Lecture Notes in Computational Vision and Biomechanics 20

Jianhua Yao
Ben Glocker
Tobias Klinder
Shuo Li *Editors*

Recent Advances in Computational Methods and Clinical Applications for Spine Imaging

Springer

Lecture Notes in Computational Vision and Biomechanics

Volume 20

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

Jianhua Yao · Ben Glocker
Tobias Klinder · Shuo Li
Editors

Recent Advances in Computational Methods and Clinical Applications for Spine Imaging

Springer

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

Tobias Klinder
Innovative Technologies
Philips Research
Hamburg
Germany

Ben Glocker
Department of Computing
Imperial College
London
UK

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

ISSN 2212-9391 ISSN 2212-9413 (electronic)
Lecture Notes in Computational Vision and Biomechanics
ISBN 978-3-319-14147-3 ISBN 978-3-319-14148-0 (eBook)
DOI 10.1007/978-3-319-14148-0

Library of Congress Control Number: 2014959524

Springer Cham Heidelberg New York Dordrecht London
© Springer International Publishing Switzerland 2015
Chapters 1 and 23 were created within the capacity of an US governmental employment. US copyright protection does not apply.
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 disc disease, spinal stenosis, scoliosis, osteoporosis, herniated discs, fracture/ligamentous injury, infection, tumor, and spondyloarthropathy. Treatment varies with the disease entity and the clinical scenario can be nonspecific. As a result, imaging is often required to help make the diagnosis. Frequently obtained studies include plain radiographs, DXA, bone scans, CT, MR, ultrasound, and nuclear medicine. Computational methods play a steadily increasing role in improving speed, confidence, and accuracy in reaching a final diagnosis. Although there has been great progress in the development of computational methods for spine imaging over the recent years, there are a number of significant challenges in both methodology and clinical applications.

The goal of this workshop on "Computational Methods and Clinical Applications for Spine Imaging" was to bring together clinicians, computer scientists, and industrial vendors in the field of spine imaging, for reviewing the state-of-the-art techniques, sharing the novel and emerging analysis and visualization techniques, and discussing the clinical challenges and open problems in this rapidly growing field. We invited papers on all major aspects of problems related to spine imaging, including clinical applications of spine imaging, computer-aided diagnosis of spine conditions, computer aided detection of spine-related diseases, emerging computational imaging techniques for spinal diseases, fast 3D reconstruction of spine, feature extraction, multiscale analysis, pattern recognition, image enhancement of spine imaging, image-guided spine intervention and treatment, multimodal image registration and fusion for spine imaging, novel visualization techniques, segmentation techniques for spine imaging, statistical and geometric modeling for spine and vertebra, spine and vertebra localization.

This is the second MICCAI workshop on this particular topic. This year we add a challenge on "Vertebra segmentation on CT." We received many high quality submissions addressing many of the above-mentioned issues. All papers underwent

a thorough double-blinded review with each paper being reviewed by three members of the paper reviewing committee. The program committee consisted of researchers who had actively contributed to the field of spine imaging in the past. From all submissions, we finally accepted 22 papers including 10 oral presentations, 6 poster presentations, and 6 challenge presentations. The papers are organized into five sessions according to the topics. The sessions are Computer Aided Diagnosis and Intervention, Spine MRI Processing, Localization, Poster session, and Vertebra Segmentation Challenge.

In order to give deeper insights into the field and stimulate further ideas, we had invited lectures held during the workshop. We are very thankful to Dr. David R. Haynor from University of Washington to give a talk on "Towards quantitative diagnosis and measurement of chronic spinal diseases: the role of image processing," and Dr. Nassir Navab from Technische Universitaet Muenchen and Johns Hopkins University to give a talk on "Domain Specific Simulation Environments for Accelerating Validation and Deployment of Novel Image-guided Spine Surgery Techniques."

We hope that with this workshop we raised attention toward this important and interesting field of computational spine imaging and would like to finally thank all contributors for their efforts in making this workshop possible. We especially thank the following institutes for their sponsorship: Journal of Computerized Medical Imaging and Graphics, GE Healthcare, Microsoft Research, Digital Imaging group of London, Imperial College London, Philips Research, and National Institutes of Health.

Jianhua Yao
Ben Glocker
Tobias Klinder
Shuo Li

Workshop Organization

Organizing Committee

Jianhua Yao, National Institutes of Health, USA
Ben Glocker, Imperial College London, UK
Tobias Klinder, Philips Research, Germany
Shuo Li, GE Healthcare and University of Western Ontario, Canada

Program Committee

Purang Abolmaesumi, University of British Columbia, Canada
Michel Audette, Old Dominion University, USA
Ulas Bagci, National Institutes of Health, USA
Jonathan Biosvert, National Research Council, Canada
Xinjian Chen, Soochow University, China
Ananda Chowdhury, Jadavpur University, India
Vipin Chaudhary, SUNY Buffalo, USA
Daniel Forsberg, Sectra and Linköping University, Sweden
Samuel Kadoury, Ecole Polytechnique de Montreal, Canada
Cristian Lorenz, Philips Research, Germany
Jianfei Liu, Duke University, USA
Xiaofeng Liu, GE Global Research, USA
Yixun Liu, Broncus Medical, USA
Nikos Paragios, Ecole Centrale de Paris, France
Alberto Santamaria Pang, GE Global Research, USA
Alexander Seitel, University of British Columbia, Canada
Sovira Tan, National Institutes of Health, USA
Tamas Ungi, Queens University, Canada

Tomaž Vrtovec, University of Ljubljana, Slovenia
Zhijie Wang, GE Healthcare, Canada
Yiqiang Zhan, Siemens Medical Solutions, USA
Guoyan Zheng, University of Bern, Switzerland

Scientific Review Committee

Purang Abolmaesumi, University of British Columbia, Canada
Michel Audette, Old Dominion University, USA
Ulas Bagci, National Institutes of Health, USA
Jonathan Biosvert, National Research Council, Canada
Vipin Chaudhary, SUNY Buffalo, USA
Xinjian Chen, Soochow University, China
Ananda Chowdhury, Jadavpur University, India
Daniel Forsberg, Sectra and Linköping University, Sweden
Ben Glocker, Imperial College London, UK
Samuel Kadoury, Ecole Polytechnique de Montreal, Canada
Tobias Klinder, Philips Research, Germany
Jianfei Liu, Duke University, USA
Xiaofeng Liu, GE Global Research, USA
Yixun Liu, Broncus Medical, USA
Cristian Lorenz, Philips Research, Germany
Alberto Pang, GE Global Research, USA
Nikos Paragios, Ecole Centrale de Paris, France
Alexander Seitel, University of British Columbia, Canada
Sovira Tan, National Institutes of Health, USA
Tamas Ungi, Queens University, Canada
Tomaz Vrtovec, University of Ljubljana, Slovenia
Zhijie Wang, GE Healthcare, Canada
Jianhua Yao, National Institutes of Health, USA
Yiqiang Zhan, Siemens Medical Solutions, USA
Guoyan Zheng, University of Bern, Switzerland

Proceedings Editors

Jianhua Yao, National Institutes of Health, USA
Ben Glocker, Imperial College London, UK
Tobias Klinder, Philips Research, Germany
Shuo Li, GE Healthcare and University of Western Ontario, Canada

Contents

Contributors

P. Abolmaesumi Department of Electrical and Computer Engineering, University of British Columbia (UBC), Vancouver, Canada

J.E. Adams Radiology, Manchester Academic Health Science Centre, Central Manchester University Hospitals NHS Foundation Trust, Manchester, UK

Michael Amann University Hospital Basel, Basel, Switzerland

Michaela Andelova University Hospital Basel, Basel, Switzerland

Rifat Aras Old Dominion University, Norfolk, VA, USA

Michel A. Audette Department of Modeling, Simulation and Visualization Engineering, Old Dominion University, Norfolk, VA, USA

David A. Besachio Naval Medical Center Portsmouth, Portsmouth, VA, USA

Roderick C. Borgie Naval Medical Center Portsmouth, Portsmouth, VA, USA

P.A. Bromiley Imaging Sciences Research Group, University of Manchester, Manchester, UK

Katja Bühler VRVis Center for Virtual Reality and Visualization, Vienna, Austria

Joseph E. Burns Department of Radiological Sciences, University of California, Irvine, Orange, CA, USA

Alvin Cabrera Department of Radiation Oncology, Duke University Medical Centre, Durham, NC, USA

Guillermo Carbajal Instituto de Ingeniería Eléctrica Facultad de Ingeniería, Universidad de la República, Montevideo, Uruguay

Isaac Castro-Mateos Department of Mechanical Engineering, Centre for Computational Imaging and Simulation Technologies in Biomedicine (CISTIB), University of Sheffield, Sheffield, UK

Philippe C. Cattin Department of Biomedical Engineering, University of Basel, Basel, Switzerland

T.F. Cootes Imaging Sciences Research Group, University of Manchester, Manchester, UK

Xiao Dong School of Computer Science and Engineering, Southeast University, Nanjing, China

Thomas Ebner Institute for Computer Graphics and Vision, BioTechMed, Graz University of Technology, Graz, Austria

