, Adams J (2005) Vertebral shape: automatic measurement with dynamically sequenced active appearance models. Proc Int Conf Med Image Comput Comput Assist Interv 2:733–740
7. Tan S, Yao J, Ward MM, Yao L, Summers RM (2008) Computer aided evaluation of ankylosing spondylitis using high-resolution CT. IEEE Trans Med Imaging (TMI) 27 (9):1252–1267
8. Criscob JJ, Sebastiana TB, Teka H, Kimia BB (2003) Segmentation of carpal bones from CT images using skeletally coupled deformable models. Med Image Anal 7(1):21–45
9. Kim Y, Kim D (2009) A fully automatic vertebra segmentation method using 3D deformable fences. Comput Med Imaging Graph 33:343–352
10. Leventon ME, Grimson WEL, Faugeras O (2000) Statistical shape influence in geodesic active contours. In: Proceedings of IEEE CVPR, pp 1316–1324
11. Chen Y, Thiruvenkadam S, Huang F, Wilson D, Geiser EA, Tagare HD (2001) On the incorporation of shape priors into geometric active contours. In: IEEE VLSM, pp 145–152

12. Huang X, Metaxas D, Chen T (2004) Metamorphs: deformable shape and texture models. In: Proceedings IEEE CVPR, pp 496–503
13. Besag JE (1986) On the statistical analysis of dirty pictures. J R Stat Soc B 48:259–302
14. North OH (1943) An analysis of the factors which determine signal/noise discrimination in pulsed-carrier systems. In: Proceedings of IEEE: RCA Labs, Princeton, NJ, Rep. PTR-6C, pp 1016–1063
15. Kumar BVKV, Savvides M, Xie C (2006) Correlation pattern recognition for face recognition. Proc IEEE 94(11):1963–1976
16. Aslan MS, Ali A, Rara H, Farag AA (2010) An automated vertebra identification and segmentation in CT images. In: Proceedings of 5th international conference on image processing (ICIP'10), pp 233–236
17. Kim Y, Kim D (2009) A fully automatic vertebra segmentation method using 3D deformable fences. Comput Med Imaging Graph 33:343–352
18. Moon T (1996) The expectation-maximization algorithm. IEEE Signal Process Mag 11:47–60
19. Borman S (2004) The expectation maximization algorithm: a short tutorial. Technical report available at http://www.seanborman.com/publications
20. Ali AM, Farag AA (2008) Automatic lung segmentation of volumetric low-dose CT scans using graph cuts. In: Proceedings of the international symposium on visual computing (ISVC'08), pp 258–267
21. Dubes RC, Jain AK (1989) Random field models in image analysis. J Appl Stat 16:131–164
22. Geman S, Geman D (1984) Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. IEEE Trans Pattern Anal Mach Intell 6:721–741
23. Derin H, Elliott H (1987) Modeling and segmentation of noisy and textured images using gibbs random fields. IEEE Trans Pattern Anal Mach Intell 9(1):39–55
24. Ali A, Farag AA, Gimel'farb G (2008) A comparative study of energy minimization methods for Markov random fields. In: Proceedings of ECCV
25. Cootes TF, Edwards GJ, Taylor CJ (1998) Active appearance models. In: Proceedings European conference on computer vision
26. El Munim HEA, Farag AA (2007) Curve/surface representation and evolution using vector level sets with application to the shape-based segmentation problem. IEEE Trans Pattern Anal Mach Intell 29(6):945–958
27. Tsai A, Yezzi A, Wells W, Tempany C, Tucker D, Fan A, Grimson WE, Willsky A (2003) A shape-based approach to the segmentation of medical imagery using level sets. IEEE Trans Med Imaging 22(2):137–154
28. Cremers D, Osher SJ, Soatto S (2006) Kernel density estimation and intrinsic alignment for shape priors in level set segmentation. Int J Comput Vision 69(3):335–351
29. Aslan MS, Mostafa E, Abdelmunim H, Shalaby A, Farag AA (2011) A novel probabilistic simultaneous segmentation and registration using level set. In: Proceedings of the international conference on image processing, pp 2161–2164
30. Aslan MS, Shalaby A, Farag AA (2013) Clinically desired segmentation method for vertebral bodies. In: IEEE 10th international symposium on biomedical imaging (ISBI), pp 840–843
31. Goldman LW (2007) Principles of CT: radiation dose and image quality. J Nucl Med Technol 35:213–225
32. Cootes TF, Edwards GJ, Taylor CJ (2001) Active appearance models. IEEE Trans Pattern Anal Mach Intell 23(6):681–685
33. Li C, Xu C, Gui C, Fox MD (2005) Level set evolution without re-initialization: a new variational formulation. In: Proceedings of the 2005 IEEE Computer Society conference on computer vision and pattern recognition
34. Ali AA, Farag AA, El-Baz AS (2007) Graph cuts framework for kidney segmentation with prior shape constraints. In: Proceedings of international conference on medical image computing and computer-assisted intervention (MICCAI'07), pp 384–392
35. Ali AM, Farag AA (2008) Density estimation using a new AIC-type criterion and the EM algorithm for a linear combination of gaussians. In: Proceedings of IEEE international conference on image processing (ICIP'08), pp 3024–3027

36. Ali AM, Farag AA (2008) A novel framework for ND multimodal image segmentation using graph cuts. In: 15th IEEE international conference on image processing, pp 729–732
37. Ali AM, Farag AA, Gimel'farb G (2008) Analytical method for MGRF potts model parameters estimation. In: Proceedings of the international conference on pattern recognition (ICPR-08)
38. Farag AA, El-Baz AS, Gimel'farb G (2006) Precise segmentation of multimodal images. IEEE Trans Image Process 15(4):952–968
39. Boykov Y, Kolmogorov V (2004) An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. IEEE Trans PAMI 26(9):1124–1137
40. Aslan MS, Ali A, Farag AA, Rara H, Arnold B, Xiang P (2010) 3D vertebral body segmentation using shape based graph cuts. In: Proceedings of the international conference on pattern recognition (ICPR-10)
41. Farag AA, Hassan H (2004) Adaptive segmentation of multi-modal 3D data using robust level set techniques. In: Proceedings of international conference on medical image computing and computer-assisted intervention, pp 143–150
42. Shalaby A, Aslan M, Abdelmunim H, Farag A (2012) 2D-PCA based shape prior for level sets segmentation framework of the vertebral body. In: Proceedings of the 6th Cairo international conference on biomedical engineering (CIBEC'12), Cairo, Egypt, pp 134–137
43. Yang J, Zhang D, Frangi AF, Yang JY (2004) Two-dimensional PCA: a new approach to appearance-based face representation and recognition. IEEE Trans Pattern Anal Mach Intell 26 (1):131–137
44. Aslan MS, Shalaby A, Abdelmunim H, Farag AA (2013) A probabilistic shape-based segmentation method using level sets. IET Comput Vision
45. Aslan MS, Ali A, Rara H, Arnold B, Farag AA, Fahmi R, Xiang P (2009) A novel 3D segmentation of vertebral bones from volumetric CT images using graph cuts. In: Advances in visual computing, Springer, Berlin, pp 519–528
46. Chan TF, Vese LA (2001) Active contours without edges. IEEE Trans Image Process 10 (2):266–277
47. Aslan MS, Abdelmunim H, Farag AA, Arnold B, Mustafa E, Xiang P (2011) A new shape based segmentation framework using statistical and variational methods. In: Proceedings of IEEE international conference on image processing (ICIP)

Part III
Image Guided Spine Intervention

Toward Virtual Modeling and Templating for Enhanced Spine Surgery Planning

Cristian A. Linte, Kurt E. Augustine, Jon J. Camp, Richard A. Robb and David R. Holmes III

Abstract Traditional 2D images provide limited use for accurate planning of spine interventions, due to their inability to display the complex 3D spine anatomy and close proximity of nerve bundles and vascular structures that must be avoided during the procedure. We have developed a platform for spine surgery planning that employs standard of care 3D pre-operative images and enables oblique reformatting and 3D rendering of individual or multiple vertebrae, interactive templating, and placement of virtual pedicle implants into the patient-specific CT data. Here we propose a combined surrogate metric—the Fastening Strength—to provide estimates of the optimal implant selection and trajectory based on implant dimension and bone mineral density of the displaced bone substrate. We conducted a retrospective clinical study based on pre- and post-operative data from four patients who underwent procedures involving pedicle screw implantation. We assessed the retrospective plans against the post-operative imaging data according to implant dimension, mean voxel intensity of implant trajectory, and Fastening Strength and showed consistency between the proposed plans and the post-operative procedure outcome. Our preliminary studies have demonstrated the feasibility of the platform in assisting the surgeon with the selection of appropriate size implant and trajectory that optimizes Fastening Strength, given the intrinsic vertebral geometry and bone mineral density. Herein we describe the platform infrastructure and capabilities, present preliminary studies conducted to assess impact on typical instrumentation procedures, and share our initial clinical experience in employing the proposed tool for the planning of several complicated spinal correction procedures for which the traditional planning approaches proved insufficient. Lastly, we also disseminate on several clinical cases and their post-operative assessment for which the proposed platform was employed by the surgical team.

C.A. Linte (✉)
Biomedical Engineering and Center for Imaging Science, Rochester Institute of Technology, Rochester, NY, USA
e-mail: Linte.Cristian@mayo.edu

K.E. Augustine · J.J. Camp · R.A. Robb · D.R. Holmes III
Biomedical Imaging Resource, Mayo Clinic, Rochester, MN, USA

© Springer International Publishing Switzerland 2015

S. Li and J. Yao (eds.), *Spinal Imaging and Image Analysis*,
Lecture Notes in Computational Vision and Biomechanics 18,
DOI 10.1007/978-3-319-12508-4_14

1 Introduction

Spinal fusion is a commonly performed procedure for a variety of conditions. Pedicle screw fixation for correction of spinal deformity has become the standard of care for stabilization of the thoracic and lumbar spine. The objective of the pedicle screw implantation procedure is to install an internal fixator for stabilization of injured vertebrae [1]. Precise screw placement is essential to avoid injury to adjacent neural structures. Patients with severe deformity or prior surgery present a challenge to the accurate placement of pedicle screws. Additionally, minimally invasive and percutaneous surgical techniques also present a greater challenge to accurate screw placement and require heavier reliance on intra-operative fluoroscopic imaging, which presents an occupational hazard for the surgeon and the operating-room (OR) staff [2]. However the techniques currently available for planning such interventions are sub-optimal. Until recently, such procedures have been traditionally planned using 2D radiographs, an approach which has proved inadequate for precise planning due to the complex 3D anatomy of the spinal column and the close proximity of the nerve bundles, blood vessels and viscera. The pedicles are anatomically close to the spinal nerve roots, forming the lateral borders of the vertebral canal and the superior and inferior margins of the intervertebral foramina [3]. The nerve roots pass directly caudal to the pedicles as they course through the respective intervertebral foramen [4–8]. Furthermore, both the sensory and motor intrathecal nerve roots follow closely the medial aspect of the pedicles and are located in the anterior-superior one third of the intervertebral foramen [4, 6, 7]. In addition, anterior to the vertebral bodies lie the aorta and vena cava, with branching of the common iliac vessels occurring in the lumbar region. Hence, penetration of the anterior cortex of the vertebral bodies could also lead to injury of one or more of these vessels. As such, significant care must be taken to avoid the risk of neural or vascular damage during intervention.

According to the Cleary et al. [9], challenges impeding the development of better guidance include adequate intra-operative imaging, fusion of images from multiple modalities, the visualization of oblique paths, percutaneous spine tracking, mechanical instrument guidance, and software architectures for technology integration. Intra-operative imaging using a high-performance mobile C-arm prototype has demonstrated a significant advance in spatial resolution and soft-tissue visibility, with the added benefit of reducing fluoroscopy reliance and enabling precise visualization via up-to-date images [10]. However, procedure planning must be conducted in the OR, using the peri-operatively acquired images, therefore adding to the procedure time.

Considering these limitations, it is critical for the surgeon to have access to superior images of the patient-specific anatomy that display the 3D relationships among these structures and enable intuitive, efficient and risk-free planning. As part of current clinical practice, 3D imaging scans, such as computed tomography (CT) and magnetic resonance imaging (MRI) are often ordered prior to spine correction procedures to help plan the intervention. During the planning process, the axial

images are reviewed and critical vertebrae are identified. The length of the vertebra is measured from the pedicles to the anterior surface of the vertebral body. Moreover, the width of the bone at the narrowest point of the pedicle is measured to ensure selection of screws which will not penetrate into the spinal column. The angle of approach is determined by an estimated deviation from the spinous process. Consistent with current clinical practice, the proposed screws and angles of insertion are documented, by hand, on a planning form. Nevertheless, the planning is limited to the review of the 2D axial slices of the anatomy. Often, the axial views cannot provide true measurements of the vertebral body or pedicle width and depth, which may in turn lead to inadequate decisions with regards to the implant size and trajectory.

In response to these challenges and driven by the motivation and insight of our orthopedic surgery collaborators, we have developed a clinician-friendly application that provides full 3D visualization for superior surgical planning. This application uses routine 3D CT or MR image data to generate detailed models and templates for better planning of pedicle screw instrumentation procedures.

The initial iteration of our spine surgery planning platform was developed simultaneously with the iPlan platform developed simultaneously by BrainLab [11] and features similar capabilities, including selection of virtual implants, trajectory planning via virtual templating, as well image segmentation, registration and overlay of multiple datasets into the templating workflow. Our platform was employed extensively within the Division of Orthopaedic Surgery at Mayo Clinic for a wide variety of non-routine clinical cases, especially pediatric cases of spine deformity correction procedures.

In this work we also describe how the initial platform is augmented by introducing and additional metric to optimize planning—the Fastening Strength. This metric is a surrogate measure of the screw holding power, as described in Sect. 3.2.3, and combines implant dimension, trajectory and bone strength into a single metric that enables pre-procedural assessment of each implanted screw. While the screw holding power is a well-known concept to the mechanical engineering design community, to our knowledge, this is its first application to implant placement and trajectory planning for surgical use. As further illustrated, its main benefit is the feasibility to also consider bone strength (i.e., density) when making decisions with respect to the implant size and trajectory, by providing a consistent relative measure in response to the dimensions, trajectory and bone mineral density characteristics of each implanted pedicle screw.

Herein we describe the platform infrastructure and capabilities, present preliminary studies conducted to assess impact on typical instrumentation procedures, and present the formulation of the Fastening Strength metric, its integration into the platform, as well, as its assessment against retrospective pre-procedural plans and post-procedural outcome in several cases. Lastly, we share our initial clinical experience in employing the proposed tool for the planning of several complicated spinal correction procedures for which the traditional planning approaches proved insufficient.

2 Spine Surgery Planning Platform Architecture

The Biomedical Imaging Resource (BIR) at Mayo Clinic has developed a clinical imaging software framework designed to provide powerful image visualization/analysis tools in an intuitive, easy-to-use interface [12]. Built upon a comprehensive, mature imaging toolkit called AVW [13], individual task-driven modules can be developed and easily added to the base software. The underlying architecture is based upon two concepts very familiar to physicians—Cases and Workflows. Each *case* is associated with a unique patient and a specific set of routine clinical tasks, or a *workflow*. This project uses a Spine Surgery Planning module developed to run within the newly developed clinician-centric application framework.

Designed by a team consisting of imaging/visualization scientists, software developers/engineers, and clinicians/surgeons, the goal at the outset was to develop tools that are powerful, yet very simple to use. Clinicians have limited time to spend learning new, complex software applications. Consequently, great care should be taken in designing advanced clinical tools to ensure widespread acceptance. The most effective way to develop such an application is to work closely with the end-user and include the targeted clinician, or "clinical champion" in the entire design process. Before a line of code was written, our collaborating orthopedic surgeon met with the development team to discuss the problems with the current methods and began outlining the requirements for the new surgical planning tools. For spine surgery, 3D visualization and manipulation is at the core of the new tools. Templating of implants in 2D has been a technique utilized in orthopedics for decades so it made sense to implement a new, advanced templating procedure that incorporates 3D imaging. Early design meetings, when the interface was first scribbled onto a whiteboard, included our clinical champion. After each teration of development, the clinical end-user reviewed the latest version with members of the development team and provided a list of recommended changes to be made to the tools. Given the busy schedule of most surgeons, finding time for repeated reviews can be extremely challenging but their input is essential to the development of tools that will be used by fellow surgeons. To successfully develop clinical tools in a timely manner, the team needs to be persistent as well as flexible. Design reviews do not take precedence over much in the working day of a surgeon so postponed meetings are routine. Reviewing the latest version on a laptop in the hospital cafeteria between surgical cases may turn out to be one's best opportunity to meet with a clinical collaborator in a several week span. In the end, the tools developed are sure to be usable by and useful to the intended clinical users, given their active involvement in the development process.

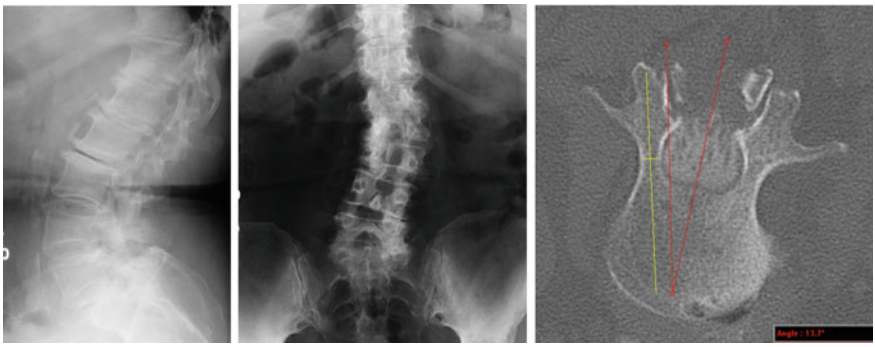


Fig. 1 On the *left* and *middle panels*, bi-planar radiographs are shown. Radiographs are sometimes used in planning spine surgeries with complex patient morphology. On the *right*, an axial image from a CT scan is shown with distance and angle measurement lines superimposed

3 Procedure Planning Workflow

3.1 Traditional Planning

Traditional methods of planning for corrective spinal surgery include the use of 2D images from radiographs or from axial, sagittal and coronal views of CT or MRI exams. Linear and angular measurements are made with simple tools from within image viewing programs, or directly on film using markers, cutouts and rulers.

Figure 1 shows a typical pair of radiographs and CT slice which would be used in the planning of a procedure. For this study, only CT data was used for planning as it provides more detail than a radiograph. During the planning process, the axial images are reviewed and critical vertebrae are identified. The length of the vertebra is measured (using a graphic line measurement tool) from the pedicles to the anterior surface of the vertebral body. In addition, the width of the bone at the narrowest point of the pedicle is measure in order to select screws which will not penetrate into the spinal column. The angle of approach is determined by an estimated deviation from the spinous process. Consistent with current clinical practice, the proposed screws and angles of insertion are documented, by hand, on a planning form.

3.2 Templating-Based Procedure Planning

3.2.1 Pre-operative Imaging

The SSP application runs on a standard desktop computer. The software can import data either directly from the file system or through an institutional PACS in the form of a high resolution CT scan acquired with standard imaging protocols.

Typical datasets consist of isotropic images with a 0.75 × 0.75 mm in-plane resolution and a 0.75 mm slice thickness. The pre-operative scan is imported into the SSP software [14], within which the surgeon "virtually" places the pedicle screws into the 3D image data, generating a virtual surgical plan which can be loaded up for visualization during the intervention. Moreover, the resulting surgical plan and image dataset can be further used to generate an appropriate anatomical model for 3D printing, resulting in a physical, 3D patient-specific model of the spine that can be used as a visual aid before and during the procedure.

3.2.2 Patient-Specific Virtual Templating

The 3D templating process is the repeated application of two steps for each vertebra of interest. First, to effectively plan spine surgery using 3D templating tools, it is necessary to reorient each vertebral body so the axial image plane runs perpendicular to its central axis. To accomplish this task, the user simply identifies a bounding box for each vertebra, using the sagittal and coronal views. The top and bottom sides of the bounding box are aligned with the vertebral endplates, making sure they extend far enough to include the entire vertebral body and any part of the implants that will extend outside the vertebra (e.g., the screw heads). This can quickly be carried out with placement and manipulation of a simple GUI tool with 2–3 mouse clicks in each of the two views, as illustrated in Fig. 2. The pedicle lengths and angles are determined in the local vertebral space to ensure that the measurements correctly represent the anatomy. Also during the vertebral identification process, it may be necessary to reorient each vertebral sub-volume into a consistent frame of references. Specifically, a rotation of the axial image may be

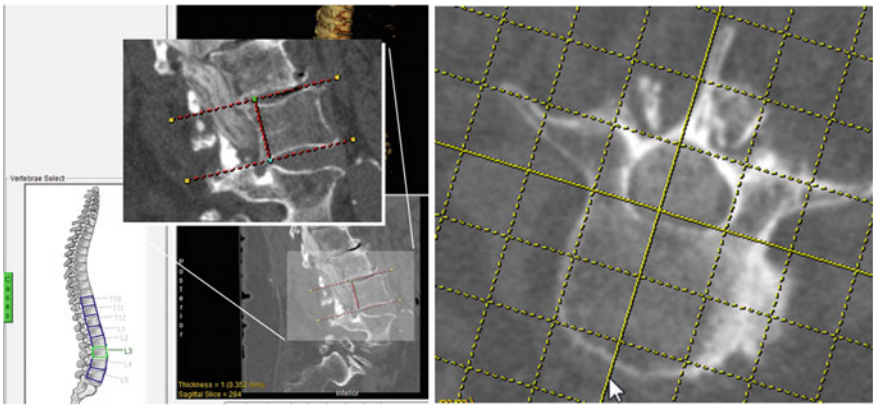


Fig. 2 Vertebral body extraction and alignment. During the process of vertebral body extraction, a user manually places I-beams around the vertebra of interest (*left*). After the vertebra has been extracted, it can be reoriented along the spinous process using an interactive grid (*right*)

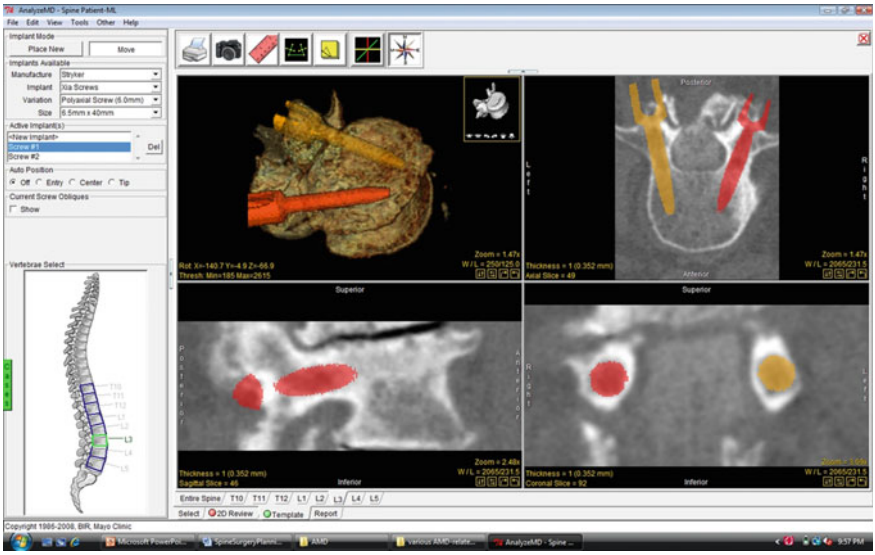


Fig. 3 Pedicle screw template placement. Screw templates are placed interactively into the image data with the mouse. Each template corresponds to a particular prosthetic implant manufacturer. The screws are evaluated in the orthogonal images and rendering to ensure that the screws are of the appropriate length and thickness

required if the spinous process does not line up parallel to the y-axis in that view. The realignment is also shown in Fig. 2.

In the second step, digital templates of screws are selected and inserted into the 3D data. During the selection step, the appropriate type of pedicle screw is chosen and a size is selected. The template is then placed into the axial image that includes the widest portion of the pedicle. The virtual pedicle screw can now be translated or rotated interactively in any of the three orthogonal views to achieve optimal placement within the vertebra, as shown in Fig. 3. Exact dimensions and angle placement for each screw placed is automatically recorded for use in the final report, which is the last step in the 3D spine surgery planning process. The report provides a list of each screw templated, including the vertebra in which they are placed, the manufacturer of the screw, the dimensions of the implant, and the precise location within the vertebral body based on the axial and sagittal angles. In addition to the implant list, a number of images are automatically generated and added to the report for visual verification.

3.2.3 Fastening Strength Formulation

According to strength of materials principles and theories of failure, each screw withstands a maximum force before it can be torn away from the material after its insertion. The holding power of a typical screw depends on the dimensions of the

screw, the threaded insertion depth, and the properties (typically characterized by specific gravity) of the material in which the screw is inserted. By transposing this theory to the pedicle screw implantation procedure, the holding power (i.e., fastening strength) of a pedicle screw is directly proportional to the screw diameter, the length of the threaded portion of the screw inserted within the bone, and the specific gravity of the pedicle body, typically characterized by its bone mineral density (BMD). Based on this relationship, we define Fastening Strength as a surrogate measure for the screw holding power that combines both implant dimension and trajectory estimated based on the virtually-templated images using the relationship below:

$$Fastening\ Strength = \int_0^L \int_0^{2\pi} \int_0^{D/2} r \cdot I(r, \theta, z) dr\, d\theta\, dz,$$

where L is the length of the in-bone threaded portion of the screw, D is the screw diameter, and $I(r; \theta; z)$ is the image intensity at each voxel within bone volume displaced by the virtual screw. The above relationship represents the intensity—area product evaluated in transverse slices (defined by the in-plane cylindrical coordinates—radial distance r and angular increment $d\theta$) throughout the extent of the insertion depth.

Studies [15] have revealed a linear correlation between the image intensity and BMD measurements, based on calibrations of known BMD CaHA (calcium hydroxyapatite) phantoms against the dynamic intensity range: BMD = $\alpha \cdot$ Intensity, where $\alpha = 0.8 \pm 0.03$. Moreover, the typical BMD of spinal cortical bone (hard shell coating the pedicle surface) was reported as 192 ± 10 mg/cm^3 [16], and the BMD of the cancellous bone (spongy bone near the pedicle core) was reported as ~ 140 mg/cm^3.

We refer to the Fastening Strength as a surrogate measure for the screw holding power primarily because the voxel intensity is used to characterize the bone mineral density (BMD) of the pedicle body segment displaced by the screw. Moreover, the Fastening Strength is not intended to be interpreted or employed as an absolute metric, but rather a relative measure to compare the expected holding power provided by implants of different dimensions and inserted along different trajectories within the bone.

3.2.4 Physical 3D Model Generation

After the pre-surgical plan is completed, the original CT data and the SSP results are imported into Analyze [17–20] for additional processing prior to 3D model printing. The spine is segmented in the CT images using basic thresholding. If the quality of the scans is low, additional manual techniques are required to correctly delineate the spine. Following the detailed spine segmentation step, the pedicle screw placement data generated by the SSP are incorporated by inserting

representative voids in the segmentation which correspond to the precise screw locations. The virtual spine model with the templated screw holes in place is tiled into a surface using an adaptive deformation algorithm [21–23] and exported as a stereo-lithography (STL) file. To improve the stability of the delicate spinal structure, a narrow ribbon is added to the anterior spine model before printing. The resultant STL file is printed using a ZCorp Spectrum Z510 printer. This printer can generate large-volume, full color models.

The pre-surgical plan report is used in advance to prepare the instrumentation inventory for the operation. The 3D patient-specific model, along with the report, is used in the procedure room to provide real-time visualization and guidance for accurate pedicle screw placement.

3.2.5 Proposed Surgical Planning Workflow

To better illustrate the functionality and capabilities of the spine surgery planning platform, we follow a hypothetical patient through the workflow associated with the proposed interventional planning protocol. Following diagnosis based on a routine CT scan, the patient is typically recommended for surgery, in which case additional imaging exams may be ordered, to better examine a specific region of interest in the spine anatomy. Based on the pre-operative spine image dataset, the surgeon or physician assistant will conduct the surgical plan using the proposed virtual platform. Each vertebral segment that needs to be instrumented will be realigned according to the true vertebral axis in order to enable true size measurements of the pedicles and vertebral body. These preliminary measurements of the pedicle width and length will serve as initial estimated of the virtual pedicle screws that are to be selected from the available database that compiles a wide variety of screws according to the specifications of several different manufacturers, so they closely match the vertebral anatomy. Each vertebral segment is then instrumented by "inserting" the selected virtual screw (in the form of an object map—a virtual representation of the physical screw) in the image dataset through the pedicle body mimicking the actual intra-operative implantation procedure. The position and orientation of each screw is then evaluated by panning through the image data containing the virtual pedicle screw to ensure the screw is fully contained within the pedicle (i.e., ne pedicle rupture) and does not interfere with the surrounding anatomy. In the event that more than one screw type (i.e., of close diameter and length) or several trajectories are permissible, the *Fastening Strength* is evaluated for each screw and each trajectory and based on the results, the screw and trajectory yielding the largest *Fastening Strength* will be selected, again, under the constraints that no interference exists between the screw and surrounding anatomy.

These steps are repeated for each vertebral segment that needs to be instrumented. Following completion of the plan, the platform provides a report that lists each instrumented vertebral segment specifying the implant gauge (i.e., screw diameter and length), implant trajectory (measured with respect to the axial and sagittal angles), and, if needed, the *Fastening Strength* computed for each implant. Note that

once a decision is made with respect to the size and trajectory of a specific implant, the *Fastening Strength* provides minimal information, as it is a relative measure based on which the implant trajectory and size is optimized and its absolute magnitude is meaningless unless compared against other values computed for the same implant.

Based on the output report, the instrumentation inventory is prepared for the upcoming procedure. Also, the virtual plan can be saved as either a surface or volume rendered model either including or excluding the virtual screws that can be displayed in the procedure room for analysis and review during the intervention. If a life-size model of the instrumented spine is required by the surgeon prior to the procedure, the virtual plan is saved as a stereo-lithography (STL) file and a physical model can be generated using a 3D printer or a rapid prototyping device.

In the event that a surgical navigation platform is employed to guide the intervention, the virtual plan can be registered to the patient's anatomy in the OR by using a variety of registration techniques—the most suitable and straight-forward being a landmark-based rigid body registration for each individual vertebral segment. This approach will minimize any uncertainties introduced by registering an entire region of the spine to the patient, provided a slightly different position or orientation of the patient between the pre-operative scan and intra-operative procedure. After registration, the drilling tool and pedicle implant can be guided and inserted according to the prescribed pre-operative plan.

4 Platform Evaluation and Validation

4.1 Assessment of 3D Templating Tool

A small retrospective pilot study was conducted to compare the traditional 2D method of spine surgery planning and the 3D templating method for pedicle screw placement. A cohort of 10 subjects was identified for the study, each having had a previous spinal procedure that included the implantation of pedicle screws in two or more vertebrae. Original preoperative plans were not available so two new plans were created for each subject based on the CT exams taken prior to the surgery. One was a plan based on the same type of 2D method used to carry out the original procedure. A second plan was created using the new 3D templating tools. Each of two participating surgeons generated separate plans for 5 of the 10 subjects. One participant was a skilled staff surgeon with 14 years of experience and the other was a fifth year resident. Post-surgical CT exams were also extracted from the patient record to compare both methods of planning with the post-surgical results. Metrics used in the evaluation included pedicle screw lengths, widths and angles. The pre-operative 2D planning data, pre-operative 3D planning, and post-operative images were analyzed by looking at:

(a) Differences between pre-operative 2D and pre-operative 3D
(b) Differences between pre-operative 2D and post-operative
(c) Differences between pre-operative 3D and post-operative
(d) Greater deviation between the two pre-operative to post-operative differences—from (b) and (c)
(e) Greatest deviations from (d) between resident and experienced staff surgeon

Several differences were observed in the 2D planning workflow and the 3D planning workflow. First, while best attempts were made to accurately extract the dimensions of vertebrae with the 2D approach, deformation of the spinal column and normal spine curvature contributed to significant errors in the initial estimate with the 2D method. Figure 4 shows these differences in one of the 10 cases. Because the axial slices do not necessary cut across the long axis of each vertebra, a simple measurement in one slice is inadequate for accurate assessment. In Fig. 4, the width of the vertebra appears to be 53.364 mm in the axial view (a) of this CT image, but the same linear segment in the sagittal view (b) clearly shows the misleading nature of the measurement, due to the oblique orientation of the vertebral body with respect to the entire patient. Figure 4 also shows the depth of a vertebra along the pedicle to be 66.4 mm in the axial view (c) of this CT image. The coronal view (d) of the same segment, however, shows a measurement that actually represents a corner to corner distance, which likely may not be the measurement intended. Both of these illustrations show the potential problems if screw length decisions are based solely on standard 2D axial, sagittal or coronal views. The potential for implant size and/or angulation error is increased further by the manual nature of current 2D spine surgery planning methods.

Another difference is the consistency of the planning report. Although there is a pre-defined manual entry form which is used routinely for spine planning, the use of this form is inconsistent from case to case. Figure 5 shows the plan from one of the cases. Because the process is manual, there are blank columns on the left and the scratched out numbers at the bottom. It is unclear if the surgeon neglected to fill in those angles or if the angles were 0° and therefore not entered. The values that were scratched out may possibly lead to transcription errors. In contrast, the

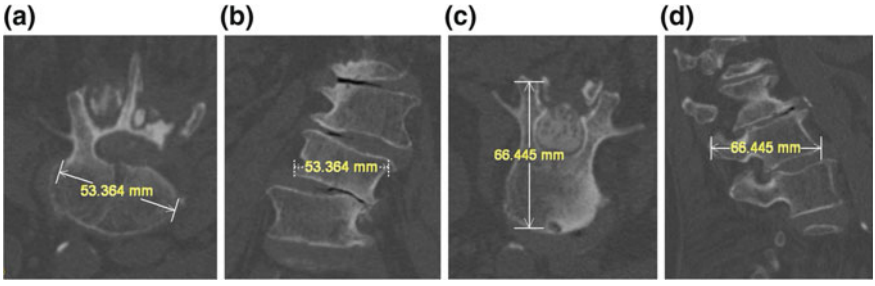


Fig. 4 The corresponding width and depth segments shown from different orientations illustrate the potential for misleading measurements when basing them on single 2D views

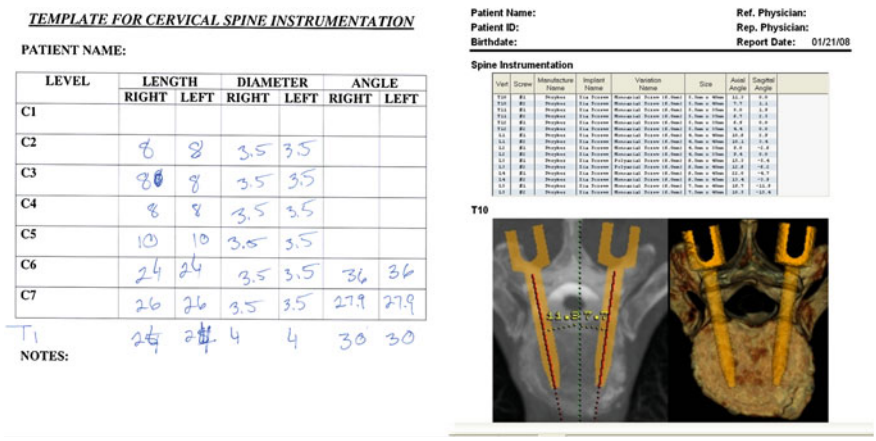


Fig. 5 The handwritten instrumentation report on the *left* illustrates the high potential for error when compared to the automatically generated report coming from the 3D templating tool

automatically generated report from the 3D planning process shown in Fig. 5 requires no manual intervention.

Each vertebral level includes an axial view with screw placement angles included and a top view 3D rendering, as well as two full spine renderings, illustrated in Fig. 6, with all screws displayed. The report can then be printed, saved, or added to the patient record as a DICOM object.

While some differences were identified in several comparisons, no strong trends were found in the pre-operative comparisons or the two pre-operative to post-operative comparisons. Given the small sample, this is not surprising. However, the different measurements generated, coupled with the participating surgeon's prevailing intuition that the 3D planning method was producing more consistent and more precise measurements, provides impetus for further study. It must also be recognized that comparison of pre-operative plans to post-operative results that were originally planned using 2D methods is not going to produce results that indicate one is "better" than the other. In fact, it could be expected that differences between 2D pre-operative planning and post-operative results should be small in these cases, since both were planned using the same method. It turns out that is exactly what occurred in the case of the experienced surgeon. The chart in Fig. 7 shows his 2D plans deviated less from the post-operative results than his 3D plans did. Interestingly, the less experienced resident's results were the opposite, as indicated by the same chart, which might indicate that 3D planning helps him plan cases that more resemble the expert. The chart in Fig. 7 also suggests that the resident's templates in the 2D plans tended to be shorter and narrower than those from the 3D plans. A plausible explanation for this is that residents tend to round down when using 2D methods since it is recognized as more of an estimate (due to oblique orientation of vertebra) and wanted to err on the side of safety. 3D planning gave a better view and provided more confidence in the template choice, thus

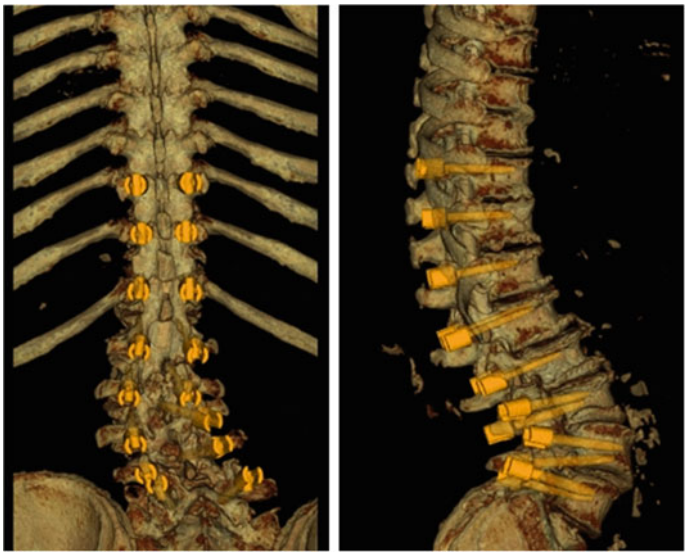


Fig. 6 The two renderings show the complete plan after all the individual vertebrae have been templated

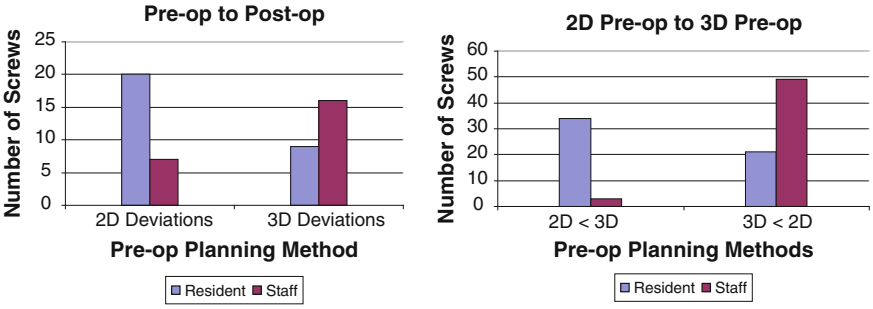


Fig. 7 The chart on the *left* shows the screw length deviations between the two different pre-operative planning methods and the post-operative results for the resident and staff surgeon. The chart on the *right* shows for each type of surgeon, the number of times the screw dimensions from one pre-operative planning method was smaller than those from the other pre-operative method

accounting for larger template sizes. The expert's templates were just the opposite, illustrated by the same chart. He tended to choose shorter narrower screws in the 3D plan compared to the same subjects in the 2D plan. Based on better views from the 3D plan, he may recognize preferred locations for the templates that would provide adequate holding strength while maximizing safety and avoiding problem areas not observed using 2D planning methods.

4.2 Assessment of Fastening Strength Estimates

We conducted an initial assessment of the developed platform using retrospective clinical data from patients who underwent spine surgery that involved the implantation of pedicle screws. For this study, we have limited our analysis to four lumbar cases, consisting of a total of 28 pedicle screws implanted in the lumbar spine of four patients. Procedures were performed by a staff surgeon following standard planning in the OR using 2D axial images from the pre-operative CT scans for implant size estimates, with no use of the virtual planning platform. All implants were deemed successful with no revisions. Post-operative CT scans routinely ordered for follow-up purposes were used as ground truth (clinical gold standard) for our assessment.

Several weeks after each procedure, a spine surgery fellow used the virtual planning platform to plan each procedure. The plans yielded a total of 30 virtual lumbar implants being suggested across all four cases. Although a total of 28 pedicle screws were implanted during the procedures, and the retrospective plans suggested 30 implants (i.e., one extra vertebral level suggested for instrumentation in one patient), only 26 implants were homologous—implanted/suggested at the same vertebral level in both procedures and plans, respectively. Therefore, for the sake of a consistent paired assessment, only the 26 homologous implants (same vertebral level in both plan and procedure) were considered in the analysis.

4.3 Implant Dimension Assessment

The dimensions of the implants selected using the virtual planning platform are automatically generated in the output report. However, to determine the dimensions of the screws implanted during the procedure, we used the virtual platform to "reverse plan" the post-operative CT such that the "reverse planned screws" match those visible in the post-operative CT images (Fig. 8). Following reverse planning, the implant dimensions were included in the automatically generated output report. This approach was chosen to eliminate any uncertainty associated with the user-conducted measurements caused by the partial volume effects and beam hardening artifacts induced by the presence of the metallic screws in the post-operative CT images.

4.3.1 Implant Fastening Strength Assessment

The Fastening Strength of the planned implants was directly assessed by estimating the intensity—area product across the pedicle volume displaced by the virtually implanted screws. To compute the Fastening Strength for the post-operative assessment, each "reverse planned" vertebra was registered to its homologous

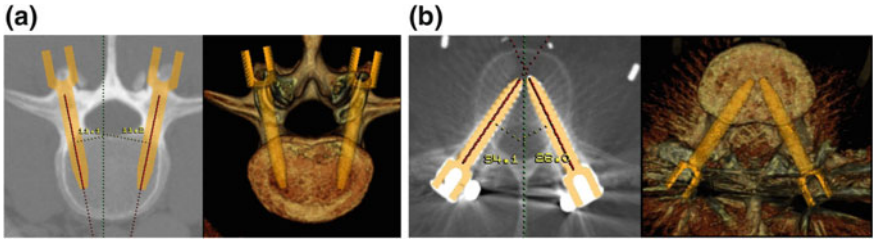


Fig. 8 Example of virtual templating and volume render representation at one vertebral segment. **a** Templating of pre-operative CT scan; **b** "reverse templating" of post-operative CT (high intensity metal implants are visible in the image)

counterpart in the pre-operative image (Fig. 9), as knowledge of the voxel intensity of the displaced bone segment is required. Given the rigid nature of individual vertebrae, an intensity-based rigid registration was used, followed by minimal manual adjustment. Following registration, the intensity—area product was computed for each pedicle bone segment displaced by the post-operative implant.