Richard Everson University of Exeter, Exeter, UK

Gabor Fichtinger Laboratory for Percutaneous Surgery, School of Computing, Queen's University, Kingston, ON, Canada

Daniel Forsberg Center for Medical Image Science and Visualization (CMIV), Linköping University, Linköping, Sweden; Sectra, Linköping, Sweden

Alejandro Frangi Department of Mechanical Engineering, Centre for Computational Imaging and Simulation Technologies in Biomedicine (CISTIB), University of Sheffield, Sheffield, UK

Takaichi Fukuda Graduate School of Medical Sciences, Kumamoto University, Kumamoto, Japan

Ketut Fundana Department of Biomedical Engineering, University of Basel, Basel, Switzerland

Yaorong Ge Department of Software and Information Systems, University of North Carolina at Charlotte, Charlotte, NC, USA

Álvaro Gómez Instituto de Ingeniería Eléctrica Facultad de Ingeniería, Universidad de la República, Montevideo, Uruguay

Kerstin Hammernik Institute for Computer Graphics and Vision, BioTechMed, Graz University of Technology, Graz, Austria

Rabia Haq Old Dominion University, Norfolk, VA, USA

Jiří Hladůvka VRVis Center for Virtual Reality and Visualization, Vienna, Austria

Corné Hoogendoorn CISTIB, Universitat Pompeu Fabra, Barcelona, Spain

Hugo Hutt University of Exeter, Exeter, UK

Bulat Ibragimov Faculty of Electrical Engineering, University of Ljubljana, Ljubljana, Slovenia

Q. Jackie Wu Department of Radiation Oncology, Duke University Medical Centre, Durham, NC, USA

Timor Kadir Mirada Medical, Oxford, UK

Samuel Kadoury CHU Sainte-Justine Research Center, MEDICAL, École Polytechnique de Montréal, Montréal, Canada

John P. Kirkpatrick Department of Radiation Oncology, Duke University Medical Centre, Durham, NC, USA

Robert Korez Faculty of Electrical Engineering, University of Ljubljana, Ljubljana, Slovenia

Ryo Kurazume Graduate School of Information Science and Electrical Engineering, Kyushu University, Fukuoka, Japan

Hubert Labelle Department of Surgery, CHU Sainte-Justine Research Center, Université de Montréal, Montréal, Canada

Aron Lazary National Center for Spine Disorder (NCSD), Budapest, Hungary

Karim Lekadir CISTIB, Universitat Pompeu Fabra, Barcelona, Spain

Shuo Li GE Healthcare, Mississauga, ON, Canada; University of Western Ontario, London, ON, Canada

Boštjan Likar Faculty of Electrical Engineering, University of Ljubljana, Ljubljana, Slovenia

Jianfei Liu Department of Radiation Oncology, Duke University Medical Centre, Durham, NC, USA

Meelis Lootus Oxford University, Oxford, UK

Le Lu Imaging Biomarkers and Computer-Aided Diagnosis Laboratory, Radiology and Imaging Sciences, National Institutes of Health Clinical Center, Bethesda, MD, USA

David Major VRVis Center for Virtual Reality and Visualization, Vienna, Austria

Judith Meakin University of Exeter, Exeter, UK

Yasushi Miyagi Department of Neurosurgery, Kaizuka Hospital, Fukuoka, Japan

Shoko Miyauchi Graduate School of Information Science and Electrical Engineering, Kyushu University, Fukuoka, Japan

Ken'ichi Morooka Graduate School of Information Science and Electrical Engineering, Kyushu University, Fukuoka, Japan

Stefan Parent Department of Surgery, CHU Sainte-Justine Research Center, Université de Montréal, Montréal, Canada

Nirmal Patel Department of Electrical and Computer Engineering, Old Dominion University, Norfolk, USA

Marco Pereañez Center for Computational Imaging and Simulation Technologies in Biomedicine (CISTIB), Universitat Pompeu Fabra, Barcelona, Spain

Franjo Pernuš Faculty of Electrical Engineering, University of Ljubljana, Ljubljana, Slovenia

Simon Pezold Department of Biomedical Engineering, University of Basel, Basel, Switzerland

Armanda Pfister University Hospital Basel, Basel, Switzerland

Thomas Pock Institute for Computer Graphics and Vision, BioTechMed, Graz University of Technology, Graz, Austria

Jose M. Pozo Centre for Computational Imaging and Simulation Technologies in Biomedicine (CISTIB), Department of Mechanical Engineering, University of Sheffield, Sheffield, UK

A. Rasoulian Department of Electrical and Computer Engineering, University of British Columbia (UBC), Vancouver, Canada

R. Rohling Department of Mechanical Engineering, UBC, Vancouver, Canada

Holger R. Roth Imaging Biomarkers and Computer-Aided Diagnosis Laboratory, Radiology and Imaging Sciences, National Institutes of Health Clinical Center, Bethesda, MD, USA

A. Seitel Department of Electrical and Computer Engineering, University of British Columbia (UBC), Vancouver, Canada

Jesse Shen Department of Surgery, CHU Sainte-Justine Research Center, Université de Montréal, Montréal, Canada

Till Sprenger University Hospital Basel, Basel, Switzerland

Darko Stern Institute for Computer Graphics and Vision, BioTechMed, Graz University of Technology, Graz, Austria

James Stieger Imaging Biomarkers and Computer-Aided Diagnosis Laboratory, Radiology and Imaging Sciences, National Institutes of Health Clinical Center, Bethesda, MD, USA

Sharmin Sultana Department of Modeling, Simulation and Visualization Engineering, Old Dominion University, Norfolk, USA

Ronald M. Summers Imaging Biomarkers and Computer-Aided Diagnosis Laboratory, Radiology and Imaging Sciences, National Institutes of Health Clinical Center, Bethesda, MD, USA

William E. Thong MEDICAL, École Polytechnique de Montréal, Montréal, Canada

Tokuo Tsuji Graduate School of Information Science and Electrical Engineering, Kyushu University, Fukuoka, Japan

Tamas Ungi Laboratory for Percutaneous Surgery, School of Computing, Queen's University, Kingston, ON, Canada

Martin Urschler Ludwig Boltzmann Institute for Clinical Forensic Imaging, Graz, Austria

Tomaž Vrtovec Faculty of Electrical Engineering, University of Ljubljana, Ljubljana, Slovenia

Jianhua Yao Imaging Biomarkers and Computer-Aided Detection Laboratory, Radiology and Imaging Sciences, National Institutes of Health Clinical Center, Bethesda, MD, USA

Fang-Fang Yin Department of Radiation Oncology, Duke University Medical Centre, Durham, NC, USA

Guoyan Zheng Institute for Surgical Technology and Biomechanics, University of Bern, Bern, Switzerland

Andrew Zisserman Oxford University, Oxford, UK

Part I
Computer Aided Diagnosis and Intervention

Detection of Sclerotic Spine Metastases via Random Aggregation of Deep Convolutional Neural Network Classifications

Holger R. Roth, Jianhua Yao, Le Lu, James Stieger, Joseph E. Burns and Ronald M. Summers

Abstract Automated detection of sclerotic metastases (bone lesions) in Computed Tomography (CT) images has potential to be an important tool in clinical practice and research. State-of-the-art methods show performance of 79 % sensitivity or true-positive (TP) rate, at 10 false-positives (FP) per volume. We design a two-tiered coarse-to-fine cascade framework to first operate a highly sensitive candidate generation system at a maximum sensitivity of ∼92 % but with high FP level (∼50 per patient). Regions of interest (ROI) for lesion candidates are generated in this step and function as input for the second tier. In the second tier we generate N 2D views, via scale, random translations, and rotations with respect to each ROI centroid coordinates. These random views are used to train a deep Convolutional Neural Network (CNN) classifier. In testing, the CNN is employed to assign individual probabilities for a new set of N random views that are averaged at each ROI to compute a final per-candidate classification probability. This second tier behaves as a highly selective process to reject difficult false positives while preserving high sensitivities. We validate the approach on CT images of 59 patients (49 with sclerotic metastases and 10 normal controls). The proposed method reduces the number of FP/vol. from 4 to 1.2, 7 to 3, and 12 to 9.5 when comparing a sensitivity rates of 60, 70, and 80 % respectively in testing. The Area-Under-the-Curve (AUC) is 0.834. The results show marked improvement upon previous work.

H.R. Roth (✉) · J. Yao · L. Lu · J. Stieger · R.M. Summers
Imaging Biomarkers and Computer-Aided Diagnosis Laboratory, Radiology and Imaging Sciences, National Institutes of Health Clinical Center, Bethesda, MD 20892-1182, USA
e-mail: holger.roth@nih.gov
J. Yao
e-mail: jyao@cc.nih.gov
L. Lu
e-mail: le.lu@nih.gov
J. Stieger
e-mail: james.stieger@nih.gov
R.M. Summers
e-mail: rms@nih.gov

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

© Springer International Publishing Switzerland (outside the USA) 2015

J. Yao et al. (eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, Lecture Notes in Computational Vision and Biomechanics 20, DOI 10.1007/978-3-319-14148-0_1

1 Introduction

Early detection of sclerotic bone metastases plays an important role in clinical practice. Their detection can assess the staging of the patient's disease, and therefore has the potential to alter the treatment regimen the patient will undergo [1]. Approximately 490,000 patients per year are affected by metastatic diseases of the skeletal structures in the United States alone [2]. More than 80 % of these bone metastases are thought to originate from breast and prostate cancer [3]. As a ubiquitous screening and staging modality employed for disease detection in cancer patients, Computed Tomography (CT) is commonly involved in the detection of bone metastases. Both lytic and sclerotic metastatic diseases change or deteriorate the bone structure and bio-mechanically weaken the skeleton. Sclerotic metastases grow into irregularly mineralized and disorganized "woven" bone [4–7]. Typical examples of sclerotic metastases are shown in Fig. 1. The detection of sclerotic metastases often occurs during manual prospective visual inspection of every image (of which there may be thousands) and every section of every image in each patient's CT study. This is a complex process that is performed under time restriction and which is prone to error. Furthermore, thorough manual assessment and processing is time-consuming and has potential to delay the clinical workflow. Computer-Aided Detection (CADe) of sclerotic metastases has the potential to greatly reduce the radiologists' clinical workload and could be employed as a second reader for improved assessment of disease [8–10].

The CADe method presented here aims to build upon an existing system for sclerotic metastases detection and focuses on reducing the false-positive (FP) number of its outputs. We make use of recent advances in computer vision, in particular deep Convolutional Neural Networks (CNNs), to attain this goal. Recently, the availability of large annotated training sets and the accessibility of affordable parallel computing resources via GPUs has made it feasible to train "deep" CNNs (also popularized under the keyword: "deep learning") for computer vision classification tasks. Great advances in classification of natural images have been achieved [11, 12]. Studies that have tried to apply deep learning and CNNs to medical imaging applications also showed promise, e.g. [13, 14]. In particular, CNNs have been applied successfully in biomedical applications such as digital pathology [15]. In this work, we apply CNNs for the reduction of FPs using random sets of 2D CNN observations. Our motivation is partially inspired by the spirit of hybrid systems using both parametric and non-parametric models for hierarchical coarse-to-fine classification [16].

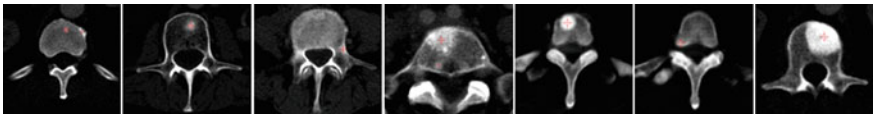

Fig. 1 Examples of sclerotic metastases as detected by the CADe candidate generation step (*red mark*) (Color figure online)

2 Methods

2.1 Sclerotic Metastases Candidate Detection

We use a state-of-the-art CADe method for detecting sclerotic metastases candidates from CT volumes [9, 17]. The spine is initially segmented by thresholding at certain attenuation levels and performing region growing. Furthermore, morphological operations are used to refine the segmentation and allow the extraction of the spinal canal. For further information on the segmentation refer to [18]. Axial 2D cross sections of the spinal vertebrae are then divided into sub-segments using a watershed algorithm based on local density differences [19]. The CADe algorithm then finds initial detections that have a higher mean attenuation then neighboring 2D sub-segments. Because the watershed algorithm can cause over-segmentation of the image, similar 2D sub-segments detections are merged by performing an energy minimization based on graph cuts and attenuation thresholds. Finally, 2D detections on neighboring cross sections are combined to form 3D detections using a graph-cut-based merger. Each 3D detection acts as a seed point for a level-set segmentation method that segments the lesions in 3D. This step allows the computation of 25 characteristic features, including shape, size, location, attenuation, volume, and sphericity. A committee of SVMs [20] is then trained on these features. The trained SVMs further classify each 3D detection as 'true' or 'false' bone lesion. Example of bone lesions candidates using this detection scheme are shown in Fig. 1. Next, true bone lesions from this step are used as candidate lesions for a second classification based on CNNs as proposed in this paper. This is a coarse-to-fine classification approach somewhat similar to other CADe schemes such as [16].

2.2 CNN Training on 2D Image Patches

A Region-of-Interest (ROI) in a CT image is extracted at each bone lesion candidate location (see Fig. 2). In order to increase the variation of the training data and to avoid overfitting analogous to the data augmentation approach in [11], each ROI is translated along a random vector v in axial space. Furthermore, each translated ROI is rotated around its center N_r times by a random angle $\alpha = [0°, \ldots, 360°]$. These translations and rotations for each ROI are computed N_s times at different physical scales s (the edge length of each ROI), but with fixed numbers of pixels. This procedure results in $N = N_s \times N_t \times N_r$ random observation of each ROI—an approach similar to [21]. Note that 2.5–5 mm thick-sliced CT volumes are used for this study. Due to this relative large slice thickness, our spatial transformations are all drawn from within the axial plane. This is in contrast to other approaches that use CNNs which sample also sagittal and/or coronal planes [13, 14]. Following this procedure, both the training and test data can be easily expanded to better scale to this

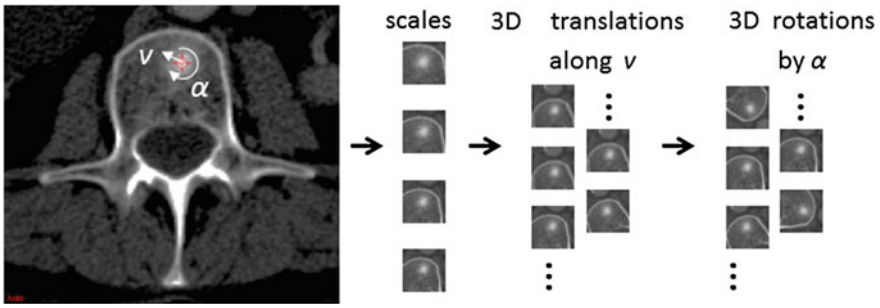

Fig. 2 Image patches are generated from CADe candidates using different scales, 2D translations (along a random vector v) and rotations (by a random angle α) in the axial plane

type of neural net application. A CNN's predictions on these N random observations $\{P_1(x), \ldots, P_N\}$ can then be simply averaged at each ROI to compute a per-candidate probability:

$$p\,(x|\{P_1(x), \ldots, P_N(x)\}) = \frac{1}{N}\sum_{i=1}^{N} P_i(x). \tag{1}$$

Here, $P_i(x)$ is the CNN's classification probability computed one individual image patch. In theory, more sophisticated fusion rules can be explored but we find that simple averaging works well. This proposed random resampling is an approach to effectively and efficiently increase the amount of available training data. In computer vision, translational shifting and mirroring of 2D image patches is often used for this purpose [11]. By averaging the N predictions on random 2D views as in Eq. 1, the robustness and stability of CNN can be further increased as shown in Sect. 3.

2.3 CNN Architecture

A CNN derives its name from the convolutional filters that it applies to the input images. Typically, several layers of convolutional filters are cascaded to compute image features. Other layers of a CNN often perform max-pooling operations or consist of fully-connected neural networks. Our CNN ends with a final 2-way softmax layer for 'true' and 'false' classification (see Fig. 3). The fully connected layers are typically constrained in order to avoid overfitting. We use "DropConnect" for this purpose. "DropConnect" is a method that behaves as a regularizer when training the CNN [22]. It can be seen as a variation of the earlier developed "DropOut" method [23]. GPU acceleration allows efficient training of the CNN. We use an open-source implementation by Krizhevsky et al. [11] with the DropConnect extension by Wan et al. [22]. Further execution speed-up for both training and evaluation is achieved by using rectified linear units as the neuron model instead of the traditional

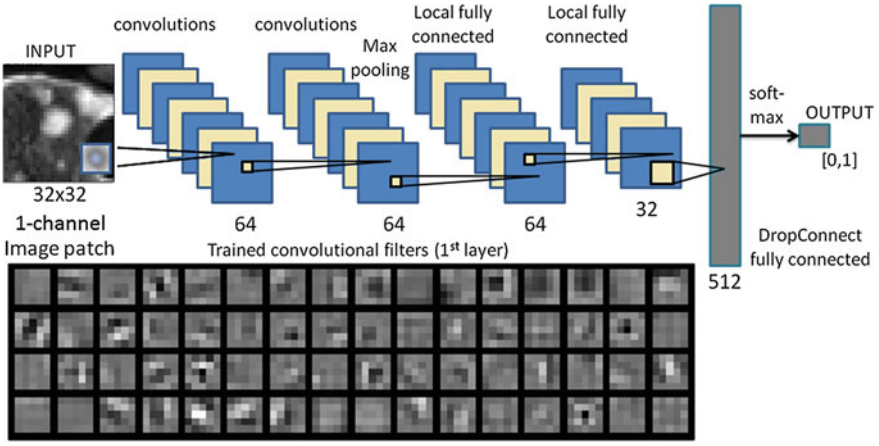

Fig. 3 The proposed convolution neural network consists of two convolutional layers, max-pooling layers, locally fully-connected layers, a DropConnect layer, and a final 2-way softmax layer for classification. The number of filters, connections for each layer, and the first layer of learned convolutional kernels are shown

neuron model $f(x) = \tanh(x)$ or $f(x) = (1 + e^{-x})^{-1}$ [11]. At the moment, it is still difficult to design a theoretically optimal CNN architecture for a particular image classification task [12]. We evaluate several CNNs with slightly different layer architectures (independently to the later evaluations) in order to find a suitable CNN architecture for our classification task, using a small number of CT cases within the training data subset. We find relatively stable behavior over model variations and hence fix the CNN architecture for subsequent experiments performed in this study. Recently, approaches have been proposed that aim to visualize the feature activations of CNNs in order to allow better CNN design [12]. Potentially, these approaches allow better understanding of how CNNs behave at a given task. This could lead to improved CNN architecture design compared to the heuristic approach applied in this work.