Table 1 summarizes the implant dimension (diameter and length), the Fastening Strength, computed displaced bone volume, as well as the mean voxel intensity across the displaced bone volume, for both the pre-operative planned and post-operative procedure. Since the assessment consist of the comparison of homologous implants in both the retrospective plans and procedures with respect to different parameters (implant dimensions, displaced bone volume, mean voxel intensity of

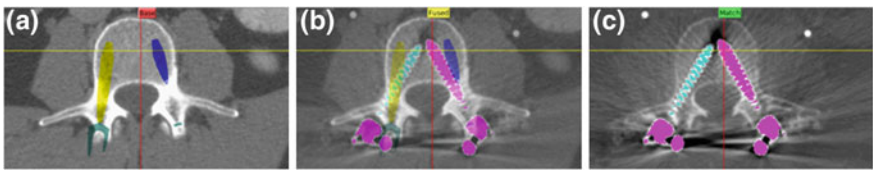


Fig. 9 Registration of post-operative "reverse plan" (**c**) to the pre-operative plan (**a**) for assessment of displaced bone volume to estimate Fastening Strength. The planned and post-operatively inserted pedicle screws can be seen in the fused image (**b**)

Table 1 Summary of planned versus post-operative measurements for implant diameter and length, as well as Fastening Strength, displaced bone volume, and mean voxel intensity of displaced bone volume

Measure (units)	Pre-operative plan	Post-procedure assessment
Screw diameter (mm)	5.5 ± 1.2	6.9 ± 0.8
Screw length (mm)	40.0 ± 2.0	47.1 ± 5.0
Fastening strength ($mm^3 \cdot HU$)	$3.6 \times 10^7 \pm 0.7 \times 10^7$	$4.7 \times 10^7 \pm 0.8 \times 10^7$
Displaced bone volume (mm^3)	$2.5 \times 10^3 \pm 0.4 \times 10^3$	$3.3 \times 10^3 \pm 0.5 \times 10^3$
Mean bone intensity (HU)	164 ± 65	167 ± 59

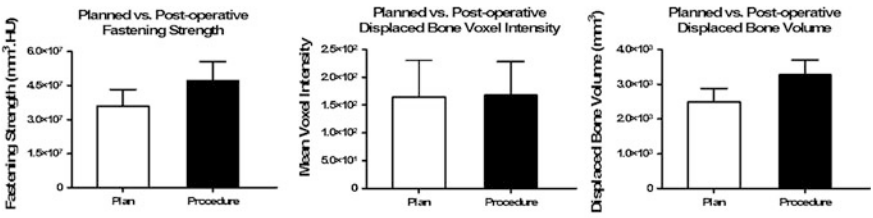


Fig. 10 Comparison between Fastening Strength, displaced bone volume and mean voxel intensity, showing no statistical difference ($p > 0.1$) between plan and procedure, therefore confirming consistency between the planned and clinical standard Fastening Strength

displaced bone volume, and Fastening Strength), we used the pair Student t-statistic to compare the paired (procedure vs. plan) results. While the average difference between the plan-suggested and procedurally-implanted screws was on the order of 1 mm in diameter and 5 mm in length (i.e., typically one screw size, given the dimensions of the available screws provided by the manufacturers), the paired Student t-statistic revealed no difference between the parameters estimated from the virtual plan and actual procedure ($p > 0.1$), therefore indicating that the results suggested by planning platform were in agreement with those assessed based on the post-procedure images, treated as clinical gold standard.

Figure 10 compares the Fastening Strength, displaced bone volume, and mean voxel intensity of displaced bone volume between the pre-operative plan and post-procedure outcome. While no statistical difference was noted between the plan and procedure, higher Fastening Strength correlated with larger displaced bone volume, and slightly larger implant dimensions.

The relationship used to estimate the Fastening Strength is available in mechanical engineering and machine component design literature [24] and traditionally used to determine the holding power of screws, bolts and other fasteners. Here we adapted to this concept the spine surgery application by relating the material strength to a surrogate measure of bone mineral density derived from CT image intensity. The Fastening Strength is computed based on the volume of bone displaced by the screw, assuming a cylindrical model whose diameter is measured across the thread, not just the screw shaft. To further emphasize the consistency and utility of the Fastening Strength, we further analyzed the observed Fastening Strength and displaced bone volume, in response to selected implant dimension. Figure 11 illustrates the correlation between Fastening Strength and implant dimension. As shown, as much as half of the holding power can be lost by undersizing the implant diameter by up to 3.5 mm, and as much as 35 % of the holding power can be lost by undersizing the implant length by up to 20 mm. These measurements are consistent with the displaced bone volume measurements, which are directly proportional to the implant dimension variability.

Moreover, as also revealed in Fig. 11, given similar implant dimensions (no difference in diameter or length), uncertainties on the order of 5 % were observed in the displaced bone volume measurements, which, in turn, led to 8 % differences in

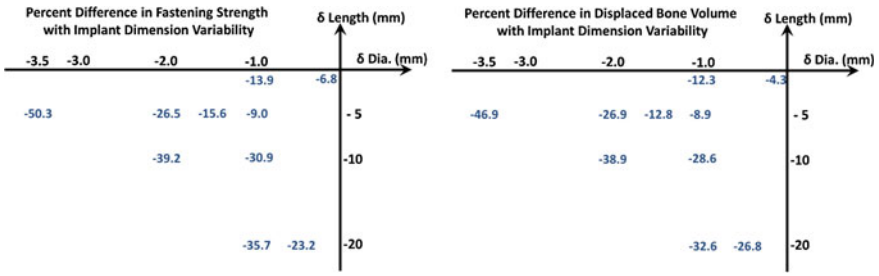


Fig. 11 Percent difference in Fastening Strength and displaced bone volume with variations in implant dimension. The negative values indicate loss of holding power and underestimated displaced bone volume due to undersized implants

fastening strength. While the difference in displaced bone volume measurements for identical size implants are mainly artificial, primarily due to partial volume effects inherent to the CT image resolution, the remaining differences in Fastening Strength may be real and mainly due to the difference in voxel intensity within the displaced bone volume.

With regards to the mean voxel intensity of the displaced bone volume, the plan and procedure both have comparable ranges, and consistent with image intensity range of the cancellous bone. Recall that cortical spine BMD was estimated as 192 ± 10 mg/cm^3, while cancellous bone BMD averaged ~ 140 mg/cm^3. Given the linear relationship between BMD and image intensity, cortical bone features a $\sim 235 \pm 14$ intensity range, while cancellous bone averages a mean voxel intensity of ~ 175. As shown in Table 1, the mean voxel intensity of the displaced bone volume in both the plan and procedure was on the order of 165 ± 60, which spans the mean voxel intensity of cancellous bone, but also extends into voxel intensity range associated with the cortical bone (~ 200). From a physical interpretation view point, the screw shaft is immersed into the cancellous bone while the "tip of the thread" extends into the cortical bone located toward the edge of the pedicle body (Fig. 12), therefore providing added strength that a thinner implant would not, since only spanning the cancellous bone region.

As documented [25, 26] and also suggested by the orthopedic surgery team, a pedicle screw implant is typically deemed optimal if the screw fully "taps" into the cancellous pedicle region and the edge of the screw thread "digs" into the hard cortical pedicle shell for improved holding power [27]. Based on the mean voxel intensity measurements of the displaced bone segment (in both the plans and post-operative assessments), the screws fully "tapped" into the cancellous bone, and also "grabbed" onto the cortical bone, documented by the upper tail of the voxel intensity range of the measured displaced bone volume (Fig. 12), confirming clinical requirement. Therefore, our proposed formulation of the Fastening Strength metric can provide quantitative information whether this clinical requirement has been met. This demonstrates that the plans, within their inherent limitations

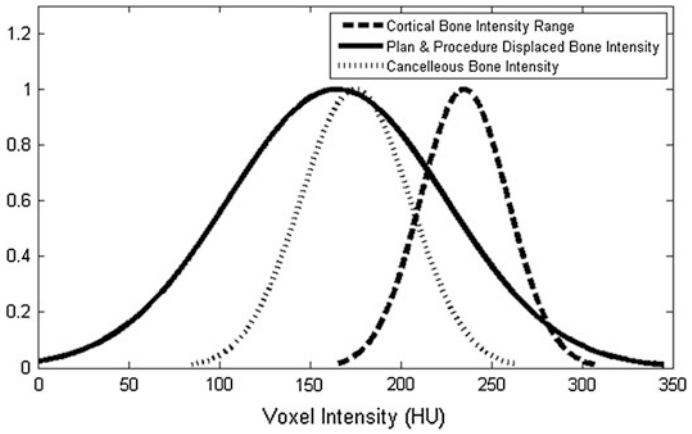


Fig. 12 Voxel intensity distribution for the cancellous and cortical bone according to the linear relationship between bone mineral density and image intensity. Superimposed voxel intensity range of the bone volume displaced by the implanted pedicel crews and computed according to the Fastening Strength formulation

introduced by the surgeon's versus fellow's skill level and screw measurement variability, provide similar Fastening Strength to the actual procedures and the proposed Fastening Strength correlates positively with implant dimension.

5 Current Clinical Experience and Relevant Cases

Given the perceived benefit of the 3D planning method, seven Mayo Clinic orthopedic surgeons have used the 3D templating tools to pre-operatively plan several cases.

Case Study 1: A 4 year old male presented with progressive congenital scoliosis associated with VATER syndrome and neurofibromatosis. The scoliosis was present at two levels. The patient had a complex cervicothoracic curve which progressed from 25° to 30° over one year. Hemivertebrae also caused thoracolumbar scoliosis which progressed from 25° to 35° over a one year period, with a focal kyphosis measure of 22°. For this patient, pre-operative CT scans were ordered to more precisely determine the pathologic anatomy and to permit 3D templating. The CT images clearly illustrated pathologic congenital anatomy of cervical and thorocolumbar congenital scoliosis, as shown in Fig. 13. 3D templating confirmed that the thorocolumbar pedicles could safely accommodate 3.5 and 4.0 diameter screws, also shown in Fig. 13, but the cervicothoracic vertebrae would not safely accommodate standard implants. The 3D templates allowed for straightforward generation of a physical model using a commercial 3D printer. This model was used intra-operatively, showing the precise starting points and trajectory for

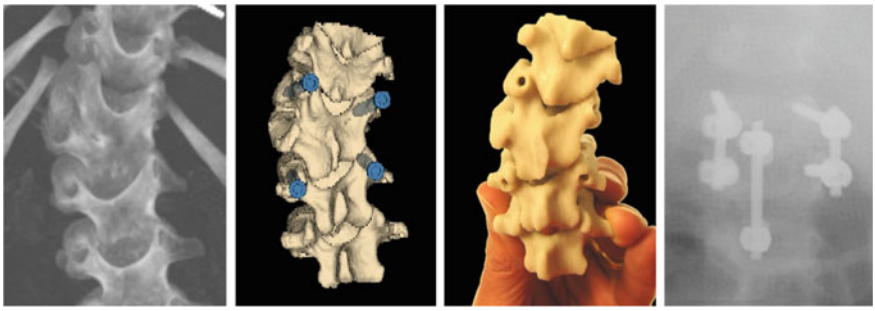


Fig. 13 Pre-operative thick maximum intensity projection CT of patient with thorocolumbar congenital scoliosis, rendering of 3D templating results developed with the spine surgery planning tools, the physical model printed from the plan, and the post-operative radiograph

each pedicle screw. Although the CT data was used to determine that hemivertebrae resection could be carried out with posterior pedicle screw and hook instrumentation from T12 to L1 shown in Fig. 4, it was not possible to model this process in the software. Instead the surgeon utilized the renderings and model to visually assess the resection process.

Case Study 2: A 4 year old presented with congenital scoliosis at the cervical-thoracic junction and at the thoracolumbar spine associated with a complex constellation of medical concerns such as dextrocardia, feeding difficulty, developmental delay and gastroesophageal reflux, as shown in Fig. 14. The congenital scoliosis at the thoracolumbar spine was caused by a fully segmented hemivertebrae. Curve progression from 25° to 35° over 2 years time warranted hemivertebrae resection. CT scan was indicated to evaluate the congenital vertebral anomalies and to measure the vertebral and pedicle dimensions. A 3-D rapid prototype model was generated from the pre-operative plan shown in Fig. 14 to confirm that the vertebrae

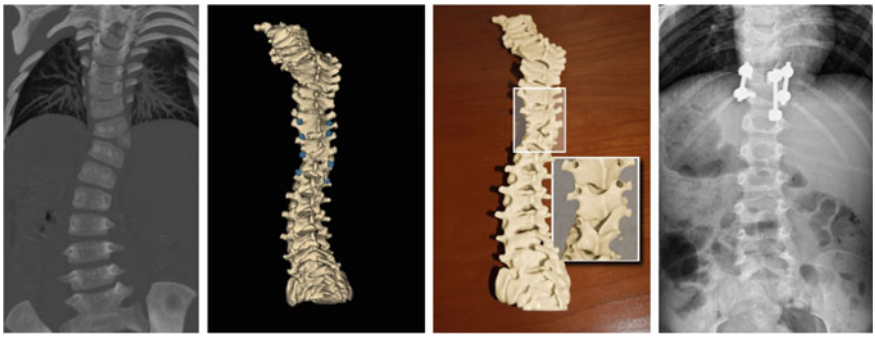


Fig. 14 Pre-operative thick maximum intensity projection CT of patient with cervical-thoracic and thoracolumbar congenital scoliosis (*upper-left*), rendering of 3D templating results developed with the spine surgery planning tools (*upper-right*), the physical model (*lower-left*) printed from the plan, and the post-operative radiograph (*lower-right*)

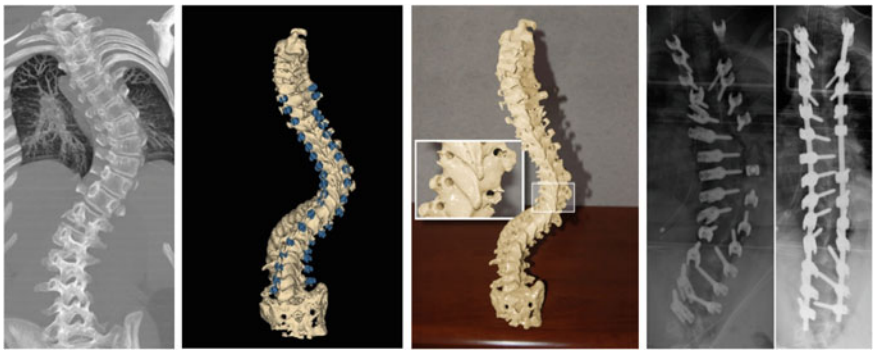


Fig. 15 Pre-operative thick maximum intensity projection CT of patient with severe progressive scoliosis, rendering of 3D templating results developed with the spine surgery planning tools, the physical model printed from the plan, and the inter-op, post-operative radiographs

adjacent to the congenital hemivertebrae could be safely instrumented with pedicle screws. The model in Fig. 14 illustrated correct starting point location and trajectory for pedicle screw placement, permitting safe instrumentation of the very small and abnormal spine. Surgery was performed without complication and an excellent clinical and radiographic outcome was achieved, as displayed in Fig. 14. As with Case 1, the viability of hemivertebrae resection was confirmed with CT; however, the software was unable to model this process. In addition, the surgeon planned for the placement of 4 additional screws located more distal from the hemivertebrae; however, it was determined intra-operatively that this was unnecessary.

Case Study 3: An 11 year old male presented with severe progressive scoliosis measuring 100° associated with Sprengel's Deformity, developmental delay, and several congenital vertebral abnormalities which can be seen in Fig. 15 (upper-left). The severe curve magnitude and congenital vertebral abnormalities warranted CT imaging to better understand the complex vertebral anatomy, especially pertaining to pedicle morphology. Three dimensional templating displayed in Fig. 15 (upper-right) confirmed that vertebral anatomy at most levels would safely accommodate pedicle screw fixation, and allowed identification of those vertebrae where pedicle screws could not be placed. Adding screw templates to the 3-D model permitted identification of pedicles which were "out of line" and could not be included in the instrumentation construct. The three dimensional model in Fig. 15 (lower-left) was used in the operating room and used as a reference for pedicle starting point and trajectory during the operation. Surgery was performed safely in an efficient manner with an excellent clinical and radiographic outcome, shown in Fig. 15 (lower-right). A total of 37 screws were planned for the surgery; however, the patient only required 26 screws to be placed, based on the pre-surgical plan and intra-operative assessment of the outcome.

Additional Cases: Four additional pediatric patients, all of which presented similarly complex deformities, have been operated on using the same planning methodology. One of the four presented severe thoracolumbar kyphotic deformity

and midthoracic level spina bifida, while the other three all presented severe progressive scoliosis. In each case, spinal instrumentation was required for fusion. Aided by the 3-D templating software, precise size and location of pedicle screws for safe implantation were identified, and full-sized models were generated from the templated image data (Fig. 16). The models were used in the operating room as a reference in all four cases, helping the surgeon to easily transfer the precise plan directly to the procedure. Clinical and radiographic outcomes were considered excellent in each of the four cases and each surgery was performed without complication.

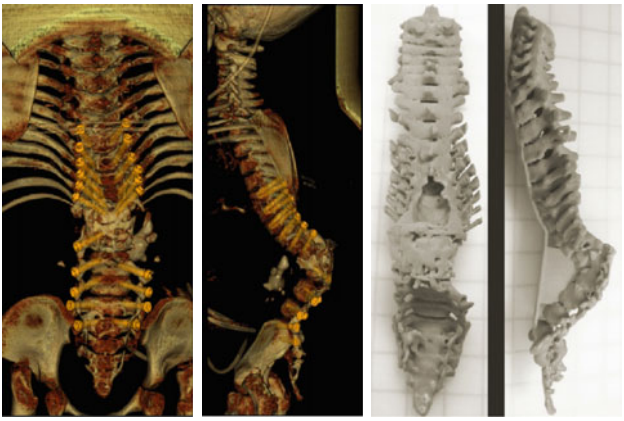
6 Discussion

This work demonstrates that accurate 3D pre-surgical planning for complex pedicle screw placement in spinal deformity correction interventions can be achieved using pre-acquired high resolution volumetric images and user-interactive guidance. Moreover, the virtual pre-surgical plan can be used to rapidly prototype a 3D physical anatomical model of the patient's spine in its corrected configuration. Surgeons reported that the planning software and 3D models provided significant information which increased the surgeon's ability to plan several concurrent surgical approaches, and, therefore, consider several viable options in the procedure room.

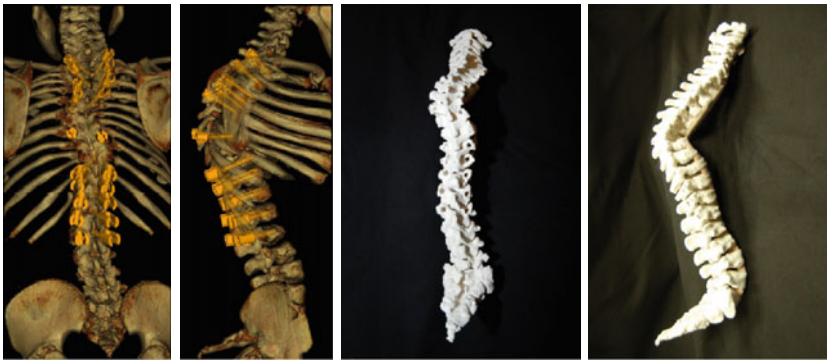
Due to the broad variety of implants used in this procedure (including hooks and rods), future work will include the incorporation of different types of instrumentation. Moreover, we envision to further improve the capabilities of the currently developed platform to enable a better integration of the planning module with the intra-operative guidance. To date, the SSP application is solely used to plan the procedure and the resulting data is available for visualization, either virtually or physically, in the operating room. However, we believe that the current workflow can be further enhanced by providing the surgeon with a direct spatial relationship that enables the translation and implementation of the pre-operative plan into surgical guidance. One approach is to perform a virtual model-to-patient registration using a surgical instrument localization system, and to make use of the tracked instruments to better guide the instrumentation, ensuring that the insertion point and trajectory of the pedicle screw follow the procedure plan.

Several other techniques have been explored by different groups in parallel with the development of the Spine Surgery Planning platform. As mentioned in the introduction, the initial iteration of our spine surgery planning platform was developed simultaneously with the iPlan platform developed simultaneously by BrainLab [11] and features similar capabilities, including selection of virtual implants, trajectory planning, as well image segmentation, registration and overlay of multiple datasets into the templating workflow.

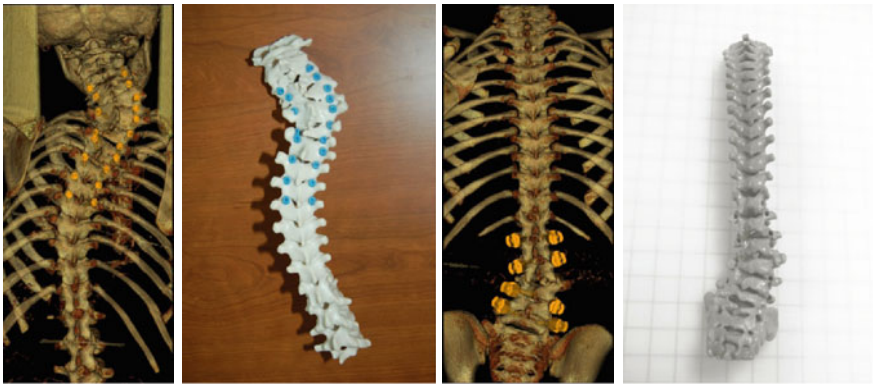
Bichlmeier et al. [1] disseminated on a new method for navigated spine surgery using a stereoscopic video see-through head-mounted display (HMD) and an



Patient 4



Patient 5



Patient 6 Patient 7

Fig. 16 Renderings of 3D templating results (*left* for each patient set) developed with the spine surgery planning tools for each of four additional cases (Patient 4–7 as labeled above), and the physical models (*right* for each patient set) printed from the plan

optical tracking system. Vertebrae are segmented from volumetric CT data and visualized in situ. A surgical drilling device was virtually extended with a mirror for intuitive planning of the drill canal, control of drill direction and insertion depth. The system was evaluated using realistic replica of lumbar vertebrae against the classical, monitor-based navigation system providing three orthogonal slice views on the operation site. Outcome was assessed according to the procedure time and accuracy measurements recorded based on the post-procedure CT images of the drilled vertebral models.