3 Evaluation and Results on Sclerotic Metastases

In our evaluation, radiologists label a total of 532 sclerotic metastases ('positives') in CT images of 49 patients (14 female, 35 male patients; mean age, 57.0 years; range, 12–77 years). A lesion is only labeled if its volume is greater than 300 mm^3. The CT scans have reconstruction slice thicknesses ranging between 2.5 mm and 5 mm. Furthermore, we include 10 control cases (4 female, 6 male patients; mean age, 55.2 years; range, 1970 years) without any spinal lesions.

Any false-positive detections from the candidate generation step on these patients are used as 'negative' candidate examples for training the CNN. All patients were

randomly split into five sets at the patient level in order to allow a 5-fold cross-validation. We adjust the sample rates for positive and negative image patches in order to generate a balanced data set for training (50 % positives and 50 % negatives). This proves to be beneficial for training the CNNs—no balancing was done during testing. Each three-channel image patch was centered at the CADe coordinate with 32×32 pixels in resolution. All patches were sampled at 4 scales of $s = [30, 35, 40, 45]$ mm ROI edge length in physical image space, after iso-metric resampling of the CT image (see Fig. 2). We used a bone window level of $[-250, 1250$ HU$]$. Furthermore, all ROIs were randomly translated (up to 3 mm) and rotated at each scale ($N_s = 4$, $N_t = 5$ and $N_r = 5$), resulting in $N = 100$ image patches per ROI. Training each CNN model took 12–15 hours on a NVIDIA GeForce GTX TITAN, while running 100 2D image patches at each ROI for classification of one CT volume only took circa 30 s. Image patch extraction from one CT volume took around 2 min on each scale.

We now apply the trained CNN to classify image patches from the test data sets. Figures 4 and 5 show typical classification probabilities on random subsets of positive and negative ROIs in the test case. Averaging the N predictions at each sclerotic metastases candidate allows us to compute a per-candidate probability $p(x)$, as in Eq. 1. Varying thresholds on probability $p(x)$ are used to compute Free-Response Receiver Operating Characteristic (FROC) curves. The FROC curves are compared in Fig. 6 for varying amounts of N. It can be seen that the classification performance saturates quickly with increasing N. This means the run-time efficiency of our second

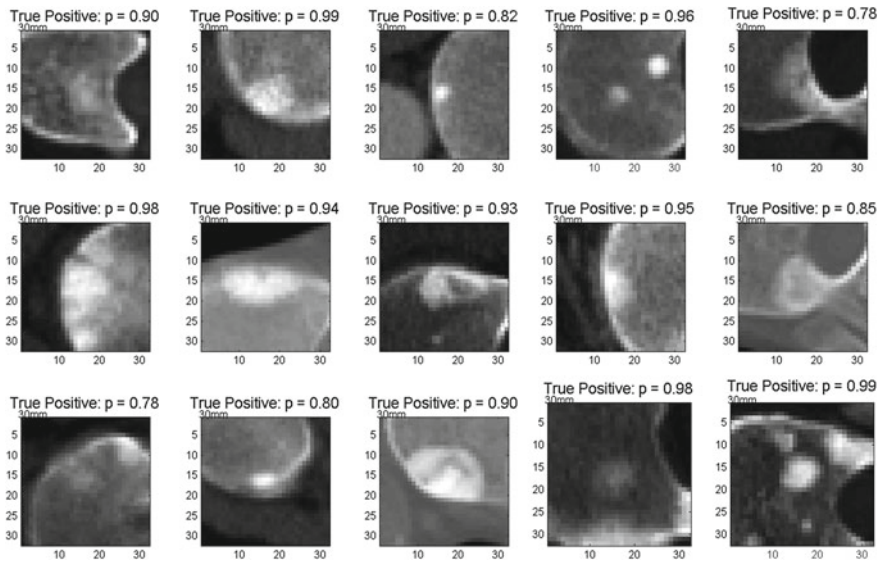

Fig. 4 Test probabilities of the CNN for being sclerotic metastases on 'true' sclerotic metastases candidate examples (close to 1.0 is good)

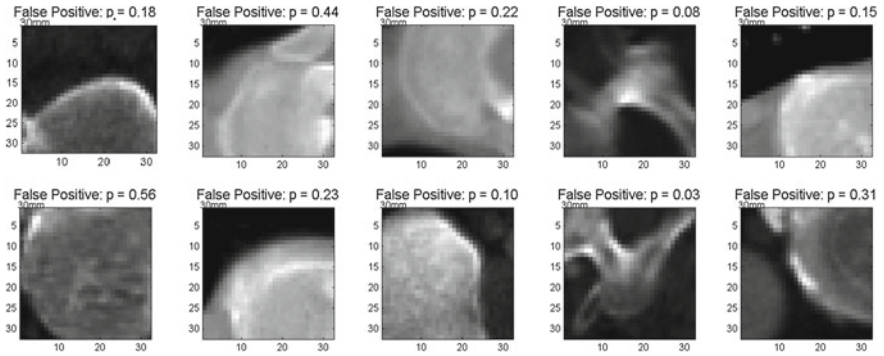

Fig. 5 Test probabilities of the CNN for being sclerotic metastases on 'false' sclerotic metastases candidate examples (close to 0.0 is good)

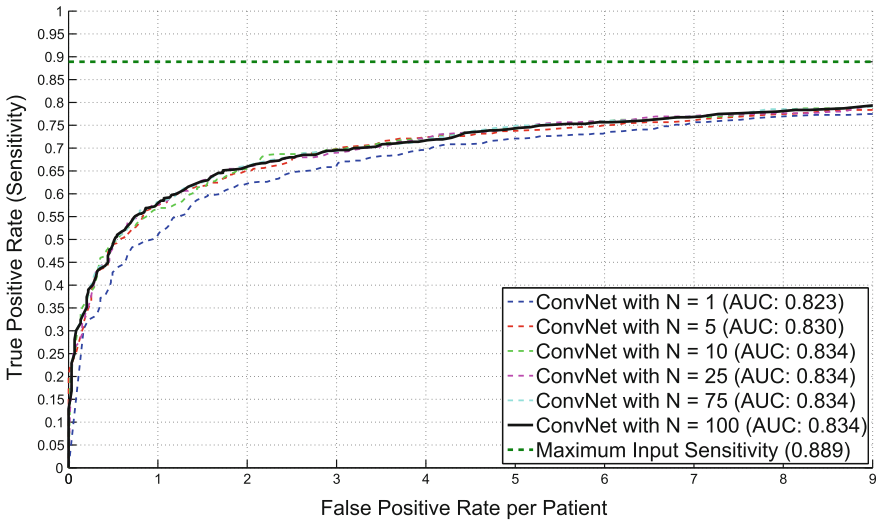

Fig. 6 Free-Response Receiver Operating Characteristic (FROC) curves for a 5-fold cross-validation using a varying number of N random CNN observers in 59 patients (49 with sclerotic metastases and 10 normal controls). AUC values are computed for corresponding ROC curves

layer detection could be further improved without losing noticeable performance by decreasing N. The proposed method reduces the number of FP/vol. of the existing sclerotic metastases CADe systems [9] from 4 to 1.2, 7 to 3, and 12 to 9.5 when comparing a sensitivity rates of 60, 70, and 80 % respectively in cross-validation testing (at $N = 100$). This has the potential to greatly reduce radiologists' clinical workload when employing the proposed CADe system as a second reader. The Area-Under-the-Curve (AUC) shows a value of 0.834 at this number of N.

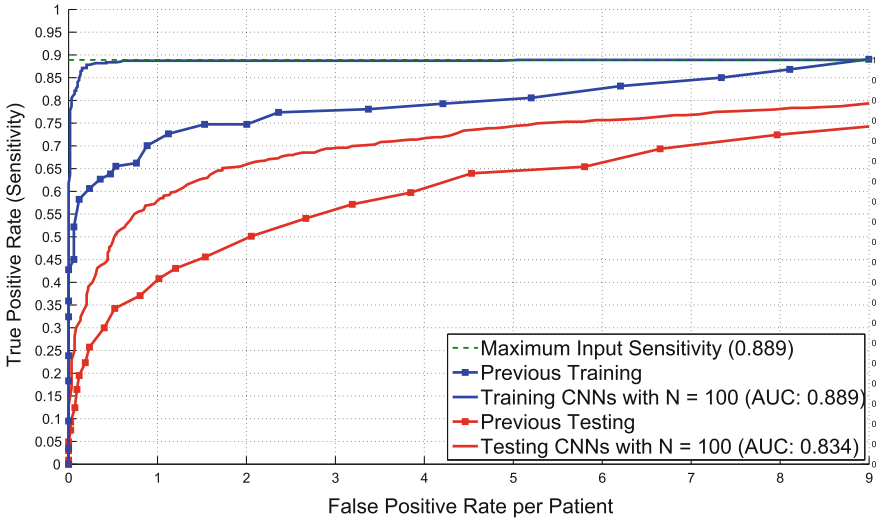

Fig. 7 Comparison of Free-Response Receiver Operating Characteristic (FROC) curves of the first layer bone lesion candidate generation (*squares*) and the second layer classification using $N = 100$ CNN observers (*lines*) for both training and testing cases. Result are computed using a 5-fold cross-validation in 59 patients (49 with sclerotic metastases and 10 normal controls)

Figure 7 compares the FROCs from the initial (first layer) CADe system [9] and illustrates the progression towards the proposed coarse-to-fine two tiered method in both training and testing datasets. This clearly demonstrates a marked improvement in performance.

4 Discussion and Conclusions

This work demonstrates that deep CNNs can be generalized to tasks in medical image analysis, such as effective FP reduction in Computer-aided Detection (CADe) systems. This is especially true, since the main drawback of current CADe developments often generates too many false positive detections at clinically relevant sensitivity levels. We show that a random set of CNN classifications can be used to reduce FPs when operating an existing method for CADe of sclerotic metastases (bone lesions) at a particular point its FROC curve. Different scales, random translations, and rotations around each of the CADe detections can be utilized to increase the CNN's classification performance. The FROC curves show a marked reduction of the FP/vol. rate at clinically useful levels of sensitivity. These results improve upon the state-of-the-art.

The average of CNN classification probabilities was chosen in this work for simplicity (see Eq. 1), but this approach shows to be very efficient and effective.

Future work will investigate more sophisticated methods of label fusion from the CNNs. A similar 2.5D generalization of CNNs also shows promise in the detection of lymph nodes in CT images (see [14]). In this work, we decide against a 2.5D or full 3D approach due to the relative large slice thicknesses of ∼5 mm in the used CT data. This prevents reformatting the data in sufficient detail in any other than the axial plane. However, the improvements achieved in this study and other methods utilizing CNNs in medical image computing show promise for a variety of applications in computer-aided detection of 2D and 3D medical images. Our mainly 2D approach may adapt and generalize particularly well to the current trend of low-dose, low-resolution (slice thickness) CT imaging protocols, compared to direct 3D based methods that require volumetric medical images of higher resolution.

Acknowledgments This work was supported by the Intramural Research Program of the NIH Clinical Center.

References

1. Msaouel, P., Pissimissis, N., Halapas, A., Koutsilieris, M.: Mechanisms of bone metastasis in prostate cancer: clinical implications. Best Pract. Res. Clin. Endocrinol. Metabol. **22**(2), 341–355 (2008)
2. Hitron, A., Adams, V.: The pharmacological management of skeletal-related events from metastatic tumors. Orthopedics **32**(3), 188 (2009)
3. Coleman, R.: Metastatic bone disease: clinical features, pathophysiology and treatment strategies. Cancer Treat. Rev. **27**(3), 165–176 (2001)
4. Saylor, P., Smith, M.: Bone health and prostate cancer. Prostate Cancer Prostatic Dis. **13**(1), 20–27 (2010)
5. Keller, E.T., Brown, J.: Prostate cancer bone metastases promote both osteolytic and osteoblastic activity. J. Cell. Biochem. **91**(4), 718–729 (2004)
6. Lee, R.J., Saylor, P.J., Smith, M.R.: Treatment and prevention of bone complications from prostate cancer. Bone **48**(1), 88–95 (2011)
7. Guise, T.A., Mundy, G.R.: Cancer and bone 1. Endocr. Rev. **19**(1), 18–54 (1998)
8. Wiese, T., Yao, J., Burns, J.E., Summers, R.M.: Detection of sclerotic bone metastases in the spine using watershed algorithm and graph cut. In: SPIE Medical Imaging, International Society for Optics and Photonics (2012) 831512–831512
9. Burns, J.E., Yao, J., Wiese, T.S., Muñoz, H.E., Jones, E.C., Summers, R.M.: Automated detection of sclerotic metastases in the thoracolumbar spine at ct. Radiology **268**(1), 69–78 (2013)
10. Hammon, M., Dankerl, P., Tsymbal, A., Wels, M., Kelm, M., May, M., Suehling, M., Uder, M., Cavallaro, A.: Automatic detection of lytic and blastic thoracolumbar spine metastases on computed tomography. Eur. Radiol. **23**(7), 1862–1870 (2013)
11. Krizhevsky, A., Sutskever, I., Hinton, G.: Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems 25 (2012)
12. Zeiler, M.D., Fergus, R.: Visualizing and understanding convolutional neural networks. arXiv:1311.2901 (2013)
13. Prasoon, A., Petersen, K., Igel, C., Lauze, F., Dam, E., Nielsen, M.: Deep feature learning for knee cartilage segmentation using a triplanar convolutional neural network. MICCAI (2013)
14. Roth, H.R., Lu, L., Seff, A., Cherry, K.M., Hoffman, J., Wang, S., Liu, J., Turkbey, E., Summers, R.M.: A new 2.5D representation for lymph node detection using random sets of deep convolutional neural network observations. MICCAI 2014 (in-print), arXiv:1406.2639 [cs.CV] (2014)

15. Cireşan, D.C., Giusti, A., Gambardella, L.M., Schmidhuber, J.: Mitosis detection in breast cancer histology images with deep neural networks. MICCAI (2013)
16. Lu, L., Liu, M., Ye, X., Yu, S., Huang, H.: Coarse-to-fine classification via parametric and nonparametric models for computer-aided diagnosis. In: Proceedings of the 20th ACM international conference on Information and knowledge management, ACM (2011) 2509–2512
17. Wiese, T., Burns, J., Yao, J., Summers, R.M.: Computer-aided detection of sclerotic bone metastases in the spine using watershed algorithm and support vector machines. In: Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on, IEEE (2011) 152–155
18. Yao, J., O'Connor, S.D., Summers, R.M.: Automated spinal column extraction and partitioning. In: Biomedical Imaging: Nano to Macro, 2006. 3rd IEEE International Symposium on, IEEE (2006) 390–393
19. Yao, J., O'Connor, S.D., Summers, R.: Computer aided lytic bone metastasis detection using regular ct images. In: Medical Imaging, International Society for Optics and Photonics (2006) 614459–614459
20. Yao, J., Summers, R.M., Hara, A.K.: Optimizing the support vector machines (svm) committee configuration in a colonic polyp cad system. In: Medical Imaging, International Society for Optics and Photonics (2005) 384–392
21. Göktürk, S.B., Tomasi, C., Acar, B., Beaulieu, C.F., Paik, D.S., Jeffrey, R.B., Yee, J., Napel, Y.: A statistical 3-d pattern processing method for computer-aided detection of polyps in ct colonography. IEEE Trans. Med. Imag. **20**, 1251–1260 (2001)
22. Wan, L., Zeiler, M., Zhang, S., Cun, Y.L., Fergus, R.: Regularization of neural networks using dropconnect. In: Proceedings of International Conference on Machine Learning (ICML-13) (2013)
23. Hinton, G.E., Srivastava, N., Krizhevsky, A., Sutskever, I., Salakhutdinov, R.R.: Improving neural networks by preventing co-adaptation of feature detectors. arXiv preprint arXiv:1207.0580 (2012)

Stacked Auto-encoders for Classification of 3D Spine Models in Adolescent Idiopathic Scoliosis

William E. Thong, Hubert Labelle, Jesse Shen, Stefan Parent and Samuel Kadoury

Abstract Current classification systems for adolescent idiopathic scoliosis lack information on how the spine is deformed in three dimensions (3D), which can mislead further treatment recommendations. We propose an approach to address this issue by a deep learning method for the classification of 3D spine reconstructions of patients. A low-dimensional manifold representation of the spine models was learnt by stacked auto-encoders. A K-Means++ algorithm using a probabilistic seeding method clustered the low-dimensional codes to discover sub-groups in the studied population. We evaluated the method with a case series analysis of 155 patients with Lenke Type-1 thoracic spinal deformations recruited at our institution. The clustering algorithm proposed 5 sub-groups from the thoracic population, yielding statistically significant differences in clinical geometric indices between all clusters. These results demonstrate the presence of 3D variability within a pre-defined 2D group of spinal deformities.

Supported by the CHU Sainte-Justine Academic Research Chair in Spinal Deformities, the Canada Research Chair in Medical Imaging and Assisted Interventions, the 3D committee of the Scoliosis Research Society, the Natural Sciences and Engineering Research Council of Canada and the MEDITIS program.