Given the fastening strength estimates, we believe it is not critical to employ a complex finite element model under hypothetical loading conditions, especially given the current formulation suits any loading condition and provides a surrogate measure for the screw holding power. This approach provides consistent trends with implant dimension within the limits defined by the clinical standard implantation procedures conducted with no computer assistance, chosen as gold standard for assessment of the plans. Claims that the proposed planning approach would lead to superior outcome compared to the clinical standard would be very difficult to make, as they would invalidate the quality of health care currently delivered in the clinic via the standard axial CT image-based planning approach. However, we claim that the proposed approach provides objective measures for planning (i.e., the effect of implant dimension and trajectory combined into the Fastening Strength metric), and enables planning prior to the procedure and, if desirable, outside the OR, potentially leading to shorter procedure and anesthesia time.

The virtual templating and planning can be completed any time before the procedure, once the patient CT image dataset is available. Following data importing into the surgery planning platform, the actual planning task requires 30–60 min of effort (5–10 min for each vertebral level), depending on the complexity of the case. The planning output consists of an automatically generated report that lists each instrumented vertebral level, selected implant dimensions and trajectory, accompanied by orthogonal views and 3D volume rendered representations. Following the addition of the Fastening Strength feature to the platform, measures such as displaced bone volume, mean voxel intensity and Fastening Strength can also be made available in the report; however their knowledge is more beneficial during the actual planning process, when selecting different size implants and assessing their optimal trajectories to optimize holding power. Lastly, if a full scale physical spine model is required, additional processing time (1 h) is necessary to generate a stereolithography (STL) file containing the surface model information; the 3D printing process may require up to 24 h to complete, but once set up, the printing is mostly automated.

One limitation of the current study, besides the small sample size available for analysis, is the comparison of the virtual planning outcome to traditionally conducted procedures, therefore making it difficult to account for any potential deviations from the plan that could have occurred during the intervention. In addition, the retrospective virtual planning was performed by a (less experienced) fellow, while the actual procedures were performed by a staff surgeon, which explains the fellow's more conservative implant selection during planning—slightly thinner and

shorter screws to avoid pedicle rupture or protrusion outside the vertebral body. To address these limitations, we intend to conduct a double cohort study that enables the translation of the plan into the OR for appropriate comparison of both virtual and traditional plans followed all the way to post-procedure outcome. This study will provide a larger sample size for analysis and enable us to investigate any differences between experienced and novice surgeons as far as implant selection and planning.

For the sake of a consistent comparison of homologous implants in both the retrospective plan and procedure with respect to different parameters (implant size, displaced bone volume, mean voxel intensity of displaced bone volume and lastly the fastening strength), we used the paired Student t-statistic to compare the paired results. However, we also computed the correlation between the retrospective plan and post-operative outcome for each of the five parameters mentioned above. A moderate correlation was found between the retrospective plan and post-procedure outcome for the implant diameter, implant length and displaced bone volume (0.63, 0.59 and 0.52, respectively), and a higher correlation was revealed between the plan and procedure with respect to the Fastening Strength and Mean Voxel Intensity of the displaced bone volume (0.82 and 0.79, respectively). These correlations confirm the more conservative plans performed by the fellow (selection of slightly thinner and shorter implants) compared to the implants used during the actual procedures by the staff surgeon. Nevertheless, the trajectory of the implants and their positioning within the vertebral body was consistent between the retrospective plans conducted by the fellow and the procedures performed by the surgeon.

Clinical Limitations and Impact: We recognize ongoing efforts in computer-assisted spine surgery to assist the surgeon with implant positioning during the procedure, and we believe that such endeavours, although valuable in the intra-operative setting, make limited effort to improve procedure planning and eventually enabling the planning process be conducted out of the OR. The platform described here is intended to complement the intra-operative endeavours and enable planning to be performed once a CT scan is available, outside of the OR, to reduce anaesthesia time, and overall procedure time and costs. Hence it is not counter-intuitive to conclude that by performing the planning prior to the procedure, the time under anesthesia, overall OR time and all associated costs could be reduced. Moreover, given the planning platform utilizes typical standard of care pre-operative image datasets, it can be seamlessly integrated with both manual implantation procedures, as well as computer-assisted navigation, providing the added bonus of precisely translating the planned trajectories from the pre-operative planning stage to the intra-operative stage by means of plan-to-patient registration and tracking of the surgical instruments and implants.

7 Summary and Future Work

In this paper we have described the augmentation of our existing spine surgery planning platform with a new metric—the Fastening Strength—that provides a surrogate measure for the screw holding power. The proposed metric is evaluated in conjunction with the virtual implant selection and 3D templating of the patient-specific 3D CT dataset and enables optimal selection and trajectory planning by taking into consideration the effect of implant dimension and geometric path, as well as the strength (i.e., bone mineral density) of the bone substrate, which is critical to achieve improved fixation. The conducted experiments showed agreement between the implants suggested by the retrospective plans and the actual procedure outcome.

In addition to the enhanced planning platform led to similar decisions as far as implant dimension selection and trajectory planning, it also provides the surgeons with the ability and choice to perform the planning pre-operatively, outside of the OR and rely on objective measures for safe and secure implant positioning that combine both implant dimension, trajectory and strength of the bone substrate into a single metric.

Future directions will involve further evaluation via both retro- and prospective studies, as well as the integration of virtual planning platform with existing computer-assisted navigation platforms [28]. In addition, we plan to compare the screw fastening strength outcome yielded by our surrogate formulation to that predicted by classical finite element models under different loading conditions, and demonstrate the feasibility of employing the proposed Fastening Strength formulation as a validated predictor of the holding power of the implanted pedicle screws.

Acknowledgments The authors would like to thank all members of the Biomedical Imaging Resource who have helped with the development and implementation of this project, especially Alan Larson, Bruce Cameron, Phillip Edwards, and Dennis Hanson. Also, we would like to acknowledge our clinical collaborators for their continuous support: Dr. Paul M. Huddleston, Dr. Jonathan Morris, Dr. Jane Matsumoto, and Dr. Shyam Shridharani.

References

1. Bichlmeier C, Heining SM, Rustaee M, Navab N (2007) Virtually extended surgical drilling device: virtual mirror for navigated spine surgery. In: Ayache N, Ourselin S, Maeder A (eds) Medical image computing and computer-assisted intervention, vol 4791. Lecture notes in computer science, pp 434–441
2. Sukovich W, Brink-Danan S, Hardenbrook M (2006) Miniature robotic guidance for pedicle screw placement in posterior spinal fusion: early clinical experience with the SpineAssist. Int J Med Robot 2:114–122
3. Weinstein J, Rydevik B, Rauschning W (1992) Anatomic and technical considerations of pedicle screw fixation. Clin Orthop Relat Res 284:34–46
4. Cohen M, Wall E, Brown R, Rydevik B, Garfin S (1990) Cauda equina anatomy II: extrathecal nerve roots and dorsal root ganglia. Spine 15:1248–1251

5. Rauschning W (1983) Computed tomography and cryomicrotomy of lumbar spine specimens: a new technique for multi-planar anatomic correlation. Spine 8:170–180
6. Rauschning W (1987) Normal and pathologic anatomy of the lumbar root canals. Spine 12:1008–1019
7. Rydevik B, Brown M, Lundborg G (1984) Pathoanatomy and pathophysiology of nerve root compression. Spine 9:7–15
8. Wall E, Cohen M, Massie J, Rydevik B, Garfin S (1990) Cauda equine anatomy I: intrathecal nerve root organization. Spine 15:1244–1247
9. Cleary K, Clifford M, Stoianovici D, Freedman M, Mun SK, Watson V (2002) Technology improvements for image-guided and minimally invasive spine procedures. IEEE Trans Inf Technol Biomed 6:249–261
10. Schafer S, Nithiananthan S, Mirota DJ, Uneri A, Stayman JW, Zbijewski W, Schmidgunst C, Kleinszig G, Khanna AJ, Siewerdsen JH (2011) Mobile C-arm cone-beam CT for guidance of spine surgery: image quality, radiation dose, and integration with interventional guidance. Med Phys 38:45–63
11. Bamba Y, Nonaka M, Nakajima S, Yamasaki M (2011) Three-dimensional reconstructed computed tomography-magnetic resonance fusion image-based preoperative planning for surgical procedures for spinal lipoma or tethered spinal cord after myelomeningocele repair. Neurol Med Chir (Tokyo) 51:397–402
12. Augustine K, Holmes D, Hanson D, Robb R (2006) Comprehensive, powerful, efficient, intuitive: a new software framework for clinical imaging applications. In: Proceedings of SPIE —medical imaging 2006, 11–16, San Diego, CA
13. Hanson D, Robb R et al (1997) New software toolkits for comprehensive visualization and analysis of three-dimensional multimodal biomedical images. J Digit Imaging 10(2):1–2
14. Augustine K, Holmes III DR, Hanson D, Robb RA (2006) Comprehensive, powerful, efficient, intuitive: a new software framework for clinical imaging applications, vol 6141. In: Proceedings of SPIE medical imaging, 61410 N-10
15. Homolka P, Gahleitner A, Prokop M, Nowotny R (2002) Bone mineral density measurement with dental quantitative CT prior to dental implant placement in cadaver mandibles: pilot study. Radiology 224:247–252
16. Jiang Y, Zhao J, Augat P, Ouyang X, Lu Y, Majumdar S, Genant HK (1998) Trabecular bone mineral and calculated structure of human bone specimens scanned by peripheral quantitative computed tomography: relation to biomechanical properties. Biomed Eng Online 13:1783–1790
17. Robb RA, Barillot C (1989) Interactive display and analysis of 3-D medical images. IEEE Trans Med Imaging 8(3):217–226
18. Robb RA, Hanson DP (1990) ANALYZE: a software system for biomedical image analysis. In: Proceedings of the first conference on visualization in biomedical computing, May 22–25, pp 507–518
19. Robb RA, Hanson DP (1995) The ANALYZE^{TM} software system for visualization and analysis in surgery simulation. In: Lavallée S, Taylor R, Burdea G, Mösges R (eds) Computer integrated surgery. MIT Press, Cambridge, pp 175–190
20. Hanson DP, Robb RA, Aharon S, Augustine KE, Cameron BM, Camp JJ, Karwoski RA, Larson AG, Stacy MC, Workman EL (1997) New software toolkits for comprehensive visualization and analysis of three-dimensional multimodal biomedical images. J. Digital Imaging 10(2):1–2
21. Cameron BM, Manduca A, Robb RA (1996) Patient specific anatomic models: geometric surface generation from 3 dimensional medical images using a specified polygonal budget. In: Sieburg H, Weghorst S, Morgan K (eds) Health care in the information age. IOS Press and Ohmsha, Washington, pp 447–460
22. Robb RA, Cameron BM, Aharon S (1997) Efficient shape-based algorithms for modeling patient specific anatomy from 3D medical images: applications in virtual endoscopy and surgery. In: Proceedings of shape modeling and applications, pp 97–108

23. Lin W, Robb RA (1999) Dynamic volume texture mapping and model deformation for visually realistic surgical simulation. Proc Med Meets Virtual Reality 62:198–204
24. Juvinall RC, Marchek KM (2011) Fundamentals of machine component design, 5th edn. Wiley, New York
25. Vaccaro AR, Rizzolo SJ, Allardyce TJ, Ramsey M, Salvo J, Balderston RA, Cotler JM (1995) Placement of pedicle screws in the thoracic spine. Part I: morphometric analysis of the thoracic vertebrae. J Bone Joint Surg Am 77:1193–1199
26. Misenhimer GR, Peek RD, Wiltse LL, Rothman SL, Widell EH Jr (1989) Anatomic 450 analysis of pedicle cortical and cancellous diameter as related to screw size. Spine 14:367–372
27. Halvorson TL, Kelley LA, Thomas KA, Whitecloud TS III, Cook SD (1976) Effects of bone mineral density on pedicle screw fixation. Spine 19:2415–2420
28. Ortmaier T, Weiss H, Döbele S, Schreiber U (2006) Experiments on robot-assisted navigated drilling and milling of bones for pedicle screw placement. Int J Med Robot. 2:350–363

Tracked Ultrasound in Navigated Spine Interventions

Tamas Ungi, Andras Lasso and Gabor Fichtinger

Abstract Ultrasound is an increasingly popular imaging modality in image-guided interventions, due to its safety, accessibility, and low cost. But ultrasound imaging has a steep learning curve, and requires significant coordination skills from the operator. It is difficult to interpret cross-sectional anatomy in arbitrary angles, and even more challenging to orient a needle with respect to the ultrasound plane. Position tracking technology is a promising augmentation method to ultrasound imaging. Both the ultrasound transducer and the needle can be tracked, enabling computer-assisted navigation applications in ultrasound-guided spinal interventions. Furthermore, the patient can also be tracked, which enables fusion of other imaging modalities with ultrasound. In this chapter, we first present the technical background of tracked ultrasound. We will review how to build research systems from commercially available components and open-source software. Then we will review some spine-related applications of tracked ultrasound modality, including procedural skills training, needle navigation for anesthesia, surgical navigation, and other potential applications.

1 Introduction

Ultrasound is becoming a ubiquitous imaging tool in many medical specialties due to its safety, portability, and low cost. Recent ultrasound devices fit in the physicians' pockets, and instantly provide real-time images of almost all anatomical regions without radiation risks to the patient or physician. Spine is, however, one of the particularly difficult areas for visualization with ultrasound. Bones and ligaments are close to the skin, and they cast acoustic shadows by reflecting the majority of the ultrasound waves, not letting through enough for visualization of deeper anatomical structures. Furthermore, stiff tissue layers of spine muscles

T. Ungi (✉) · A. Lasso · G. Fichtinger
Queen's University, Kingston, Canada
e-mail: ungi@queensu.ca

© Springer International Publishing Switzerland 2015

S. Li and J. Yao (eds.), *Spinal Imaging and Image Analysis*,
Lecture Notes in Computational Vision and Biomechanics 18,
DOI 10.1007/978-3-319-12508-4_15

attenuate the energy of the ultrasound more than other tissues with more water content. One can still find sonographic landmarks along the spine that can be used to obtain limited view of the anatomy. These landmarks are often used during interventions, as the operator finds the way of the needle based on these points.

Ultrasound combined with position tracking is a promising technology that has recently reached the clinical device market. It allows needle navigation methods that show the 3D position or projection of the tracked needle relative to the tracked ultrasound image. This visual aid enhances the accuracy of needle insertions when the target is directly visible on ultrasound. Some commercial ultrasound machines recently offer fusion of CT or MR images to real-time ultrasound, which is also a very promising avenue in computer assisted spine interventions. The real-time nature of ultrasound combined with the resolution and contrast of other image modalities may revolutionize image-guided spine interventions, enabling more procedures to be performed in a minimally invasive way. In this chapter, we will focus on the tracked ultrasound technology, and show some of its promising applications that may become routine procedures in the hands of surgeons, anesthesiologists, or interventional radiologists.

2 Ultrasound in Spinal Needle Guidance

Ultrasound has been in use for decades in guidance of invasive procedures in the spine. Although most needle insertion procedures that are commonly performed, can be completed blindly with knowledge of the anatomy. The procedural difficulty of spine interventions has a wide range depending on target structures and individual patients. For example, the most common procedure is lumbar puncture, needle insertion into the spinal canal between two lumbar vertebrae. Lumbar puncture is generally thought of as a simple procedure that every physician is able to perform without image guidance or other forms needle guides. However, in obese patients or degenerative spines, even this procedure can be so difficult that it requires ultrasound or fluoroscopic guidance. There are significantly more difficult procedures, such as selective nerve blocks, that are only attempted using CT or MRI guidance.

The most common use of spine ultrasound is to find vertebral interspaces for lumbar puncture in difficult cases. Ultrasound is helpful when the spine is covered by thick fat tissue, or when spine pathologies prevent conventional navigation by palpation. In these cases, ultrasound scanning can be done either before needle insertion, or during needle insertion to provide real time guidance as the needle approaches its target. The first technique uses landmarking. Ultrasound is used before needle insertion to find the space between two spinous processes, and marking it with a pen on the patient's skin. The needle is introduced at the marked point, which has a high probability of leading to the space between two vertebrae. In case of the second technique, imaging can be performed simultaneously during needle insertion too, to provide real time visual feedback on the needle position.

Real time guidance requires more experience and coordination skills, because the two hands of the operator are engaged in different tasks, and the attention is divided between image interpretation and needle manipulation. The difficulty in learning this complex skill is probably the only disadvantage of ultrasound-guided needle insertions in the spine region [1].

3 Tracked Ultrasound Systems

Although ultrasound has proven to be a great help in needle insertions, the combination of ultrasound imaging and position tracking, called tracked ultrasound, offers as many opportunities in the hands of interventionists as a new imaging modality. Tracked ultrasound systems have just reached the clinical market, and their future role in clinical practice will be subject to how much evidence will be found on its benefits. But the future looks promising for tracked ultrasound. It is one of the most affordable imaging modalities, and prices will drop with future generations of devices. It helps spatial coordination of the needle relative to ultrasound image position, which is one of the most challenging skills in medical interventions; therefore probably many operators will take advantage of this technology. Tracked ultrasound systems are relatively easy to build in research laboratories, and are exciting tools in experimental and clinical research. Therefore, we dedicate this section to the technical details of tracked ultrasound systems, with the goal of making them easily reproducible for a wide audience. We focus on the adaptability to existing ultrasound and tracking devices, rather than recommending a single set of hardware components. We encourage every reader who has access to an ultrasound machine and a position tracker to try assembling tracked ultrasound, because most medical specialties can take advantage of such an enhancement of ultrasound imaging in the guidance of interventions.

4 Position Tracking in Ultrasound-Guidance

Position tracking technologies evolved rapidly in the past decades, and have made it possible to track the ultrasound transducer, as well as the needle during interventions. This allowed development of navigation software for needle guidance. Medical navigation applications are much like GPS navigators developed for cars. They take advantage of position tracking by showing the user where they are on a geographical map. This makes the map extremely easy and intuitive to use. Medical navigation software enhances traditional medical images and image-guided interventions by showing the real-time positions of medical instruments on these images. Although the medical interventionist community is more careful accepting new technologies than car drivers.

Table 1 Summary of advantages and disadvantages of optical and electromagnetic tracking technologies

	Optical tracking	Electromagnetic tracking
Advantages	• Accuracy ~1–0.1 mm • Does not depend on objects in its environment • Large range (several meters) • Wireless position markers	• Can track without line of sight (inside body) • Position sensors can be small to fit in needles and catheters (~0.5 mm)
Disadvantages	• Requires line of sight • Optical markers are relatively large	• Accuracy ~1–2 mm • Limited range (typically 20–60 cm) • Affected by ferromagnetic metals in its environment • Wired position sensors

Common position tracking devices in medicine are using either optical or electromagnetic technology (Table 1). Optical tracking uses cameras and optical position markers that the computer automatically detects on the camera images. The main advantages of optical tracking are its accuracy and robustness. The main disadvantage is that the position markers need to be relatively large and in the line of sight of the cameras. An emerging alternative to optical tracking is electromagnetic tracking technology that uses an electromagnetic field generator, and wired position sensors that detect their position relative the field generator. Electromagnetic trackers generate a known changing magnetic field, and measure the currents in sensor coils that are induced by the changing magnetic field. A signature of currents in the sensor is unique to its position relative to the field generator. The main advantage of electromagnetic tracking is that it does not require line of sight, although it is less accurate and is sensitive to certain metal objects, especially electric devices in its environment.

Tracking the ultrasound transducer expands the possibilities in ultrasound-guided needle interventions. By attaching a position tracker to the ultrasound transducer and the needle, their relative positions can be computed and visualized, even when the needle is not in the ultrasound imaging plane. Such a tracked system can be further enhanced by attaching another position sensor to the patient. This allows visualization of the needle not only relative to the ultrasound image, but relative to pre-procedural CT, MRI, or other models of patient anatomy.

There are other technologies for needle tracking in ultrasound-guided interventions beyond optical and electromagnetic. The most simple and oldest way is mechanical tracking is to attach a passive needle guide to the ultrasound transducer. Ultrasound guidance methods for abdominal interventions use mechanical needle guides, but they constrain the needle motion to a single line relative to the ultrasound imaging plane. This line is displayed on the ultrasound display, so the operators see where the needle will be inserted relative to the image. The needle target can be chosen by moving the transducer with the fixed needle guide. But in the spine, the target areas are only visible from a limited range of angles. And the

needle usually has to go through a narrow space. Therefore, spinal interventions require more freedom of motion of both the transducer and the needle, so mechanical needle guides are typically not suitable for these procedures. Optical and electromagnetic position tracking, however, allows any position and angle of the needle relative to the ultrasound transducer. Using the tracked position information, navigation software can display the needle relative to the ultrasound image in real time.

5 Hardware Components

Experimental tracked ultrasound systems have been studied for over a decade in spinal needle guidance applications. But the first products approved for clinical use only appeared recently on the market. In this section we describe the architecture of tracked ultrasound systems in general, and how research prototypes can be built from low-cost components.

Tracked ultrasound hardware systems are composed of a conventional ultrasound machine and an added position tracker. In an experimental setting, there is often a dedicated computer for tracked ultrasound data processing, because ultrasound machines either restrict installation of research software or their hardware is not powerful enough for running additional applications. We will discuss a system design with a dedicated computer for our research application, because it can be easily built from existing components in any research laboratory (Fig. 1).

The majority of tracked ultrasound systems use electromagnetic technology for position tracking. Although optical tracking can also be used, the line of sight often breaks when the transducer is moved around the patient. This causes loss of

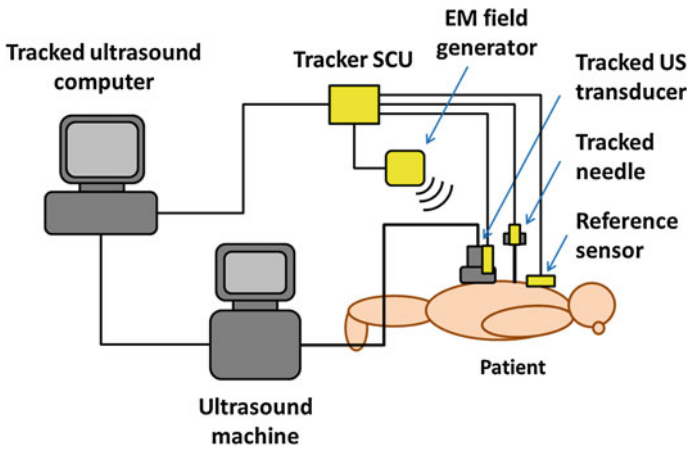


Fig. 1 Schematic layout of tracked ultrasound systems using electromagnetic (*EM*) position trackers

tracking signal, which is inconvenient for the operating staff. Electromagnetic trackers do not need line of sight, and—if the field generator can be placed close enough to the operating region—it is usually accurate enough.

When choosing an ultrasound machine for a tracked ultrasound system, we should first consider systems that are already integrated with position tracking, and have research interface that provides real-time access to the ultrasound image and tracking data streams. If tracking is not already available in the chosen ultrasound machine, an external tracker needs to be attached to the transducer. Even if the ultrasound machine does not offer digital access to the images and imaging parameters, most ultrasound machines have a standard video output that can be tapped into using a video grabber device.

Fixing the tracking sensor on the ultrasound transducer is not difficult using glue or a rigid clip. If sterile environment is needed, the transducer along with the sensor can be placed in a sterile bag. The reference position sensor needs to be fixed to the patient as rigidly as possible. Since the reference sensor provides the link between the patient and the navigation coordinate system, it makes the system more convenient to use if anatomical directions are marked on the reference sensor, so it can be placed in the same orientation. A reference sensor holder can provide the anatomical markers, along with an interface that can be firmly attached to the skin using an adhesive sheet (Fig. 2). Tracking the needle is the most challenging task, especially if the needle is thin (smaller than about 17 Ga) and bends during insertion. A larger, more accurate sensor can be clipped to the needle using a disposable plastic interface. But when the needle bends, a clipped sensor at the hub will not give accurate information on the tip position. Smaller sensors can be integrated in the needle stylet to provide direct tip tracking. Some companies offer electromagnetically tracked stylets approved for clinical use. However, such small sensors have a very limited (around 200 mm) usable range around the field generator, which can make the system hard to set up around the patient.

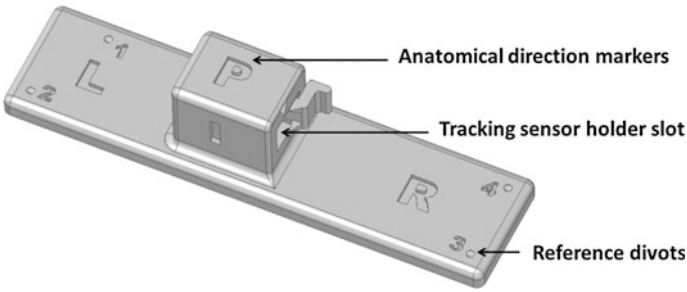


Fig. 2 Reference sensor holder

6 System Calibration

Ultrasound imaging differs significantly from other imaging modalities traditionally used in image-guided interventions. Both the contents and the positions of ultrasound images change rapidly in time, while CT and MRI images have static content and well-defined positions. Therefore ultrasound tracking requires special practices to ensure a maintainable navigation software design. We describe the coordinate systems that need to be represented in tracked ultrasound systems, and best practices in finding the transformations between the coordinate systems. In other words, we discuss calibration between components of the system.

In a full featured navigation system, there are three dynamic and three static coordinate transformations (Fig. 3). The dynamic transformations are shown in orange color, and the static ones in blue. The dynamic transformations change rapidly as the tracking sensors move relative to the Tracker coordinate system. The Tracker coordinate system is most commonly the electromagnetic field generator. The static transformations are equally important, but they do not change significantly during the intervention.

All transformation chains eventually end in a common Right-Anterior-Superior (*RAS*) anatomical coordinate system. When a CT or MRI image is loaded in the needle navigation scene, their *RAS* coordinate system is used. In ultrasound-only cases, the *RAS* can be defined at an arbitrary position with the coordinate axes directions matching the patient anatomical directions.

Spatial calibration of the system entails the computation of the static transformations. *Reference* to *RAS* transform is typically obtained by landmark registration. In this method the transform is determined by minimizing the difference between points defined in the pre-procedural CT or MRI image and the same points marked

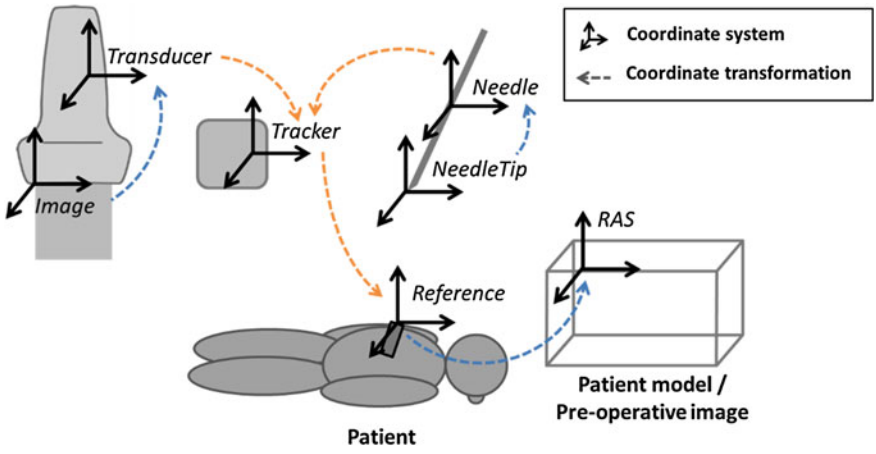


Fig. 3 Coordinate systems and transformations in a tracked ultrasound-guided needle navigation scene

on tracked ultrasound images. The method is very simple, the computation is immediate, and usually accurate enough, but finding the corresponding anatomical locations on different imaging modalities requires experience. Although there have been promising attempts to automate this process by image-based registration. Automatic methods may require less skills from the users and might be more accurate (by matching large number of points or surface patches), but so far these methods do not seem to be able to match the speed, simplicity, and robustness of the manual registration method.

Computation of the *NeedleTip* to *Needle* transform is straightforward, typically performed using a simple pivot calibration. The tracked needle is pivoted around its tip for a couple of seconds and the transform that minimizes the dislocation of the needle tip is computed. Usually the calibration has to be performed only once for each needle type that may be used in the procedure.

Determining the *Image* to *Transducer* transform (also known as *probe calibration*) accurately is a difficult task, mostly because of the 3D point localization by ultrasound is inherently inaccurate, due to the "thickness" of the ultrasound beam (Fig. 4). Beam width causes objects to appear in the ultrasound image that are several millimeters away from the ideal imaging plane and blurring of object boundaries on the images.

The *Image* to *Transducer* transform can be determined by moving a tracked pointing device (such as a needle or stylus) to various points in the image and recording the pointer tip position in both the *Transducer* coordinate system and the *Image* coordinate system (Fig. 5). The transform can be determined by a simple landmark registration. The advantages of the method are that it is simple, reliable, requires just an additional tracked stylus, and can be performed in any medium where a needle can be inserted. However, positioning the pointing device's tip in the middle of the image plane and finding the tip position in the image requires an experienced operator and therefore the accuracy and speed of the calibration heavily depends on the operator.

Automatic methods have been proposed to reduce the operator-dependency and increase the accuracy of the probe calibration. These methods extract features (such

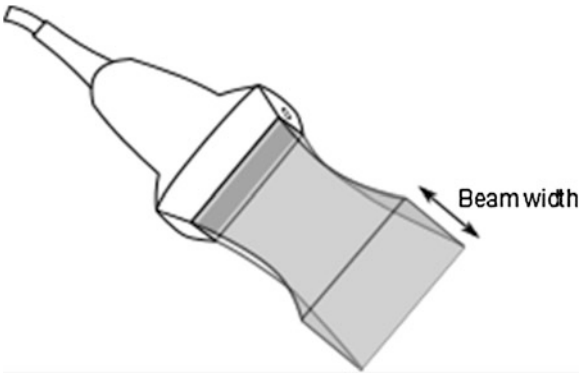


Fig. 4 Anything inside the thick ultrasound beam will appear in the acquired ultrasound image

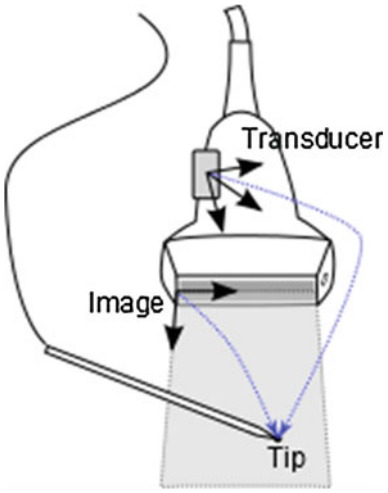


Fig. 5 Spatial calibration of the transducer can be performed by recording the pointer tip position in the *Transducer* coordinate system and marking them in the image

as intersection points or lines) from the image automatically, then compute the transform that minimizes the difference between the expected and the measured positions of the features.

Intersection of a thin linear object (such as a wire or needle) and the image plane show up on the image clearly, as a bright spot. Automatic detection of small bright spots in an image is a relatively simple task and the position of the spot usually can be determined very accurately, therefore many calibration phantoms contain a number of wires at known positions. A particularly interesting setup is when wires are arranged in multiple N-shaped patterns (Fig. 6), because if the wire positions are known in 3D and the relative distances of the intersection points in the image are known in 2D, the position of the middle wire intersection can be computed in 3D [2]. Arranging wires in planes have the additional advantage that the intersection points in the image are collinear, which can be used for automatically rejecting bright spots in the image that do not correspond to an actual wire intersection point (Fig. 7). Having 3 N-shaped wire pattern is shown to be enough to reach submillimeter calibration accuracy [2]. Fully automatic, open-source implementation of the N-wire-based probe calibration is available in the Plus toolkit [3]. The advantage of the method is that is fully automatic, therefore a large number of calibration points can be collected and so the effect of random errors can be reduced, the results not depend much on the operator, and the calibration can be completed within a few minutes. The disadvantage of the method is that it requires measurement of the wire positions in the tracker coordinate system (typically by landmark registration of the calibration phantom), requires phantom fabrication, and attention has to be paid to set imaging parameters that allow accurate automatic detection of the wire intersections.

Other automatic methods have been proposed that use a simpler calibration phantom. For example, it is possible to compute the probe calibration just by imaging a flat surface while completing certain motion patterns with the transducer. This method is called *single-wall calibration*. The advantage of the method that it

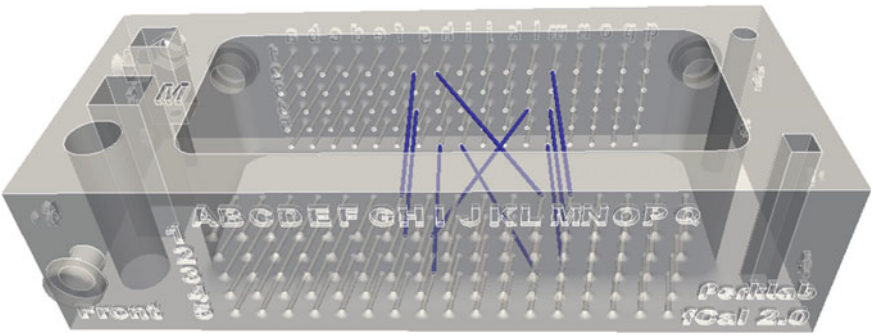


Fig. 6 Calibration phantom containing 3 N-wires. 3D-printing-ready CAD model, instructions, and calibration software are all available in the Plus toolkit [3]

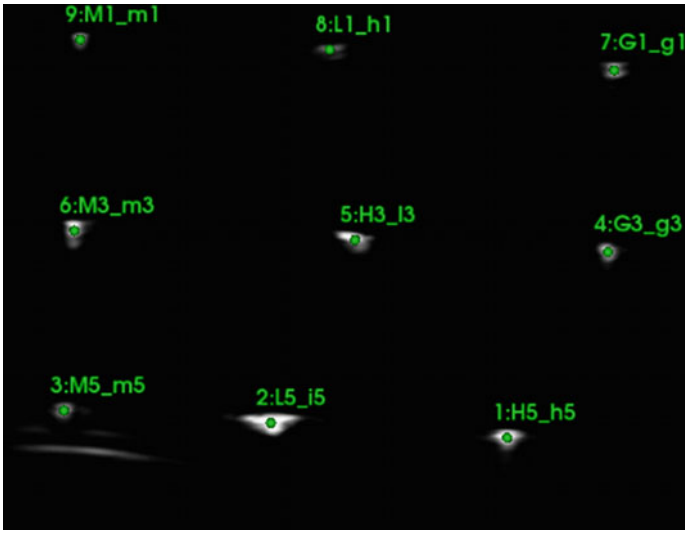


Fig. 7 Ultrasound image of the calibration phantom containing 3 N-wires with an overlay showing the results of the automatic marker detection algorithm

just require a simple flat diffusively reflecting surface as calibration phantom, however the method is not very robust and can provide very inaccurate results if the motion patterns are not completed carefully or not optimal imaging parameters are used.

The ultrasound imaging system is typically only loosely coupled to the position tracking system and there can be temporal misalignments between tracking and imaging data that is recorded at the same time. The goal of *temporal calibration* is to detect and compensate such temporal misalignments. Accurate temporal calibration is needed when images are acquired while moving the transducer. High accuracy and reliability is achievable using hardware triggers. If hardware-based

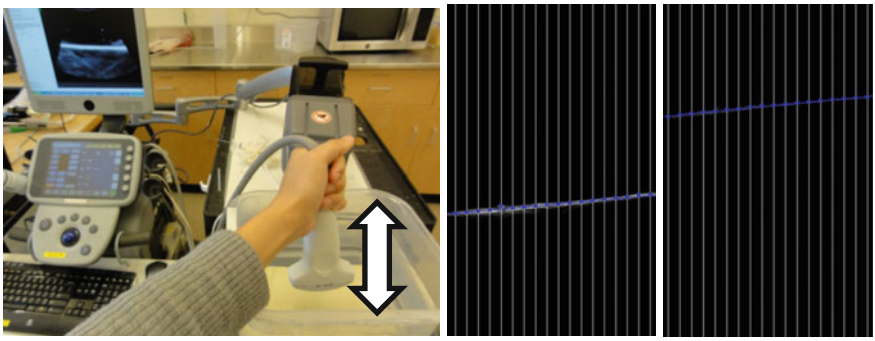


Fig. 8 Moving the transducer up/down repeatedly for acquiring tracking and imaging data for temporal calibration (*left*). Position of the water tank bottom is automatically detected in the ultrasound image and used as position signal for the image data. Position of the water tank bottom is shown for the *top* and *bottom* positions (*center, right*)

synchronization is not available but the acquisition rate and latency is constant in both the imaging and tracking device then software-based method can be used to compute the fixed time offset. Methods based on detecting certain events (such as sudden motion) have been proposed. These methods are easy to implement, but inaccurate or require lengthy data acquisition, because acquisition of a single measurement sample takes a few seconds. Correlation-based methods require the operator to perform quasi-periodic motion with the transducer for a few seconds and during this time imaging and tracking data is recorded (Fig. 8). Then position signal is extracted from the data and the time offset is computed that results in the highest correlation value between the position signals (Fig. 9). Position signal from the 3D pose information can be computed as position along the first principal axis of the motion. Position signal from the image data can be retrieved by detecting the position of a feature (such as the bottom of the water tank) and use the position along a chosen axis. The correlation-based temporal calibration method is accurate, reliable, and a free, open-source implementation is available in the Plus toolkit [3].

7 Volume Reconstruction of Tracked Ultrasound

Position of recorded ultrasound images can be used to reconstruct 3-dimensional ultrasound volumes. Reconstructed volume data can be in the same format as other volumetric images (CT or MRI), but the intensity values of voxels still highly depend on the direction of sound propagation. Therefore, processing and visualization of such volumetric images are difficult. Intensity values in ultrasound are not characteristic to tissue types, and are often attributed to artifacts (including scatter and shadow), rather than anatomical structures. Image quality and parameters also depend on the settings of the ultrasound scanner, the size of the patient, and motion patterns of the transducer during image recording.

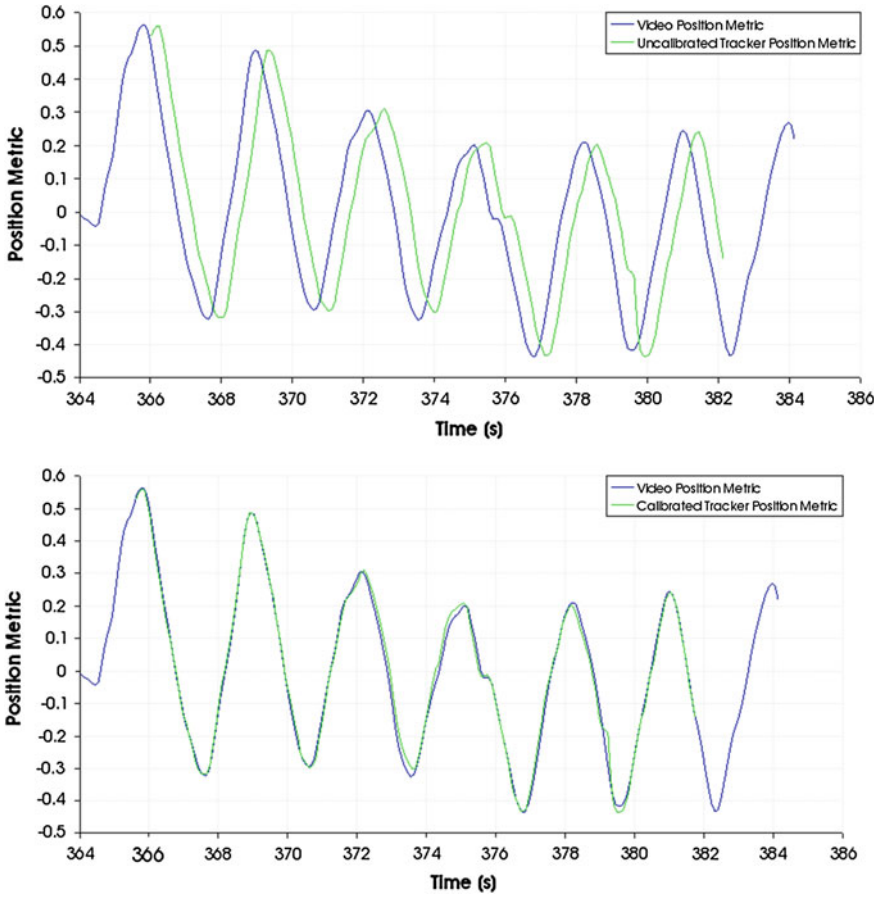


Fig. 9 Without temporal calibration the video and tracking data are misaligned (*top*). Temporal calibration minimizes the misalignment (*bottom*)

Reconstructed image volumes are often used in cross-modality image registration for fusion of ultrasound with pre-procedural CT or MRI images. These promising applications are still in research phase, but they may have a significant role in clinical practice in the future, as they combine the excellent tissue visualization features of other modalities with the safety, portability, and accessibility of ultrasound.

The quality of reconstructed ultrasound volumes depend on many factors, including the quality of the input images, calibration accuracy of the transducer tracker, the accuracy of temporal synchronization between image acquisition and position tracking, and the algorithms applied for filling voxels in the reconstructed volume where a recorded image is not available. Fortunately, there are a number of open-source implementations for ultrasound volume reconstruction algorithms.

8 Open-Source Software Tools for Rapid System Development

The complexity of image-guided needle navigation systems requires continuous software development and maintenance. Regular tasks include fixing errors, adding features, modifying the user interface, and adding support for new imaging and tracking hardware. Reliable software requires so much resources that it can only be achieved through a collaborative common platform that is shared between research groups and commercial partners. A medical engineering research group, or a medical device company would not develop a computer operating system, a programming language, or a computer graphics library. Similarly, they do not need to spend efforts on re-implementing device interfaces, calibration algorithms, or visualization methods, etc. To maximize productivity, they should focus on implementing new methods, building on previous results. Unfortunately previous results are typically published in journal and conference papers, which are not suitable to archive software methods. These publication are most effective if they are accompanied by an implementation of the published methods in an open-source software platform.