W.E. Thong
MEDICAL, École Polytechnique de Montréal, Montréal, Canada
e-mail: william.thong@polymtl.ca

H. Labelle · J. Shen · S. Parent
Department of Surgery, CHU Sainte-Justine Research Center,
Université de Montréal, Montréal, Canada
e-mail: hubert.labelle@umontreal.ca

J. Shen
e-mail: jesse.shen@umontreal.ca

S. Parent
e-mail: stefan.parent@umontreal.ca

S. Kadoury (✉)
CHU Sainte-Justine Research Center, MEDICAL,
École Polytechnique de Montréal, Montréal, Canada
e-mail: samuel.kadoury@polymtl.ca

© Springer International Publishing Switzerland 2015

J. Yao et al. (eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, Lecture Notes in Computational Vision and Biomechanics 20, DOI 10.1007/978-3-319-14148-0_2

1 Introduction

Adolescent idiopathic scoliosis (AIS) refers to a complex deformation in three dimensions (3D) of the spine. Classification of the rich and complex variability of spinal deformities is critical for comparisons between treatments and for long-term patient follow-ups. AIS characterization and treatment recommendations currently rely on the Lenke classification system [10] because of its simplicity and its high inter- and intra-observer reliability compared with previous classification systems [8]. However, these schemes are restricted to a two-dimensional (2D) assessment of scoliosis from radiographs of the spine in the coronal and sagittal planes. Misinterpretations could arise because two different scoliosis deformities may have similar 2D parameters. Therefore, improvements in the scoliosis classification system are necessary to ensure a better understanding and description of the curve morphology.

Computational methods open up new paths to go beyond the Lenke classification. Recent studies seek new groups in the population of AIS using cluster analysis [4, 5, 15, 17] with ISOData, K-Means or fuzzy C-Means algorithms. Their common aspect is founded upon the clustering of expert-based features, which are extracted from 3D spine reconstructions (Cobb angles, kyphosis and planes of maximal deformity). This methodology stems from the fact that clustering algorithms are very sensitive to the curse of dimensionality. Still, these parameters might not be enough to tap all the rich and complex variability in the data. Computational methods should be able to capture the intrinsic dimensionality that explains as much as possible the highly dimensional data into a manifold of much lower dimensionality [2, 11]. Hence, another paradigm for spine classification is to let the algorithm learn its own features to discriminate between different pathological groups. This implies directly analyzing the 3D spine models instead of expert-based features as it has been experimented previously. To our knowledge, only one study tried to learn a manifold from the 3D spine model [7] using Local Linear Embeddings (LLE). In this study, we propose to use stacked auto-encoders—a deep learning algorithm—to reduce the high-dimensionality of 3D spine models in order to identify particular classes within Lenke Type-1 curves. This algorithm was able to outperform principal component analysis (PCA) and LLE [6, 11].

Recent breakthroughs in computer vision and speech processing using deep learning algorithms suggest that artificial neural networks might be better suited to learn representations of highly non-linear data [2]. Training a deep neural network has been a challenging task in many applications. Nowadays, this issue is tackled by leveraging more computation power (i.e. parallelizing tasks), more data and by a better initialization of the multilayer neural network [6]. Deep neural networks promote the learning of more abstract representations that result in improved generalization. Network depth actually helps to become invariant to most local changes in the data and to disentangle the main factor of variations in the data [2].

We propose a computational method for the classification of highly dimensional 3D spine models obtained from a group of patients with AIS. The core of the methodology, detailed in Sect. 2, builds a low-dimensional representation of the 3D

spine model based on stacked auto-encoders that capture the main variabilities in the shape of the spine. The low-dimensional codes learnt by the stacked auto-encoders are then clustered using the K-Means++ algorithm. Finally, a statistical analysis assesses the relevance of the clusters identified by the framework based on clinical geometrical indices of the spine. Experiments conducted with this methodology are shown and discussed in Sect. 3, while Sect. 4 concludes this paper.

2 Methods

The proposed method consists of four main steps: (1) reconstruction of a 3D spine model from biplanar X-rays for each patient; (2) dimensionality reduction of each high-dimensional model to a low-dimensional space; (3) clustering of the low-dimensional space; (4) analysis of the clusters obtained with the clinical data.

2.1 3D Spine Reconstruction

A 3D model for each patient's spine was generated from anatomical landmarks with a semi-supervised statistical image-based technique built in a custom software in C++ [13]. Seventeen 3D anatomical landmarks were extracted per vertebra (12 thoracic, 5 lumbar): center, left, right, anterior and posterior of superior and inferior vertebral endplates (10 landmarks); left and right transverse process (2 landmarks); spinous process (1 landmark); and tips of both pedicles (4 landmarks). All 3D spine models were normalized with regards to their height and rigidly translated to a common referential at the L5 vertebra. Hence, each observation contains 867 features, which corresponds to the concatenation of the 3D coordinates of all the landmarks into an observation vector.

2.2 Stacked Auto-encoders for Dimensionality Reduction

An auto-encoder is a neural network that learns a hidden representation to reconstruct its input. Consider a one hidden layer auto-encoder network. First, the input vector $\mathbf{x}$ of dimension d, representing the 3D coordinates of a spine model, is mapped by an encoder function f into the hidden layer $\mathbf{h}$, often called a code layer in the case of auto-encoders:

$$\mathbf{h} = f(\mathbf{x}) = s(-\mathbf{W}^{(1)}\mathbf{x} + \mathbf{b}^{(1)}) \tag{1}$$

where $\mathbf{W}^{(1)}$ is the encoding weight matrix, $\mathbf{b}^{(1)}$ the bias vector and $s(\cdot)$ the activation function. Note that this one hidden layer auto-encoder network corresponds to a principal component analysis if the activation function is linear. Then, the code

representation is mapped back by a decoder function g into a reconstruction $\mathbf{z}$:

$$\mathbf{z} = g(f(\mathbf{x})) = s(-\mathbf{W}^{(2)}\mathbf{h} + \mathbf{b}^{(2)}) \tag{2}$$

where $\mathbf{W}^{(2)}$ is the decoding weight matrix. Tying the weights ($\mathbf{W}^{(2)} = \mathbf{W}^{(1)^T}$) has several advantages. It acts as a regularizer by preventing tiny gradients and it reduces the number of parameters to optimize [2]. Finally, the parameters $\theta = \{\mathbf{W}^{(1)}, \mathbf{b}^{(1)}, \mathbf{b}^{(2)}\}$ are optimized in order to minimize the squared reconstruction error:

$$L_2(\mathbf{x}, \mathbf{z}) = \|\mathbf{x} - \mathbf{z}\|^2 . \tag{3}$$

In the case of dimensionality reduction, the code layer $\mathbf{h}$ has a smaller dimension than the input $\mathbf{x}$. One major drawback comes from the gradient descent algorithm for the training procedure that is very sensitive to the initial weights. If they are far from a good solution, training a deep non-linear auto-encoder network is very hard [6]. A pre-training algorithm is thus required to learn more robust features before fine-tuning the whole model.

The idea for initialization is to build a model that reconstructs the input based on a corrupted version of itself. This can either be done by Restricted Boltzmann Machines (RBMs) [6] or denoising auto-encoders [18] (used in this study). The denoising auto-encoder is considered as a stochastic version of the auto-encoder [18]. The difference lies in the stochastic corruption process that sets randomly some of the inputs to zero. This corrupted version $\tilde{\mathbf{x}}$ of the input $\mathbf{x}$ is obtained by a stochastic mapping $\tilde{\mathbf{x}} \sim q_D(\tilde{\mathbf{x}}|\mathbf{x})$ with a proportion of corruption v. The denoising auto-encoder is then trained to reconstruct the uncorrupted version of the input from the corrupted version, which means that the loss function in Eq. 3 remains the same. Therefore, the learning process tries to cancel the stochastic corruption process by capturing the statistical dependencies between the inputs [2]. Once all the layers are pre-trained, the auto-encoder proceeds to a fine-tuning of the parameters θ (i.e. without the corruption process).

2.3 K-Means++ Clustering Algorithm

Once the fine-tuning of the stacked auto-encoder has learnt the parameters θ, low-dimensional codes from the code layer can be extracted for each patient's spine. Clusters in the codes were obtained using the K-Means++ algorithm [1], which is a variant of the traditional K-Means clustering algorithm but with a selective seeding process. First, a cluster centroid is initialized among a random code layer $\mathbf{h}$ of the dataset χ following a uniform distribution. Afterwards, a probability is assigned to the rest of the observations for choosing the next centroid:

$$p(\mathbf{h}) = \frac{D(\mathbf{h})^2}{\sum_{h \in \chi} D(\mathbf{h})^2} \tag{4}$$

where $D(\mathbf{h})^2$ corresponds to the shortest distance from a point $\mathbf{h}$ to its closest cluster centroid. After the initialization of the cluster centroids, the K-Means++ algorithm proceeds to the regular Lloyd's optimization method.

The selection of the right number of clusters k is based on the validity ratio [14], which minimizes the intra-cluster distance and maximizes the inter-cluster distance. The ratio is defined as $validity = intra/inter$. The intra-cluster distance is the average of all the distances between a point and its cluster centroid:

$$\text{intra} = \frac{1}{N} \sum_{i=1}^{k} \sum_{\mathbf{x} \in C_i} \|\mathbf{h} - \mathbf{c}_i\|^2. \tag{5}$$

where N is the number of observations in χ and $\mathbf{c}_i$ the centroid of cluster i. The inter-cluster distance is the minimum distance between cluster centroids.

$$\text{inter} = \min(\|\mathbf{c}_i - \mathbf{c}_j\|^2) \tag{6}$$

where $i = 1, 2, \ldots, k-1$ and $j = i+1, \ldots, k$.