A commonly used software platform for tracked ultrasound system consists of two main parts (Fig. 10). The Public Software Library for Ultrasound Research (PLUS) implements lower level software components, including device interfaces,

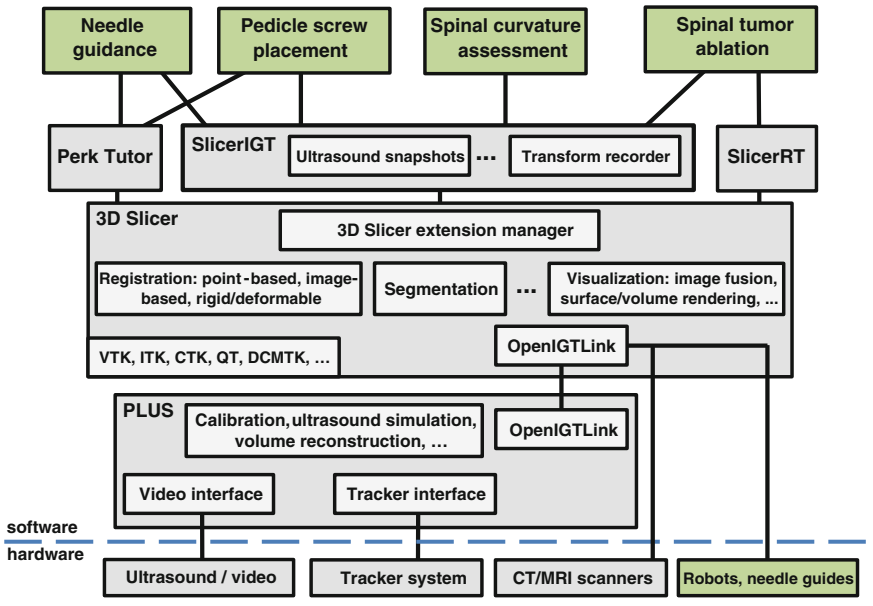


Fig. 10 Architecture overview of image-guided spinal disease diagnosis and treatment systems made from reusable software components

calibration methods, data acquisition, and data processing methods (e.g. 3-D volume reconstruction) [3]. PLUS is distributed under a permissive open-source license that allows both academic and commercial use without restrictions (www.plustoolkit.org). PLUS provides real-time data streams to end-user applications. Applications can be rapidly prototyped in the 3D Slicer framework (www.slicer.org). The advantage of 3D Slicer is that hundreds of medical image processing algorithms are implemented and deployed in this framework. They are readily available, and can be used for visualization that best helps intervention navigation.

9 Tracked Ultrasound in Interventions Training

Long learning curve is probably the only disadvantage of ultrasound guidance in spinal needle placement procedures [1]. The interpretation of musculoskeletal ultrasound images is difficult, and the operator has to do it in real time during interventions, while manipulating the ultrasound transducer in one hand and insert a needle with the other hand. This challenge is largely related to visuospatial coordination skills. Ideally, these necessary skills are learned before they are first performed on patients. Learning in a simulated environment on phantom models is not only safer for patients, but is also shown to improve the learning process [4]. Phantom models are proven tools in teaching spinal needle insertions to prepare medical residents for patient encounters [5]. Needle coordination skills in difficult procedures can be improved by providing augmented reality visual feedback while practicing the procedures on phantom models [6, 7].

Objective measurement of operator skills is of utmost importance in procedural skills training. Medical training is currently transforming according to the principles of competence-based medical education. The goal of this trend is to assure proper acquisition of skills before physicians perform interventions on patients. This demand requires simulation-based training and quantitative performance feedback for the trainees, as well as quantitative evaluation of skills. Teaching of ultrasound-guided spine interventions can greatly benefit from tracked ultrasound technology, both as an augmented reality system for improving visuospatial skills, and using tracking to objectively analyze hand motion data for skills evaluation. Systems with position tracking are inherently able to record motion trajectories that can be analyzed for qualitative and quantitative measures of procedural proficiency. Algorithms borrowed from artificial intelligence are shown to be able to classify motion gestures [8] and skill levels of the operator [9, 10].

Sonographic anatomy of the spine is difficult to master due to poor visibility and the complex shape of vertebrae. Tracked ultrasound along with tracked needle offers an excellent augmented reality training system. 3-D anatomical models of the training phantoms can be registered to the navigation scene to show what structures are responsible for characteristic features on the ultrasound image. When the needle is inserted incorrectly, the 3-D scene shows the trainee the exact relation of the actual needle position and the target point in relation to the spine anatomy. The

Fig. 11 A simulator with tracked ultrasound, tracked needle, and registered 3-D anatomical model for learning spatial coordination in spinal interventions

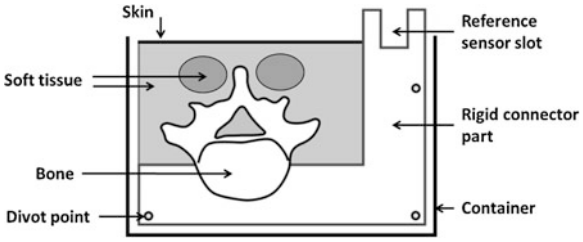


Fig. 12 Components of an ultrasound-guided spine intervention training phantom

spatial relations of tools and anatomy can be learned with such tracked systems [11] to improve needle coordination skills (Fig. 11).

The rest of this section gives an overview of how to build augmented reality training systems using position tracking to develop the skill of mental projection of the ultrasound image and needle trajectory on the patients in clinical procedures.

Commercial suppliers offer more and more spine simulation training models, but they can also be prepared from low-cost components (Fig. 12). A spine model can be rapid prototyped, or purchased from a supplier. It should be rigidly fixed by a connecting part to a reference tracking sensor holder, and some divot points should be marked on this rigid part for landmark registration. The space around the spine can be filled with organic or soft plastic gel, and the skin can be simulated by a rubber sheet.

Although there are several commercial and free products for ultrasound-guided spinal interventions, finding the best ways to teach and evaluate these skills is still subject to intensive research. Open-source platforms allow fast setup of research

prototypes that can be modified for new visualization techniques or evaluation metrics with minimal additional development work, such as the Perk Tutor platform [11].

Skill levels and performance scores of trainees are essential in any training program. Access to position data in tracked ultrasound and needle systems can be used to record tool trajectories, which correspond to hand motions of trainees. Recorded tool trajectories can be used in many ways to compute objective performance metrics. The most common performance metrics are total procedure time and needle insertion time. The latter corresponds to the total amount of time when the needle was inside the phantom. An important motion economy parameter is total needle path inside the phantom. Longer needle paths add up from multiple reinsertions and probing. These are clinically proven risks for infection and bleeding complications, therefore they are always good to be treated as primary measures of skill. Novice operators often do sideways or rotating motions with the needle, which is not recommended because the needle inside the tissue bends, which cannot be directly seen, so aiming at the correct target becomes more difficult. Sideways needle motion can be measured using the *potential tissue damage* parameter [7]. Procedures have specific success criteria that can be measured or observed during practice insertions to compute success rate. In case of lumbar puncture phantoms, the artificial spinal canal is usually filled with water, so the backflow of that water through the needle defines successful completion of the procedure. In facet joint injections or other nerve blocks the position of the needle tip may define success or failure. These metrics are readily implemented in the Perk Tutor platform.

Cost of the training system can be reduced by simulated ultrasound. Low cost training simulators are important because none of the training enhancement technologies substitute a good amount of hands-on practice. Trainees should ideally be given opportunity to deliberately practice until leaning objectives are met. Trackers are typically an order of magnitude less expensive devices compared to ultrasound machines. And ultrasound compatible training phantoms wear out over hundreds of needle insertions, which deteriorate ultrasound image quality. Simulated ultrasound can be generated from the tracked position of a needle and a non-functional ultrasound transducer. Simulated ultrasound has been shown to be useful in learning ultrasound skills usable with real ultrasound [12]. An open-source ultrasound simulator is available in the PLUS software library.

10 Extending Needle Navigation Techniques

Position tracking of the ultrasound and the needle extends the possibilities in ultrasound-guided needle insertion techniques. Direct visual aiming is only possible with conventional ultrasound when the needle is parallel to the ultrasound imaging plane. This is called the in-plane insertion technique (Fig. 13a). The out-of-plane technique (Fig. 13b) is more challenging, because the transducer needs to be moved back and forth, and the needle position needs to be assessed mentally from multiple

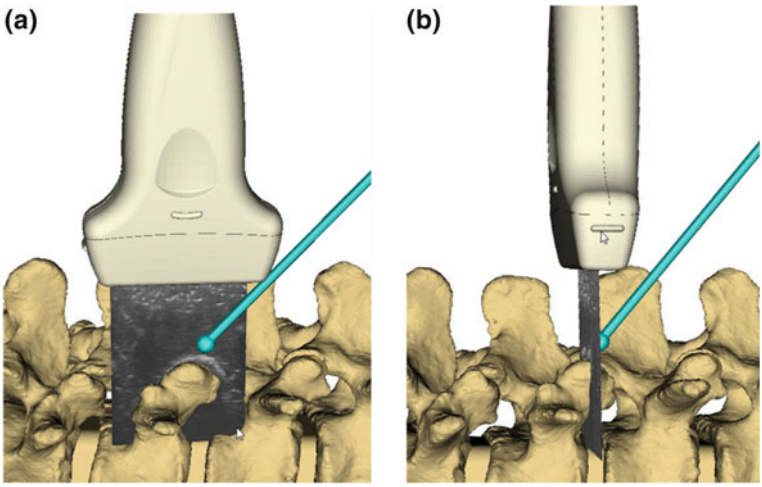


Fig. 13 In-plane and out-of-plane techniques in ultrasound-guided needle insertions

scanned images. But position tracking allows 3-dimensional visualization of both the ultrasound and the needle, allowing accurate needle aiming regardless of the ultrasound image orientation (Fig. 13).

11 Tracked Ultrasound Snapshot Technique

Simultaneous handling of the ultrasound transducer and the needle has two main disadvantages. It requires significant hand coordination skills, and the transducer physically limits the range of motion of the needle. The acoustic shadows of vertebrae limit angles and positions of the ultrasound transducer. The ideal, shortest path for the needle is often blocked by the transducer in real time ultrasound guidance. Therefore, the operator may sacrifice the ideal needle path for real time imaging. But tracked ultrasound offers separation of imaging and needle insertion in time. The optimal ultrasound image can be recorded relative to the patient anatomy. This image can be displayed for navigation when the transducer is removed from the patient, and the tracked needle can be guided along the recorded ultrasound snapshot. This technique, called tracked ultrasound snapshot (TUSS) guidance simplifies the hand coordination task, because the operator has to do only one thing at a time, imaging or needle insertion. TUSS also allows needle insertion at the same location that was used for imaging.

12 Facet Joint Injections with Tracked Ultrasound Snapshots

Facet joint injections are done routinely on a relatively large patient population with chronic back pain. The current standard of practice is either fluoroscopic or CT-guided needle placement. Ultrasound offers a radiation-free alternative to image guidance [13, 14], but it has not become a routine clinical procedure due to its difficulty. Tracked ultrasound improves the accuracy of needle placement when it is fused with a previous CT scan [15]. However, a CT scan is not always available for these procedures. In this section we describe the TUSS-guided facet joint injections, which potentially facilitates ultrasound-only guidance in facet joint injections.

Since needles can access the facet joints only in a constrained range of angles, real-time ultrasound guidance is inconvenient. TUSS allows the procedure to be separated in an imaging phase and a needle insertion phase. Initially, the operator finds the target facet joint, and records one or more ultrasound snapshots at the target. Then the ultrasound transducer is not needed during needle insertion, as the operator guides the needle tip to the targets defined on the snapshots.

Operator performance in TUSS-guided facet joint injections was compared to conventional ultrasound guidance in a cadaveric lamb model [16] (Fig. 14). Success rate and insertion time improved significantly in a pilot study (Table 2).

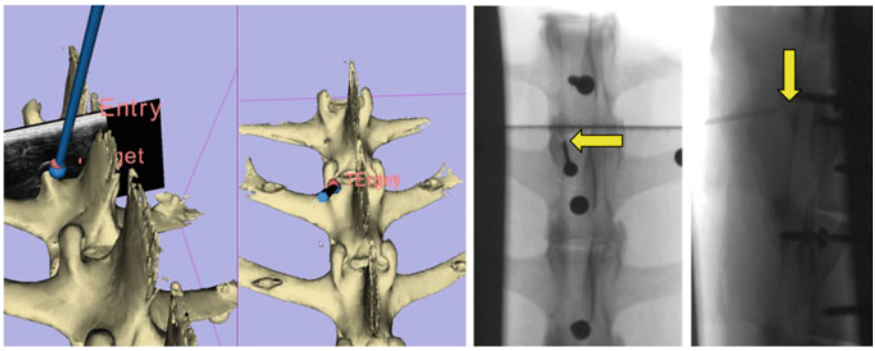


Fig. 14 Dual 3-D navigation scene for facet joint injection with registered CT-derived spine model in a lamb specimen. Radiographs on the right confirm the needle position (*arrows* point at the needle)

Table 2 Operator performance in TUSS-guided versus conventional US-guided facet joint injections in a lamb model

	TUSS guidance	US guidance
Number of insertions	50	50
Success rate (%)	*94	44
Insertion time (s ± SD)	*36.1 ± 28.7	47.9 ± 34.2

*$p < 0.05$ versus Freehand US guidance

The most important limitation of ultrasound and TUSS guidance in the spine is limited visibility of bone structures in ultrasound images. Visual enhancement of the spine could be achieved by fusion of a previous CT image to the tracked ultrasound [17], however, ultrasound-only procedures are preferred to reduce radiation risks and cost. Vertebra visibility could be improved in the needle navigation display by fitting a deformable general vertebra shape model to automatically detected bone contours [18]. Although shape model fitting is still in the experimental phase, and will likely have limitations in certain pathological cases, it may greatly enhance the potentials in ultrasound guidance in the spine.

13 Spinal and Epidural Anesthesia with Tracked Ultrasound Snapshots

Spinal and epidural anesthesia are similar procedures; the needle is just pushed a little further in case of spinal anesthesia. Both are performed to numb the lower body for surgery while the patient remains awake. These procedures are preferred over general anesthesia, having lower risks and the contributing to faster recovery after surgery. Spinal and epidural needles are both placed in the spinal canal. Spinal anesthesia is injected inside the dura sac, where the medicine takes effect immediately, and is usually used in shorter and simpler procedures. Epidural injections are placed just outside the dura sac. A catheter can be left in the epidural space to provide continuous administration of medicine for longer procedures. From the needle guidance point of view, the needle should be similarly navigated in the spinal canal between two lumbar vertebrae in both cases (Fig. 15).

Spinal and epidural anesthesia is routinely performed without image guidance, as the vertebral interspaces are palpable in the average patient. However, some pathological conditions may cause the narrowing of the interspaces, making it difficult or impossible to lead a needle to the spinal canal. In less severe cases, conventional ultrasound may help identify the interspaces where needle insertion can be attempted with higher probability of success, but in extreme cases, only a CT image-based guidance may provide enough information for needle navigation. Tracked ultrasound offers the accuracy of CT-guided navigation, using a pre-operative CT image, registered to the patient using landmarks visible on ultrasound images.

The most intuitive display for needle navigation is when vertebrae and the needle are represented with surface models. Surface models can be generated from CT images using a threshold-based segmentation, but pathological spines may require manual slice-by-slice contouring, especially in the presence of metallic implants. The CT-derived surface models can be registered to the needle navigation coordinate system using landmark points. The landmarks should be rigidly fixed to the vertebrae, and should be easy to identify on ultrasound images. Natural landmarks can be the facet joints, or transverse processes. In case of implanted vertebral screws, the screw heads are excellent landmarks (Fig. 16).

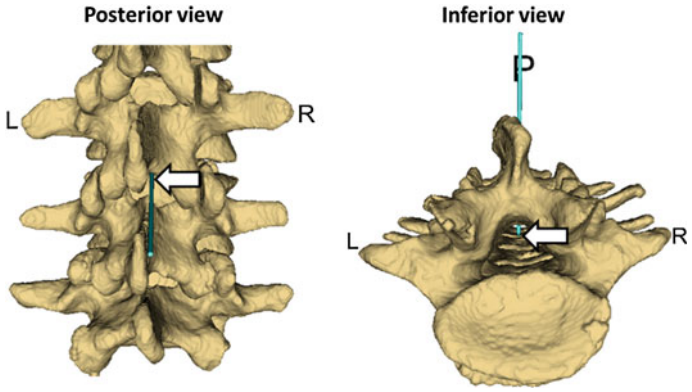


Fig. 15 Illustration of needle position in spinal and epidural injections relative to the lumbar spine in posterior and inferior views. The *arrow* points at the needle tip in both images

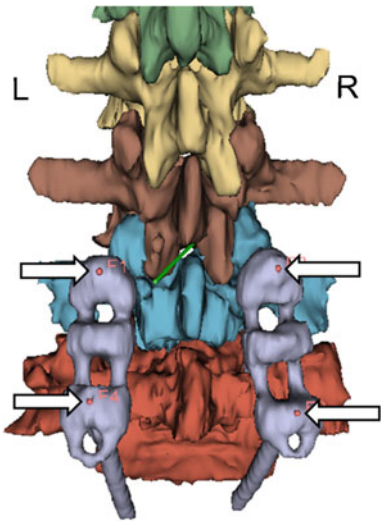


Fig. 16 Needle navigation scene for spinal anesthesia. In this patient, vertebral screws provide landmark points (*arrows*) for registration. The *green stick* shows the operator the ideal direction of needle insertion. In such degenerative spines, different colors for individual vertebra models make image interpretation easier

14 X-ray Dose Reduction in Pedicle Screw Navigation

One of the most popular subjects for computer-aided surgical navigation techniques is pedicle screw placement. There is an abundance of evidence that computerized navigation of surgical tools improves the outcomes of the surgeries, and reduces the probability of complications. Different navigation techniques share a common task, which is the spatial registration of the actual patient with the virtual model of the patient. The pedicle screw position is typically planned with respect to a pre-operative CT image. But the CT image needs to be registered with the patient on the

surgical table, so the navigation system knows where the screws positions are planned with respect to the patient.

Ultrasound can be tracked using the same tracking system that is used for surgical navigation. This allows ultrasound to identify landmarks for registration of the pre-operative plan to the surgical navigation system. The vertebra anatomy offers many unique surface landmarks, but few are convenient to identify in ultrasound images. The spinous process is hard to localize with ultrasound because of the prominent echo signal from the supraspinous ligament. The second closest structure to the skin that has a face perpendicular to the ultrasound propagation direction is the set of articular processes. The four articular processes are relatively easy to find in ultrasound images, and they surround the vertebra, therefore are excellent points for landmark registration.

The pre-operative CT can be accurately registered to intraoperative tracking using the articular processes as landmarks [19]. More landmarks can be defined to further reduce the effect of landmark position errors (Fig. 17), although at the cost if increasing the total procedure time.

15 Spinal Curvature Monitoring with Tracked Ultrasound Snapshots

Kyphoscoliosis is a condition with pathological curvatures of the spine. The most common cause of this condition is a disease called adolescent idiopathic kyphoscoliosis. It affects 1 individual in 1,000, and is typically discovered in the early adolescent age. It requires regular monitoring of the pathological curvatures, to be able to decide on treatment options in time. Spinal curvature measurement may also be needed during surgery to provide feedback on achieving the surgical plan. Spinal curvature measurements are currently performed on X-ray images in the clinical practice. However, regular examinations with X-ray have been linked to an increased risk of cancer [20–22]. Therefore, an alternative measurement method without ionizing radiation would be ideal for monitoring kyphoscoliosis angles.

In the current clinical practice, measurements are made on X-ray radiographs. The reader selects two vertebrae that are most angled at the superior and inferior end of the curvature. A line is drawn on superior end-plate of the superior vertebra, and on the inferior end-plate of the inferior vertebra. The angle between these lines is called the Cobb-angle, which is the most common measure of spinal curvatures. Minor curvature angles can also be defined besides the most prominent major angle. However, lots of factors cause variance in the Cobb-angle. The posture of the patient, the angle of X-ray imaging, and these curvatures are reported to increase within a day, begin up to 5° larger in the afternoon compared to measurements in the morning [23]. Since variability between different readers is reported to be 2°–7° even on the same images, spinal curvature differences less than 5° are generally not considered significant when estimating disease progression [24, 25].

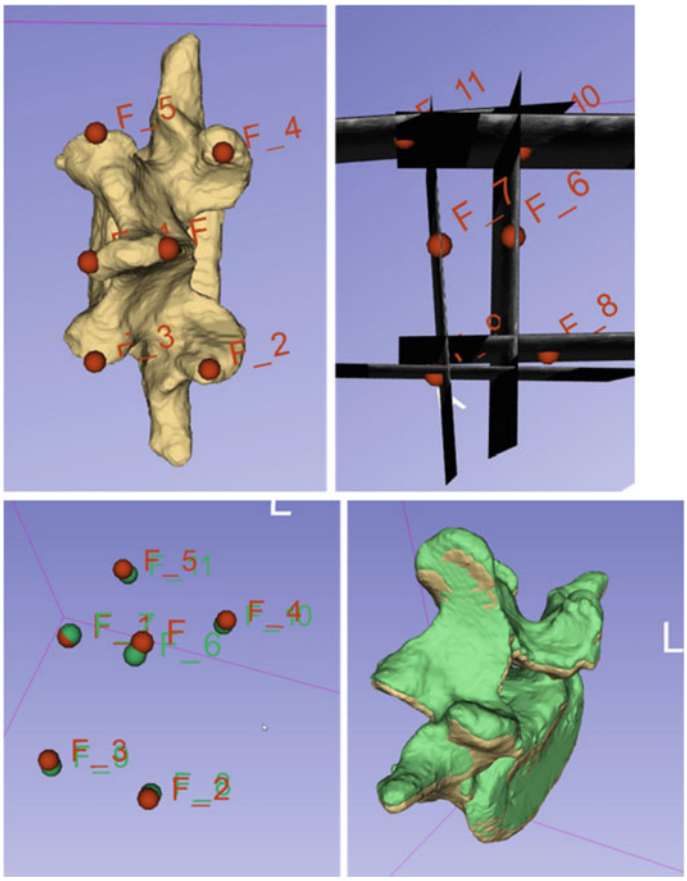


Fig. 17 Landmarks defined for registration on the CT-derived model of a lumbar vertebra (*top left image*), and the same landmarks defined on tracked ultrasound snapshots (*top right image*). The two sets of landmarks are registered (*lower left image*), and the registered vertebra position (*green*) is localized close to the ground truth position (*yellow*) in the *lower right image*

Tracked ultrasound offers accurate spatial localization of vertebra landmarks visible on ultrasound images. These landmarks are suitable for measurement of spinal curvature and vertebra rotation without ionizing radiation. Spinal curvatures are measured between two vertebrae that are rotated in the coronal plane at the largest angle. The angle is defined between two lines in the coronal plane. Both lines can be defined by two symmetric points on each vertebra. The points can be transverse processes on tracked ultrasound snapshots, as these points are visible on ultrasound images along the entire spinal column (Fig. 18).

Tracked ultrasound technique can provide as accurate spinal curvature measurements as X-ray images [26]. Although this method needs further clinical testing, as the conventional anatomical landmarks, the vertebral end-plates, cannot be

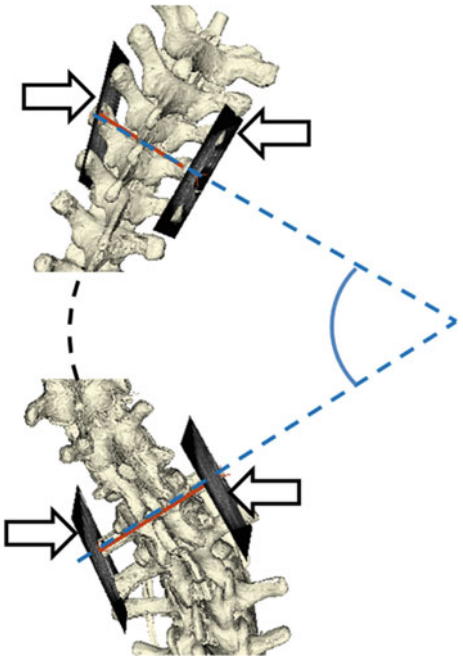


Fig. 18 Spinal curvature measurement using four landmark points from four tracked ultrasound snapshots (marked by *white arrows*). The 3D spine model illustrates the measurement principle, but it is not available in the clinical setting

used for ultrasound-based measurement that is used for X-ray measurement. The vertebral end-plates cannot be seen in ultrasound due to the acoustic shadow of the lamina and vertebral processes. Anatomical features that are accessible by ultrasound imaging and also visible in X-ray are transverse processes (Fig. 19).

Two potential advantages of using tracked ultrasound for spine curvature angle measurements are safety and accessibility. Radiation-free monitoring method in

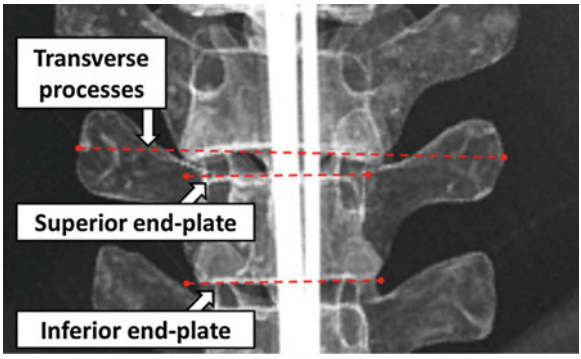


Fig. 19 Anatomical features for spinal curvature measurement. Superior and inferior end-plates are conventionally used in radiographic measurements. The transverse processes are also visible in ultrasound

adolescent kyphoscoliosis reduces the risk of cancer in these patients, as ultrasound has no known adverse side effects. Tracked ultrasound machines are also more accessible tools than X-ray machines. Portable ultrasound machines allow screening and monitoring in remote areas where permanent medical imaging facilities are not available. Therefore, tracked ultrasound may become the clinical standard for kyphoscoliosis monitoring in the future.

16 Ultrasound Image Fusion with Other Modalities

Ultrasound imaging lacks important features of CT or MRI modalities, including characteristic image intensity values for different tissues. Intensity values are relative on ultrasound due to attenuation, acoustic shadowing, and other artifacts. The ideal image guidance for the interventionist would have the standard image quality of CT and MRI, and also the convenient accessibility of ultrasound. Therefore, a great challenge for researchers and engineers is to fuse preoperative CT and MRI with ultrasound in real time during ultrasound scanning. If these preoperative images are registered to the patient anatomy, tracked ultrasound images can be enhanced by showing a corresponding slice from CT or MRI, either fused with the ultrasound, or side-by-side. Tracking ensures that both images show the same slice respective to the patient anatomy. Even though perfect spatial registration between preoperative images and intraoperative ultrasound cannot be achieved due to soft tissue deformations around the spine, and due to patient motion, physicians can mentally correct for these deformations, so the image fusion can help both even when the registration accuracy is limited.

CT-to-ultrasound or MRI-to-ultrasound fusion could also be used to eliminate needle tracking from interventional procedures. Ultrasound can be used to directly visualize the needle, and preoperative images show the target anatomical structures. Therefore, fusion of the two modalities may provide real time needle navigation in preoperative images. This potentially reduces the cost of disposable needle trackers, and extends the applicability of tracked ultrasound to interventional tools (e.g. tissue ablators) that are currently not equipped with position tracking.

Significant effort has been made to implement fusion of preoperative images with intraoperative ultrasound. The registration methods are either based on common image features between CT and ultrasound [27], or they use the surface model of the spine, which requires segmentation of the vertebrae [28]. A common problem in image registration is that the CT image is usually taken in supine patient position, while needle insertions are done while the patient is bent forward. This requires non-rigid registration of the CT image. Biomechanical constraints can be applied to account for the typical deformation of the spinal column. Unfortunately, rate of failed spine CT to ultrasound registration is reported to be significant, even under experimental conditions, both with image-based [17] and with surface-based algorithms [29]. Reported success rates are below 90 %, and clinical cases would

probably result lower success rate compared to the experimental environment, therefore, automatic registration of CT and ultrasound images require further research and development.

Ultrasound image fusion with other modalities has significant potential in transforming image-guided therapy applications. Its benefits are not limited to navigation of needle interventions. Other image-guided therapies including radiation therapy may also benefit from real-time, accurate localization of organs and pathological tissues.

References

1. Chen CP, Lew HL, Tsai WC, Hung YT, Hsu CC (2011) Ultrasound-guided injection techniques for the low back and hip joint. Am J Phys Med Rehabil 90(10):860–867
2. Carbajal G, Lasso A, Gómez A, Fichtinger G (2013) Improving N-wire phantom-based freehand ultrasound calibration. Int J Comput Assist Radiol Surg 8(6):1063–1072
3. Lasso A, Heffter T, Pinter C, Ungi T, Fichtinger G (2012) Implementation of the PLUS open-source toolkit for translational research of ultrasound-guided intervention systems. MIDAS J Med Imaging Comput (http://hdl.handle.net/10380/3367)
4. Palter VN, Grantcharov TP (2010) Simulation in surgical education. CMAJ 182 (11):1191–1196
5. Uppal V, Kearns RJ, McGrady EM (2011) Evaluation of M43B Lumbar puncture simulator-II as a training tool for identification of the epidural space and lumbar puncture. Anaesthesia 66 (6):493–496
6. Moult E, Ungi T, Welch M, Lu J, McGraw RC, Fichtinger G (2013) Ultrasound-guided facet joint injection training using Perk Tutor. Int J Comput Assist Radiol Surg 8(5):831-6
7. Yeo CT, Ungi T, U-Thainual P, Lasso A, McGraw RC, Fichtinger G (2011) The effect of augmented reality training on percutaneous needle placement in spinal facet joint injections. IEEE Trans Biomed Eng 58(7):2031–7
8. Datta V, Mandalia M, Mackay S, Chang A, Cheshire N, Darzi A (2002) Relationship between skill and outcome in the laboratory-based model. Surgery 131(3):318–323
9. Lin HC, Shafran I, Yuh D, Hager GD (2006) Towards automatic skill evaluation: detection and segmentation of robot-assisted surgical motions. Comput Aided Surg 11(5):220–230
10. Reiley CE, Hager GD (2009) Task versus subtask surgical skill evaluation of robotic minimally invasive surgery. Med Image Comput Comput Assist Interv 12(Pt 1):435–442
11. Ungi T, Sargent D, Moult E, Lasso A, Pinter C, McGraw RC, Fichtinger G (2012) Perk Tutor: an open-source training platform for ultrasound-guided needle insertions. IEEE Trans Biomed Eng 59(12):3475–3481
12. Bartha L, Lasso A, Pinter C, Ungi T, Keri Z, Fichtinger G (2013) Open-source surface mesh-based ultrasound-guided spinal intervention simulator. Int J Comput Assist Radiol Surg 8 (6):1043–51
13. Galiano K, Obwegeser AA, Bodner G, Freund M, Maurer H, Kamelger FS, Schatzer R, Ploner F (2005) Ultrasound guidance for facet joint injections in the lumbar spine: a computed tomography-controlled feasibility study. Anesth Analg 101(2):579–583
14. Loizides A, Peer S, Plaikner M, Spiss V, Galiano K, Obernauer J, Gruber H (2011) Ultrasound-guided injections in the lumbar spine. Med Ultrason 13(1):54–58
15. Moore J, Clarke C, Bainbridge D, Wedlake C, Wiles A, Pace D, Peters T (2009) Image guidance for spinal facet injections using tracked ultrasound. Med Image Comput Assist Interv 12(Pt 1):516–523

16. Ungi T, Abolmaesumi P, Jalal R, Welch M, Ayukawa I, Nagpal S, Lasso A, Jaeger M, Borschneck DP, Fichtinger G, Mousavi P (2012) Spinal needle navigation by tracked ultrasound snapshots. IEEE Trans Biomed Eng 59(10):2766–2772
17. Gill S, Abolmaesumi P, Fichtinger G, Boisvert J, Pichora D, Borshneck D, Mousavi P (2012) Biomechanically constrained groupwise ultrasound to CT registration of the lumbar spine. Med Image Anal 16(3):662–674
18. Khallaghi S, Mousavi P, Gong RH, Gill S, Boisvert J, Fichtinger G, Pichora D, Borschneck D, Abolmaesumi P (2010) Registration of a statistical shape model of the lumbar spine to 3D ultrasound images. Med Image Comput Comput Assist Interv 13(Pt 2):68–75
19. Ungi T, Moult E, Schwab JH, Fichtinger G (2013) Tracked ultrasound snapshots in percutaneous pedicle screw placement navigation: a feasibility study. Clin Orthop Relat Res 471(12):4047–4055
20. Hoffman DA, Lonstein JE, Morin MM, Visscher W, Harris BS 3rd, Boice JD Jr (1989) Breast cancer in women with scoliosis exposed to multiple diagnostic x rays. J Natl Cancer Inst 81 (17):1307–1312
21. Doody MM, Lonstein JE, Stovall M, Hacker DG, Luckyanov N, Land CE (2000) Breast cancer mortality after diagnostic radiography: findings from the U.S. Scoliosis Cohort Study. Spine (Phila Pa 1976) 25(16):2052–2063
22. Schmitz-Feuerhake I, Pflugbeil S (2011) 'Lifestyle' and cancer rates in former East and West Germany: the possible contribution of diagnostic radiation exposures. Radiat Prot Dosimetry 147(1–2):310–313
23. Beauchamp M, Labelle H, Grimard G, Stanciu C, Poitras B, Dansereau J (1993) Diurnal variation of Cobb angle measurement in adolescent idiopathic scoliosis. Spine (Phila Pa 1976) 18(12):1581–1583
24. Malfair D, Flemming AK, Dvorak MF et al (2010) Radiographic evaluation of scoliosis: review. AJR Am J Roentgenol 194(3 suppl):S8–S22
25. Sardjono TA, Wilkinson MH, Veldhuizen AG, van Ooijen PM, Purnama KE, Verkerke GJ (2013) Automatic Cobb angle determination from X-ray images. Spine (Phila Pa 1976)
26. Ungi T, King F, Kempston M, Keri Z, Lasso A, Mousavi P, Rudan J, Borschneck DP, Fichtinger G (2013) Spinal curvature measurement by tracked ultrasound snapshots. Ultrasound Med Biol (in press)
27. Yan CX, Goulet B, Tampieri D, Collins DL (2012) Ultrasound-CT registration of vertebrae without reconstruction. Int J Comput Assist Radiol Surg 7:901–909
28. Herring JL, Dawant BM, Maurer CR Jr, Muratore DM, Galloway RL, Fitzpatrick JM (1998) Surface-based registration of CT images to physical space for image-guided surgery of the spine: a sensitivity study. IEEE Trans Med Imaging 17:743–52
29. Rasoulian A, Abolmaesumi P, Mousavi P (2012) Feature-based multibody rigid registration of CT and ultrasound images of lumbar spine. Med Phys 39:3154–3166

Robotic Assistance and Intervention in Spine Surgery

Rajesh Kumar

Abstract Robotic assistance and intervention methods have now been in use in several surgical specialties for nearly three decades. While image guided surgery for orthopedics (led by such developments as the Integrated Surgical Systems Inc. ROBODOC system) lead early clinical application, it is laparoscopic and telesurgical robotic applications (such as with the Intuitive Surgical Inc. da Vinci surgical systems) that have found the widest clinical and user acceptance. Orthopedic and neurosurgical robotic systems have yet to be widely applied as robotic assistance systems (such as Mazor Robotics SpineAssist or Renaissance systems) have only been recently approved for clinical use. Given the large volume of spinal procedures such as pedicle screw placement for spinal fusion, vertebroplasties, osteotomies, biopsies and other spinal surgeries several other image-guided robotic systems are in advanced research and development. The goals of these devices include improving the efficacy and safety (including radiation safety) for both the patient and the surgeon. This survey includes recent development and results for these robotic applications.

1 Overview

Preoperative fluoroscopic imaging is commonly used to plan orthopedic surgeries as well as for enabling robotic [1] spine applications (Fig. 1). Targets in hard anatomy are typically immobilized, and imaged using commonly accessible fluoroscopy (CT, Biplanar or C-Arm) imaging. Fiducials with unique fluoroscopic signature attached to the device or anatomy are used to aid registration as well.

For intra-operative guidance or robotic automation, accurate registration to anatomy is the most important procedure step, while safety and efficacy are the most important factors from the patient perspective.

R. Kumar (✉)
Rajesh-Kumar.org, Fremont, CA 94555, USA
e-mail: rajesh@rajesh-kumar.org

© Springer International Publishing Switzerland 2015

S. Li and J. Yao (eds.), *Spinal Imaging and Image Analysis*,
Lecture Notes in Computational Vision and Biomechanics 18,
DOI 10.1007/978-3-319-12508-4_16

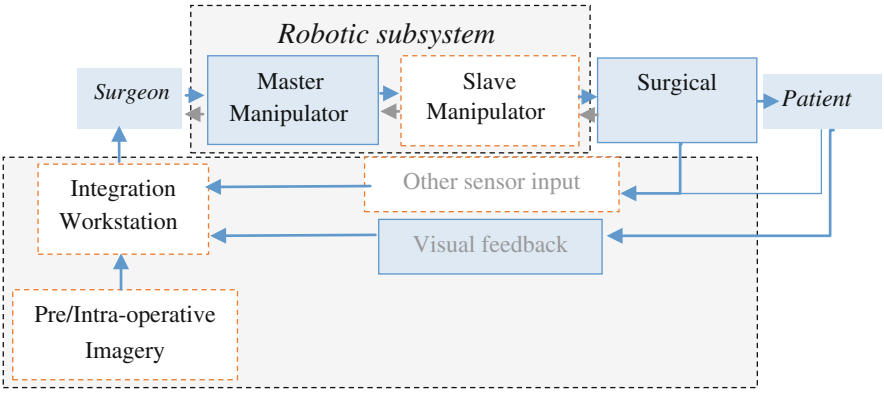


Fig. 1 The role of robots in surgery: a robotic device is an accurate, intelligent intermediary between the surgeon's intent and action of the surgical instruments on the patient. Both telesurgical (master-slave) and directly manipulated (a surgeon's assistant, where the slave subsystem is omitted) robotic paradigms have been used in augmented surgical procedures. For example, the former paradigm is used in widely used robotic surgery systems such as the da Vinci surgical system (Intuitive Surgical Inc.), and the latter in the MAKOplasty RIO orthopedic devices (MAKO Surgical Corporation, now part of Stryker). While visual feedback is always available to the surgeon in these procedures, neurosurgical and orthopedic robotic procedures typically include target guidance and navigation assistance integrated using a powerful computer workstation that uses intra-operative sensory feedback for registration of pre-operative and real-time imagery to provide augmented target visualization

In soft-tissue surgery, the lack of appropriate safety and efficacy evidence has hampered adoption in precision dependent specialties such as cardiology [2]. Similarly, neurosurgery has yet not widely used robotics [3].

By contrast, spine surgery provides a much closer approximation to other orthopedic surgery procedures (for example, the hip, knee, or the shoulder surgery) and the surgical goals and tolerances are relatively easily defined in geometric terms and robotic parameters. Recent wider acceptance of robotics among orthopedic surgeons has made some procedures both more commonly available as well as made robotics an invaluable tool. In their reviews of orthopedic robots, Mavrogenis et al. [4] present a clinical perspective on the adoption of orthopedic robots. Many reviews from engineering perspectives have also been published; for example, Sistona et al. [5] review the surgical navigation aspects for knee applications.

Orthopedic surgery performs fusing, shaping, or cutting of bones for either providing access or creating cavities for placement of implants. Many spinal surgery tasks, such as pedicle screws, spinal fusion, or disk implant placement are closely similar tasks from a robotics perspective to robotic orthopedic tasks. However, spinal surgery introduces unique challenges for the surgeon due to the complex three-dimensional anatomy, and placement of critical and delicate neural and vascular structure in close proximity to the bony anatomy being operated.

Conventional navigation mainly relies on identification of bone surface anatomy in 2D fluoroscopy images, providing robotic guidance and operation a great opportunity

to improve safety, accuracy, as well as reduction in radiation exposure. Well-defined geometric accuracy goals [6] that prevent damage to the spine or peripheral nerves can be used for equivalent safety and efficacy evaluation. Large surgical volumes in procedures such as vertebroplasties, fusions, and biopsies and other spinal procedures [7, 8] strengthen the corresponding business cases for any devices.

1.1 Background

Robotics research in surgery is now several decades old (Fig. 2). The first widespread clinical robotic application was the ROBODOC system robot designed for automated milling of the hip cavity for implant placement in a hip replacement surgery [9]. This automation was particularly well received by users. While 10,000s of procedures were performed when the system was in active use, the lack of efficacy data for comparison with convention practice, and improvements in conventional treatment options, lack of FDA clearance, among other reasons withheld the system from reaching its full potential.

A version of the system has recently received FDA clearance for milling of the implant cavity for hip surgery. ROBODOC Applications were also extended to include other joint reconstruction such as the knee, the domain of later more successful applications such as MAKOplasty. An integration workstation (Fig. 1) provided the registration and the user interface between the surgeon and the robot. The corresponding ORTHODOC planning software was designed to help a surgeon to graphically position a CAD model over a patient's CT scan.

Registration determined the intra-operative spatial relationships between surgical instruments held by the robot, pre-operative CT imaging and the milling plan and the anatomy. Implanted pin fiducials, and later intra-operative surface digitization using a robot held digitizer was used for robot/CT/bone surface registration.

Fig. 2 Examples of research systems: a NEUROMATE (Integrated Surgical Systems, Inc.) neurosurgical experimental setup at the Johns Hopkins University (*left*), and a user manipulating the first generation "steady-hand" surgeon's assistant microsurgery system (*right*)

After initial manual preparation during surgery, this milling plan was executed automatically. After the milling, the surgery proceeded without the robot for the remaining portion. A contemporary robotic system for orthopedic surgery (the German CASPER system [10]) used a similar clinical workflow as the ROBODOC and was also not successful.

Similarly, the contemporary FDA-approved neurosurgery system (the NEUROMATE; Fig. 2, left) too was not widely applied clinically [11] and now exists primarily for use in further research and development. The reviews [12, 13] detail other early orthopedic robotic research applications. By the turn of the century, robotics research was receiving active attention for many other surgical specialties, such as eye surgery (the JHU "steady-hand" system [14]) that have also yet to be commercialized.

By comparison, led by successful robotic systems such as the AESOP (Computer Motion, Inc; now owned by Intuitive Surgical, Inc) camera assistants [15] for laparoscopic procedures (a modified SCARA architecture, with a passive remote center of motion), robots have achieved a much greater user acceptance. Other robotic surgery systems including the MIRO system [16] or the University of Washington RAVEN prototypes, and commercial complete general laparoscopic surgery robotic systems such as the da Vinci telerobotic system (Fig. 3) [17–19].