2.4 Clinical Data Analysis

Once clusters were created from the low-dimensional representation of the dataset, we analyzed the clustered data points with 3D geometric indices in the main thoracic (MT) and thoracolumbar/lumbar (TLL) regions for each patient's spine. One-way ANOVA tested differences between the cluster groups with a significance level $\alpha = 0.05$. The p-values were adjusted with the Bonferroni correction. For all cases, the following 3D spinal indices were computed: the orientation of the plane of maximum curvature (PMC) in each regional curve, which corresponds to the plane orientation where the projected Cobb angle is maximal; the kyphotic angle, measured between T2 and T12 on the sagittal plane; the lumbar lordosis angle, defined between L1 and S1 on the sagittal plane; the Cobb angles in the MT and TLL segments; and the axial orientation of the apical vertebra in the MT region, measured by the Stokes method [16].

3 Clinical Experiments

3.1 Clinical Data

A cohort of 155 AIS patients was recruited for this preliminary study at our institution. A total of 277 reconstructions of the spine was obtained in 3D. From this group, 60 patients had repeat measurements from multiple clinic visits (mean = 3 visits). The mean thoracic Cobb angle was 53.2 ± 18.3° (range = 11.2–100.2°). All patients were diagnosed with a right thoracic deformation and classified as Lenke Type-1 deformity. A lumbar spine modifier (A, B, C) was also assigned to each observation, using the biplanar X-Ray scans available for each patient. The dataset included 277 observations divided in 204 Lenke Type-1A, 43 Lenke Type-1B and 30 Lenke Type-1C deformities. The training set included 235 observations, while the validation set included 42 observations (15 % of the whole dataset). Observations were randomly assigned in each set.

3.2 Hyper-Parameters of the Stacked Auto-Encoders

The neural network hyper-parameters were chosen by an exhaustive grid search. The architecture yielding to the lowest validation mean squared error (MSE) is described in Fig. 1. We used an encoder with layers of size 867-900-400-200-50 and a decoder with tied weights to map the high-dimensional patient's spine models into low-

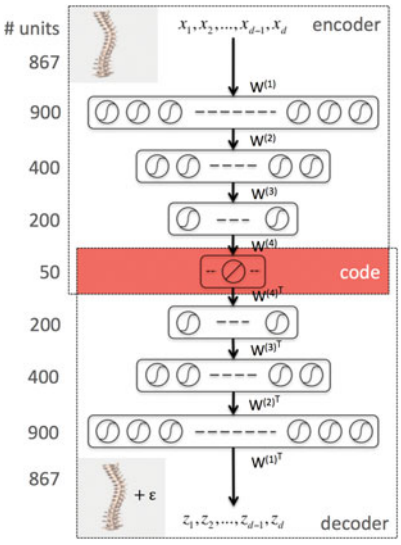

Fig. 1 Illustration of the stacked auto-encoders architecture to learn the 3D spine model by minimizing the loss function. The middle layer represents a low-dimensional representation of the data, called the code layer. An optimal layer architecture of 867-900-400-200-50 was found after a coarse grid search of the hyper-parameters

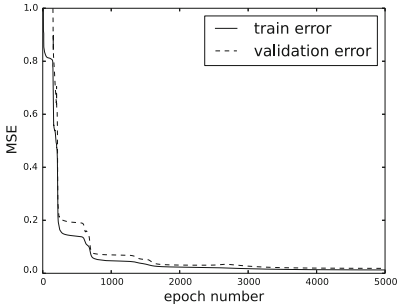

Fig. 2 Evolution of the mean squared error (MSE) with respect to the number of epochs to determine the optimal model described in Fig. 1

dimensional codes. All units in the network were activated by a sigmoidal function $s(a) = 1/(1 + e^{-a})$, except for the 50 units in the code layer that remain linear $s(a) = a$.

3.3 Training and Testing the Stacked Auto-encoders

Auto-encoder layers were pre-trained and fine-tuned with the stochastic gradient descent method using a GPU implementation based on the Theano library [3]. Pre-training had a proportion of corruption for the stochastic mapping in each hidden layer of $v = \{0.15, 0.20, 0.25, 0.30\}$ and a learning rate of $\varepsilon_0 = 0.01$. Fine-tuning ran for 5,000 epochs and had a learning rate schedule with $\varepsilon_0 = 0.05$ that annihilates linearly after 1,000 epochs. Figure 2 shows the learning curve of the stacked auto-encoder. The optimal parameters θ for the model in Fig. 1 were found at the end of training with a normalized training MSE of 0.0127, and a normalized validation MSE of 0.0186, which corresponds to 4.79 and 6.46 mm^2 respectively on the original scale. The learning curve describes several flat regions before stabilizing after 3,500 epochs.

3.4 Clustering the Codes

The K-Means++ algorithm was done using the scikit-learn library [12]. For each number of clusters k (2 through 9), the algorithm ran for 100 times with different centroid seeds in order to keep the best clustering in terms of inertia. Figure 3 depicts the validity ratio against the number of clusters. The validity ratio suggests that the optimal number of clusters should be 4 or 5. However, subsequent analysis illustrated in Table 1 indicates that 5 clusters is the right number of clusters for this dataset because all the clinical indices are statistically significant ($\alpha = 0.05$) given that the other indices are in the model. Figure 4 presents the visualization of the five

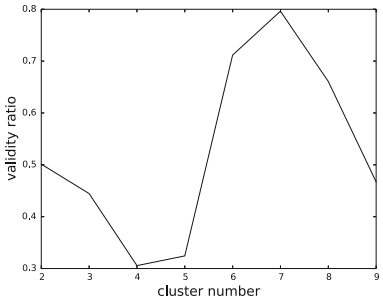

Fig. 3 Validity ratio with respect to the number of clusters, to determine the optimal number of clusters

clusters using a PCA to project the codes in 3D and 2D views. However, it should be mentioned that the clustering was performed on the codes of dimension 50. Figure 5 shows the frontal, lateral and daVinci representations of the centroid of each cluster.

3.5 Clinical Significance

Based on the five identified clusters, cluster 1 is composed of 50 Lenke Type-1A, 12 Type-1B and 2 Type-1C, representing hyper-kyphotic, hyper-lordotic profiles, with high curvatures in the sagittal plane. No sagittal rotation was detected in cluster 1. Cluster 2 is composed of 29 Lenke Type-1A, 7 Type-1B and 0 Type-1C, representing a high axial rotation of the apical vertebra, with the strongest thoracic deformation of all clusters. Moreover, those two clusters have no lumbar derotation.

Clusters 3, 4 and 5 represent the clusters with higher lumbar deformities. Cluster 3 includes 34 Lenke Type-1A, 17 Type-1B and 20 Type-1C, with a minimal thoracic deformation and the highest angulation of TLL plane of all clusters. Cluster 4 includes 23 observations, with 21 Lenke Type-1A, 2 Type-1B and 0 Type-1C. Cluster 4 is characterized by a hypo-kyphotic profile (mean = 7°) and the highest angulation of the MT plane of all clusters. Finally, cluster 5 includes 70 Lenke Type-1A, 5 Type-1B and 8 Type-1C, with a low kyphosis, and medium range thoracic deformations. While this last cluster has a higher orientation of the thoracolumbar curve, its magnitude is not significant.

Surgical strategies based on current 2D classification systems are suboptimal since they do not capture the intrinsic 3D representation of the deformation. Lenke Type-1 classification currently leads to selective thoracic arthrodesis. This very restrictive surgery treatment comes from the hard thresholds on the geometric parameters. A small change in Cobb angle could lead to two different classifications and to two different fusion recommendations subsequently [9]. Therefore, identifying groups based on their true 3D representation will help to better adjust surgery choices such as levels of fusion, biomechanical forces to apply or surgical instrumentations.

Table 1 Mean geometric clinical 3D parameters for the thoracic and lumbar regions, within all five clusters detected by the framework

Parameter (all in degrees)	Cluster 1 (n = 64)	Cluster 2 (n = 36)	Cluster 3 (n = 71)	Cluster 4 (n = 23)	Cluster 5 (n = 83)	p-value$_{ajd}$
PMC MT Cobb	50.6 ± 13.8	73.3 ± 12.7	43.3 ± 15.5	62.2 ± 9.2	52.5 ± 14.6	<0.001*
PMC MT Orient	71.6 ± 12.1	78.1 ± 7.5	72.5 ± 12.1	88.5 ± 4.1	82.0 ± 10.4	<0.001*
PMC TLL Cobb	34.1 ± 12.7	49.5 ± 11.6	36.2 ± 15.1	38.2 ± 12.1	36.6 ± 13.3	<0.001*
PMC TLL Orient	45.4 ± 18.9	51.3 ± 15.0	65.0 ± 24.0	52.8 ± 12.2	48.4 ± 17.3	<0.001*
Kyphosis	36.6 ± 10.6	32.4 ± 11.2	27.4 ± 9.8	7.3 ± 11.6	21.9 ± 11.7	<0.001*
Lordosis	–66.6 ± 8.9	–63.4 ± 14.1	–61.6 ± 12.3	–54.2 ± 13.1	–60.9 ± 9.7	0.002*
MT Cobb	46.6 ± 22.2	69.2 ± 24.8	43.3 ± 15.6	62.2 ± 9.2	50.9 ± 18.4	<0.001*
TLL Cobb	–29.9 ± 22.2	–45.5 ± 25.5	–36.3 ± 14.9	–38.2 ± 12.1	–34.8 ± 18.3	0.036*
Ax. rot. apex	–19.7 ± 10.3	–33.8 ± 7.6	–13.7 ± 8.7	–25.9 ± 8.8	–22.1 ± 12.0	<0.001*