The da Vinci Surgical system (Fig. 3) remains the dominant commercially available telerobotic minimally invasive surgery system. This telerobotic system

Fig. 3 The da Vinci surgical system (Intuitive Surgical, Inc.): a first-generation slave (patient-side manipulators) setup for a training session (*left*), and a second-generation master console (*right, top*) and the system in use by a trainee surgeon and bottom. Until the introduction of near infrared fluorescence integration recently, only visual imaging was available for feedback to the surgeon. Any additional imaging was viewable as a video (*picture in picture*) piped from external consoles (*right bottom*). A new version of the da Vinci system—the da Vinci Xi—has been introduced in 2014 that significantly improves positioning of the surgical instruments over the patient and reduces the operating room footprint

scales down the hand-motion depicting surgeon's intent between the master manipulators and their configurable associated slaves. Laparoscopic robotic surgery with the da Vinci surgical systems is now widely used beyond prostate surgery. Active specialties include complex gynecological procedures [20], partial nephrectomies and other urological procedures [21]. Development and adoption for other procedures continues in cardiac surgery, head/neck surgery [22], and many other applications.

Now in the fourth generation (the da Vinci Xi), the da Vinci consists of several parts. A surgeon's console contains the control handles (master manipulators) that are driven by the surgeon using laparoscopic instrument like grips, while viewing an auto-stereoscopic endoscopic view of the surgical site. A set of patient side manipulators hold the camera and the surgical instruments, and associated computing and stereo-endoscopic vision equipment completes the setup. With the instrument degrees of freedom included, the slave robots can be configured to have up to seven degrees of freedom in total.

A da Vinci system may mount up to four instruments, with one restricted to being the stereo endoscopic camera. The third-generation systems (the da Vinci Si) first allowed for up to two surgeon' consoles [23] to be used simultaneously. A large catalog of 8 mm wristed rigid and 5 mm articulated (snake-like) removable flexible surgical instruments can be interactively mounted during surgery for specific surgical tasks (e.g. cutting, suturing, or cautery) as needed.

As noted by Shuford et al. [18], a majority of the approximately 75,000 radical prostatectomies performed in the United States annually were performed robotically by 2007, rising from only 18,000 procedures in 2005. Multiple large population and long-term studies show comparable or favorable performance of robotic methods [19] in urology.

1.2 Limitations of Current Systems

Current robotic systems suffer many limitations in addition to the substantial initial system cost, annual maintenance expenses, and the higher cost of the disposable surgical instruments compared to laparoscopy. Significantly long learning curves for clinical proficiency have also been reported.

These existing systems can't be used in image-guided surgery as currently designed due to their size, limited accuracy, and interference with the conventional clinical workflow. Currently reported spine applications using the da Vinci systems are mostly forward-looking procedure development similar to other specialties such as head/neck surgery [22].

In such procedure development, intra-operative devices for registration have included instruments held by the robot, fiducials combined with C-Arm or other fluoroscopy, or optical and electromagnetic trackers such as the Axiem EM tracking (Medtronic Inc), or the Optotrak or Polaris systems (Northern Digital, Canada) systems. Intra-operative registration is performed similar to the spine navigation

applications or the orthopedic surgery applications discussed above, typically by digitizing sufficient numbers of corresponding anatomical landmarks necessary for accurately computing a rigid registration.

The remaining sections of this article are limited to a narrow engineering perspective on robotic surgery developments specifically related to spine surgery. Non-robotic surgical navigation and registration are addressed separately elsewhere in this collection.

We also omit simulators for surgical training that may include robotic interfaces, robots for anatomical testing or rehabilitation, and external therapy or imaging robotic devices such as those for cone-beam computed tomography.

The reader is referred to clinical reviews for perspectives on human usage, for example, Bertelsen et al. [24] present a clinical perspective of spine related robotics and Rozer et al. [25] focus on the development and use of the commercial spine robots. For an even broader engineering review of robotic surgery the reader is referred to [26], or similar reviews [27, 28], and for clinical perspectives specifically on spinal robotics developments to reviews such as [8, 24, 25].

2 Recent Research and Developments

While clinical evaluations of robotic surgery often note [21] the need for additional information overlay as it is difficult for a human to interpret multiple sources of information presented together, it is also non-trivial to establish and maintain any registration in soft-tissue while it is being manipulated.

Integration of imaging to robotic systems may also add to the radiation exposure to the surgeon, patient, and staff which can be significant for long cases, revisions, in the presence of abnormal anatomy or due to additional imaging required for registration where fiducials are not visible.

Additional procedure steps also increase total procedure time. Research and development has therefore, tried to address both clinical workflow integration as well as safety issues.

2.1 Research in Systems and Procedures

Conventional procedures still perform the majority of spine surgeries. A retrospective review [8] of a very large number of spine surgeries performed over 3 years (108,419 procedures between 2004 and 2007) at multiple medical centers notes that only 13.2 % were performed using a minimally-invasive approaches.

In current commercially available general surgery robotic systems, the surgeon is guided only by real-time endoscopic video available from the stereo endoscopic cameras, but the use of registered imaging is integral to robotic spine surgery. Registration enables mid-course corrections and iterative progress [12] since errors

can be corrected by retargeting or in the following iteration but if removed, material cannot be replaced. This is especially important if a previous surgical repair already exists in the area to be operated on.

Several research systems have been investigated for spine surgery. Comparatively easier procedures such as radiotherapy or percutaneous needle based interventions account for the majority of recent relevant research. Literature identifies over a dozen different robots in development or testing for spine procedures, however only one was specifically designed for spinal surgery [24] as therefore deserves a detailed account here.

Previously developed for percutaneous kidney access, a Georgetown robot [28] integrating RCM/PAKY modules has also been used for spine surgeries. Its registration software used preoperative CT scans and intra-operative fluoroscopic images. Intra-operative guidance permitted the surgeon to then operate the robot interactively. Other research systems of note include the MIRO [29] developed by the German Aerospace Centre DLR have also investigated for pedicle screw placement [16], the Z-KAT prototype system that preceded the MAKO Rio system [30], the Innomotion device [31], and the SpineNav device [32].

Spine robotic surgery may also require integrated haptic feedback. While force sensing and haptic feedback are active areas of research, any sensed forces are typically displayed with a PHANToM (SensAble Technologies Inc.) a common haptic device with passive end-effort and only suitable for point force displays. Research systems such as the MIRO system envision integrating force sensing in their design, and force sensing [33] and display have also been designed for the da Vinci system, though not currently enabled in the commercial product.

Of particular note here is the parallel mechanism developed by Shoham et al. [34]. This small cylindrical (5×7 cm^3, 200 g) 6DOF Miniature Robot for Surgical procedures (MARS) robot was designed for spinal pedicle screws placement, guidance for intramedullary nailing, and other spine applications. The prototype claimed both ergonomics and safety benefits from a small footprint while reporting sub-millimeter (~ 0.1 mm) accuracy and was later developed into the Mazor Robotics SpineAssist system.

2.2 Current Commercial Systems

Robotic spine procedures are still in early development. As was the case with other now widely used procedures, early spine surgery procedure development has also been attempted with the da Vinci surgical systems. We describe some spine procedure applications of this system below.

Mohr et al. [35] attempted spinal surgery using a da Vinci S Surgical Robot in a non-survival porcine surgery. These procedures were performed by a senior spine surgeon (though obviously one with very little prior robotic experience), with prototype and conventional da Vinci instruments developed for other approved surgical procedures.

The goal of these surgical procedures was to mimic conventional technique for dissection of 3 spinal lumbar levels as well as laminotomy, laminectomy, disc incisions, and dural suturing, As it was the case with most early procedures developed, video recordings were made and procedure times, ergonomic assessments, and surgical mistakes or difficulties noted for possible improvements in surgical technique.

Another report by Yang et al. [36] describes lumbar fusion using da Vinci Surgical System. Positioning the patient in a steep Trendelenburg position, this six-hour procedure was only possible while accommodating robotic arm complications related to unresponsive instruments as well as collisions of instrument holding robots outside the body due to suboptimal arm configurations necessary for the procedure. While configuration and software issues did not result in any significant nerve or vessel injury, the report confirms the challenges of adopting a robot not designed for spine surgery.

Yang et al. [37] also performed a da Vinci odontoidectomy on a cadaver with fewer ergonomics issues and reduced soft tissue damage, and note benefits from improved ergonomics and access due to increased freedom of movement at the wrist of the da Vinci instruments.

Further along in the procedure development curve, [38, 39] describe human lumbar fusion at L5–S1 using new Gelpoint access ports and new robotic instruments performed by an experienced surgeon. While reduced, the challenge of maintaining pneumoperitoneum during the discectomy, especially during placement of the interbody cage remains significant even with smaller incisions as noted by Beutler et al. [38].

Some of the challenges in this report may also relate to inexperienced surgeons who also had no prior robotic surgery experience. Surgeons in head/neck surgery and skull base procedures using transperitoneal, transthoracic, and transoral approaches [22] have adapted to some of these limitations with greater success. However, the overall assessment remains that these developments are forward looking research aimed at creating a knowledge base for robotic systems designed specifically for spine surgery.

Mohr et al. [35] also admit that many obvious challenges remain before spine surgery using such a large system could be considered a realistic application. A major challenge is arranging fluoroscopic imaging in an operating room along with large footprint robot and accommodating the additional personnel. This may require pre-operative imaging prior to the robot being moved in place, and some times prevent real image-guided surgery.

Literature also questions the large investment needed (up to $1.7 million) in system costs and annual maintenance, if expanded specifically for spine surgery, in the absence of an established or approved spine procedures, or trained surgeons.

Instead of justifying such expenses for an ergonomic aid, it is hoped that hospitals with under-utilized da Vinci surgical systems might benefit from the addition of other specialties performing procedures and helping to amortize overall system and maintenance costs. In current development approaches, intra-operative

verification is not typically attempted thereby eliminating the major accuracy advantage of a robotic system.

By Comparison, similar to other navigation tools the SpineAssist [25] and the current Renaissance system (Mazor Robotics) have been extensively used in human spine surgery. In [39] authors note that 646 pedicle screw placements post-operatively assessed by CT imaging resulted in 98.3 % meeting clinical accuracy criteria with a average deviation of 1.2 ± 1.49 and 1.1 ± 1.15 mm on the axial and sagittal planes respectively. This large retrospective 14-center study performed 3,271 total spinal implants inserted under SpineAssist guidance with no irreversible nerve damage reports.

The Renaissance system that followed SpineAssist is also receiving favorable attention in spine surgery since it may improve accuracy and result in reduced radiation exposure even for minimally invasive surgery. It has been used in over 20,000 procedures (at more than 34 centers in 2013) with implant placement accuracy reported to be better than 1.5 mm [40]. A proprietary fiducial array, and rigid attachment to the patient are credited for the accuracy. Registration is performed using two fluoroscopy images.

Peer-reviewed large volume studies are now beginning to appear in the literature. For example, the Renaissance and preceding Mazor robotic devices have been used for the placement of pedicle screws using a preoperatively planned trajectory [41, 42] for nearly a decade. This technology shows the promise of improving outcomes in both the accuracy of placement of spinal instrumentation, as well as reducing the radiation exposure.

With increasing user acceptance, other surgical uses of this robotic guidance are also appearing on the horizon. Mazor Robotics has received U.S. FDA clearance for the Renaissance system enables to be used in brain surgery and it has also been used in several brain surgeries in Europe.

3 Ongoing Work and Future Prospects

The greatest challenge to using existing general-purpose robotic surgery systems in constrained environments is their large size that makes it very difficult to integrate them in image-guided surgery clinical workflow. More compact systems being designed now, including systems that may attach to the patient bed or operating room ceiling may alleviate some of the clutter. New instrumentation and techniques will be developed along with these new systems.

As also noted in other reviews [1, 6, 7], the small number of surgical robots currently in use show great potential to improve surgical outcomes especially when accuracy and minimal invasiveness are needed or access is complicated. In several portions of the spine, target bone volume is small and the arteries, nerve roots and spinal cord are all closer to the vertebral bone complex, robotics may be especially enabling. This is particularly true of robots to be designed or adapted for the cervical spine.

Accuracy/efficacy is difficult to measure without post-operative assessment. Therefore, in addition to improving instrumentation (e.g. new laparoscopic distraction devices are in development), future systems will need to provide a lot more data, and still somehow reduce overall operating time to be comparable to conventional open or minimally-invasive surgery.

The promise of natural orifice transluminal surgery (NOTES) is exciting to both general and spine surgeons alike. However, the same key issues need to be addressed prior to the incorporation of NOTES into spine surgery practice. NOTES instrumentation has only been used for relatively simple surgeries (e.g. appendectomies and cholecystectomies).

Current research and development is mainly focused on evaluating the efficacy of the robotic devices that have been relatively recently introduced in practice. Important concerns such as optimization of device user interfaces, training for robotic devices, learning curves for these devices and corresponding training curricula are beginning to emerge. As with robotic laparoscopic surgery, these developments precede the development of standard testing and accreditation methods likely to be needed in the future.

Procedure development work identifies [35, 36] that cost, complexity, and lack of spine specific instrumentation offset many advantages of a robotic surgery platform. It is likely that the progress in these areas will continue to be made only in the context of systems and applications that are already deployed.

We envision that as next generation systems become available to spine surgeons and experience with these systems increases, the pace of development of robotic surgery systems should also hasten and lead to complete active robotic systems designed specifically for image-guided spine surgeries.

References

1. Matis GK, Silva D, Chrysou OI, Birbilis TA, Bernardo A, Stieg PE (2012) Robotics for spinal operations: reality or Alice in Wonderland? Int J Med Robotics Comput Assist Surg 8:125–126
2. Chitwood W Jr (2005) Current status of endoscopic and robotic mitral valve surgery. Ann Thoracic Surg 79(6):2248–2253
3. Sutherland GR, McBeth PB, Louw DF (2003) NeuroArm: an MR compatible robot for microsurgery. In: International congress series, vol 1256. Elsevier, pp 504–508
4. Mavrogenis AF, Aavvidou OD, Mimidis G, Papanastasiou J, Koulalis D, Demertzis N, Papagelopolos PJ (2013) Computer-assisted navigation in orthopedic surgery. Orthopedics 36 (8):631–642
5. Sistona RA, Giori NJ, Goodman SB, Delp SL (2007) Surgical navigation for total knee arthroplasty: a perspective. J Biomech 40:728–735
6. Gertzbein SD, Robbins SE (1990) Accuracy of pedicular screw placement in vivo. Spine 15 (1):11–14
7. Kantelhardt SR, Martinez R, Baerwinkel S, Burger R, Giese A, Rohde V (2011) Perioperative course and accuracy of screw positioning in conventional, open robotic-guided and percutaneous robotic-guided, pedicle screw placement. Eur Spine J 20(6):860–868

8. Hamilton DK, Smith JS et al (2011) Rates of new neurological deficit associated with spine surgery based on 108,419 procedures: a report of the scoliosis research society morbidity and mortality committee. Spine 36(15):1218–1228
9. Taylor RH, Mittelstadt BD, Paul HA, Hanson W, Kazanzides P, Zuhars ZF, Williamson B, Musits BL, Glassman E, Bargar WL (1994) An image-directed robotic system for precise orthopaedic surgery. IEEE Trans Robot Autom 10(3)
10. Paul A (1999) Surgical robot in endoprosthetics. How CASPAR assists on the hip. Fortschritte der Medizin 141(33):18
11. Xia T, Baird C, Jallo G, Hayes K, Nakajima N, Hata N, Kazanzides P (2008) An integrated system for planning, navigation and robotic assistance for skull base surgery. Int J Med Robot Comput Assist Surg 4(4):321–330
12. Yao J, Taylor RH, Goldberg RP, Kumar R, Bzostek A, Vorhis RV, Kazanzides P, Gueziec A (2000) A C-arm fluoroscopy guided progressive cut refinement strategy using surgical robot. Comput Assist Surg 5(6):373–390
13. Yao J, Taylor RH, Goldberg RP, Kumar R, Bzostek A, Vorhis RV, Kazanzides P, Gueziec A, Funda J (1999) A progressive cut refinement scheme for revision total hip replacement surgery using C-arm fluoroscopy. In: Proceedings of medical image computing and computer assisted intervention, Lecture notes in computer science (LNCS), vol 1679. pp 1010–1019
14. Kumar R (2001) An augmented steady hand system for precise micromanipulation. Ph.D. thesis, The Johns Hopkins University
15. Sackier JM, Wang Y (1996) Robotically assisted laparoscopic surgery: from concept to development. In: Computer-integrated surgery. MIT Press, Cambridge, pp 577–580
16. Ortmaier T, Weiss H, Döbele S, Schreiber U (2006) Experiments on robot-assisted navigated drilling and milling of bones for pedicle screw placement. Int J Med Robot 2(4):350–363
17. Guthart GS, Salisbury JK (2000) The intuitive telesurgery system: overview and application. In: IEEE international conference on robotics and automation (ICRA), pp 618–621
18. Shuford MD (2007) Robotically assisted laparoscopic radical prostatectomy: a brief review of outcomes. In: Proceeding Baylor University Medical Center, vol 20. pp 354–356
19. Thali R, Shah K, Patel VR (2007) Robotic surgery clinical review: applications of robots in urology. Robotic Surg 1(1):3–17
20. Boggess J (2007) Robotic surgery in gynecologic oncology: evolution of a new surgical paradigm. J Rob Surg 1(1):31–37
21. Bhayani SB, Snow DC (2008) Novel dynamic information integration during da Vinci robotic partial nephrectomy and radical nephrectomy. J Rob Surg 2(2):67–69
22. Curry M, Malpani A, Li R, Jog A, Tantillo T, Blanco R, Ha P, Califano J, Kumar R, Richmon J (2012) Objective assessment in residency based training for transoral robotic surgery. The Laryngoscope
23. Kumar R, Hoffman B, Prisco G, Larkin D, Nowlin W et al (2013) Multi-user medical robotic system for collaboration or training in minimally invasive surgical procedures. US patent 8,527,094
24. Bertelsen A, Melo J, Sánchez E, Borro D (2013) A review of surgical robots for spinal interventions. Int J Med Robot Comput Assist Surg 9:407–422
25. Roser F, Tatagiba M, Maier G (2013) Spinal robotics: current applications and future perspectives. Neurosurgery 72(S1):A12–A18
26. Kumar R (2010) Robotic surgery. In: Levine W et al (ed) The control handbook, 2nd edn. CRC Press, Boca Raton, pp 20.1–20.20
27. Faust RA (ed) (2006) Robotics in surgery: history, current and future applications. Nova Science Pub Inc., Hauppauge
28. Peters T, Cleary, K (eds) (2008) Image-guided interventions: technology and applications. Springer, Heidelberg
29. Hagn U, Ortmaier T, konietschke R, Kuebler B, Seibold U, Tobergte A, Stefanjoerg M, Hirzinger G (2008) Telemanipultion for remote minimally invasive surgery. IEEE Robot Autom Mag

30. Hagag B, Abovitz R, Kang H, Schmitz B, Conditt M (2011) RIO: robotic-arm interactive orthopedic system MAKOplasty: user interactive haptic orthopedic robotics. Surg Robot 219–246
31. Hempel E, Fischer H, Gumb L et al (2003) An MRI-compatible surgical robot for precise radiological interventions. Comput Aided Surg 8(4):180–191
32. Ju H, Zhang J, An G, et al (2008) A robot-assisted system for minimally invasive spine surgery of percutaneous vertebroplasty based on CT images. In: IEEE conference on robotics, automation and mechatronics, pp 290–295
33. Blumenkranz S, Larkin D, Kumar R (2010) Modular force sensor. US patent number 7752920
34. Shoham M, Burman M, Zehavi E, Joskowicz L, Batkilin E, Kunicher Y (2003) Bone-mounted miniature robot for surgical procedures: concept and clinical applications. IEEE Trans Robot Autom 19(5):893–901
35. Ponnusamy K, Chewning S, Mohr CJ (2009) Robotic approaches to the posterior spine. Spine 34(19):2104–2109
36. Yang MS, Yoon do H, Kim KN, Kim H, Yang JW, Yi S, Lee JY, Jung WJ, Rha KH, Ha Y (2011) Robot-assisted anterior lumbar interbody fusion in a Swine model in vivo test of the da vinci surgical-assisted spinal surgery system. Spine 36(2):E139–E143
37. Kim MJ, Ha Y, Yang MS et al (2010) Robot-assisted anterior lumbar interbody fusion (ALIF) using retroperitoneal approach. Acta Neurochir (Wien) 152(4):675–679
38. Beutler WJ, Eppelman WC Jr., DiMarco LA (2013) The da Vinci robotic surgical assisted anterior lumbar interbody fusion: technical development and case report. Spine 38(4):356–363
39. Kantelhardt SR, Martinez R, Baerwinkel S, Burger R, Giese A, Rohde V (2011) Perioperative course and accuracy of screw positioning in conventional, open robotic-guided and percutaneous robotic-guided, pedicle screw placement. Eur Spine J 20(6):860–868
40. Kantelhardt SR, Martinez R, Baerwinkel S, Burger R, Giese A, Rohde V (2014) Mazor robotics website. http://mazorrobotics.com/renaissance/clinical-evidence/. Accessed Jan 2014
41. Hu X, Lieberman IH (2013) What is the learning curve for robotic-assisted pedicle screw placement in spine surgery? Clin Orthop Related Res. E-publication available at http://link.springer.com/article/10.1007%2Fs11999-013-3291-1?_escaped_fragment_=#!
42. Sukovich W, Brink-Danan S, Hardenbrook M (2006) Miniature robotic guidance for pedicle screw placement in posterior spinal fusion: early clinical experience with the SpineAssist. Int J Med Robot 2:114–122

Author Index

© Springer International Publishing Switzerland 2015

S. Li and J. Yao (eds.), *Spinal Imaging and Image Analysis*,
Lecture Notes in Computational Vision and Biomechanics 18,
DOI 10.1007/978-3-319-12508-4