A significant star (*) indicates that the differences of geometric parameter between all groups are statistically significant at level $\alpha = 0.05$, given that the other geometric parameters are in the model

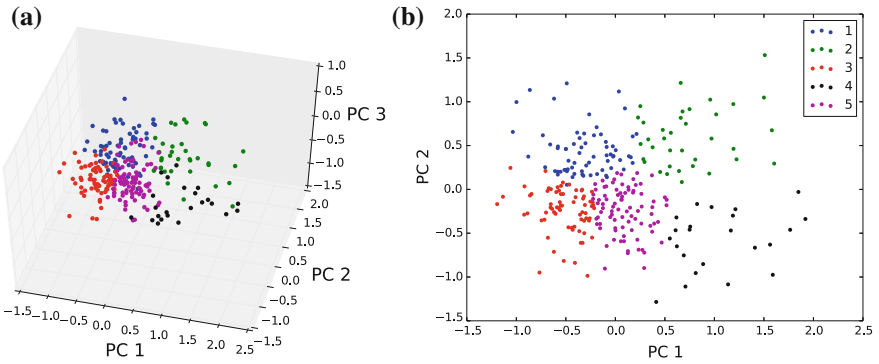

Fig. 4 Visualization of the five clusters found by the K-Means++ algorithm on low-dimensional points, by projecting the 50-dimensional codes into 3 (**a**) and 2 (**b**) principal components (PC) using PCA

In this study, the learning framework provided an optimal number of 5 clusters based on the input population. It is not possible at this stage to infer that Lenke Type-1 should be divided in 5 groups. However, this study confirms that within a defined 2D class currently used for surgical planning, there exists a number of sub-groups with different 3D signatures that are statistically significant. Therefore, each sub-group would lead to different surgical strategies. Previous studies have indicated that within Lenke Type-1 [5, 7, 15], there indeed exists 3D variability in terms of geometric parameters that could be divided in 4–6 sub-groups.

4 Conclusion

In this paper, we presented an automated classification method using a deep learning technique, namely with stacked auto-encoders, to discover sub-groups within a cohort of thoracic deformations. The code layer of the auto-encoder learns a distributed representation in low-dimension that aims to capture the main factors of variation in the training dataset. However, different examples from the distribution of the training dataset may potentially yield to high reconstruction errors [2]. Therefore, having a large and representative training dataset of AIS is critical. This will also prevent the model from overfitting.

The current study evaluated the 3D sub-classification of Lenke Type-1 scoliotic curves, suggesting that shape variability is present within an existing 2D group used in clinical practice. However, these types of approaches include complex synthetization tasks, which require sizeable datasets to improve the data representation within the code layer. Therefore, a multicentric dataset may help to significantly increase the number of cases from various sites and obtain a more reproducible model.

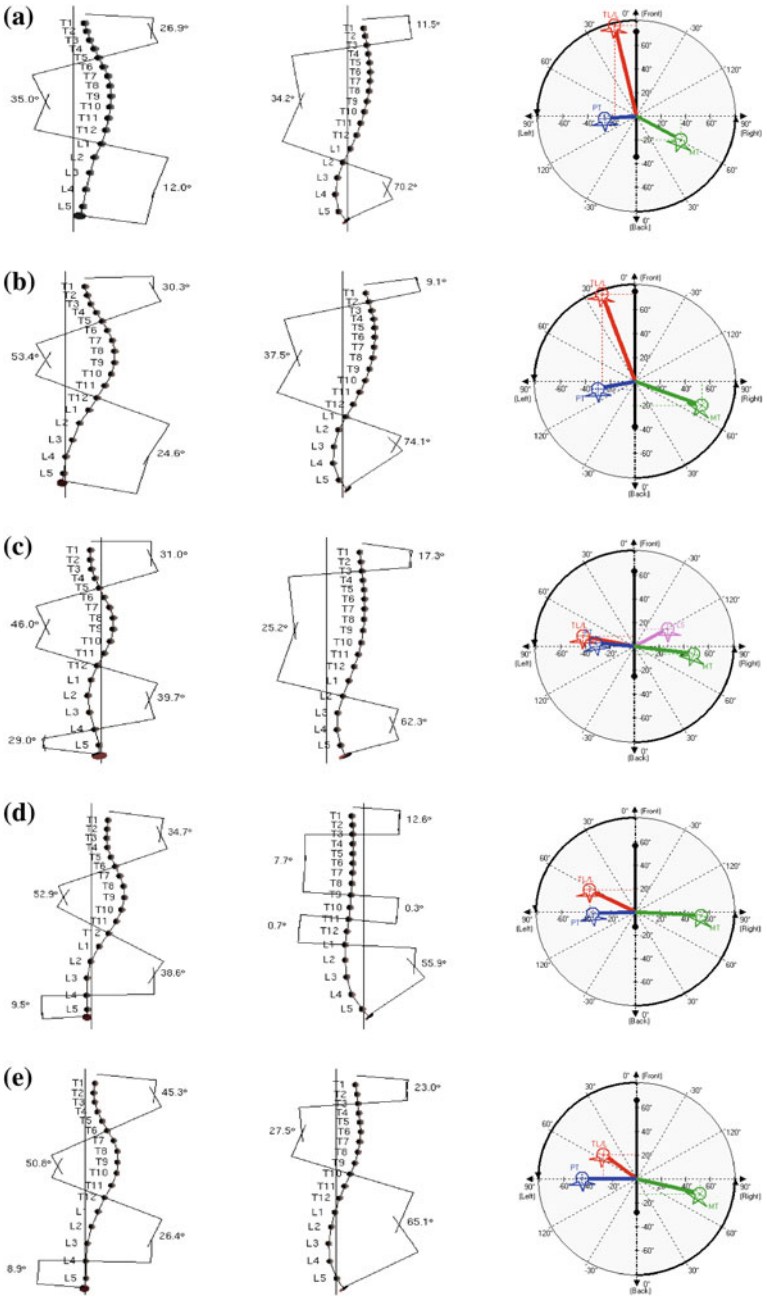

Fig. 5 *Frontal*, *lateral* and *top* view profiles of cluster centers, with daVinci representations depicting planes of maximal deformities. **a** Centroid of cluster 1. **b** Centroid of cluster 2. **c** Centroid of cluster 3. **d** Centroid of cluster 4. **e** Centroid of cluster 5

Furthermore, the development of computational methods will ultimately lead to more reliable classification paradigms, helping to identify possible cases which might progress with time. Future work will include additional Lenke types, such as double major and lumbar deformations. Other works will propose to use longitudinal data for surgical treatment planning, whereas each observation is considered independently in the current framework. Finally, a reliability study will be undertaken to evaluate the relevance of classification systems.

References

1. Arthur, D., Vassilvitskii, S.: k-means++: The advantages of careful seeding. In: Proceedings of the eighteenth annual ACM-SIAM symposium on discrete algorithms. pp. 1027–1035. Society for industrial and applied mathematics (2007)
2. Bengio, Y., Courville, A., Vincent, P.: Representation learning: A review and new perspectives. IEEE transactions on pattern analysis and machine intelligence (2013)
3. Bergstra, J., Breuleux, O., Bastien, F., Lamblin, P., Pascanu, R., Desjardins, G., Turian, J., Warde-Farley, D., Bengio, Y.: Theano: a CPU and GPU math expression compiler. In: Proceedings of the Python for scientific computing conference (SciPy) (2010)
4. Duong, L., Cheriet, F., Labelle, H.: Three-dimensional classification of spinal deformities using fuzzy clustering. Spine **31**(8), 923–930 (2006)
5. Duong, L., Cheriet, F., Labelle, H.: Three-dimensional subclassification of lenke type 1 scoliotic curves. J. Spinal Disord. and Tech. **22**(2), 135–143 (2009)
6. Hinton, G., Salakhutdinov, R.: Reducing the dimensionality of data with neural networks. Science (New York) **313**(5786), 504–507 (2006)
7. Kadoury, S., Labelle, H.: Classification of three-dimensional thoracic deformities in adolescent idiopathic scoliosis from a multivariate analysis. Euro. Spine J. **21**(1), 40–49 (2012)
8. King, H.A., Moe, J.H., Bradford, D.S., Winter, R.B.: The selection of fusion levels in thoracic idiopathic scoliosis. J. Bone and Jt. Surg. **65**(9), 1302–1313 (1983)
9. Labelle, H., Aubin, C.E., Jackson, R., Lenke, L., Newton, P., Parent, S.: Seeing the spine in 3d: how will it change what we do? J. Pediatr. Orthop. **31**, S37–S45 (2011)
10. Lenke, L., Betz, R., Harms, J., Bridwell, K., Clements, D., Lowe, T., Blanke, K.: Adolescent idiopathic scoliosis: a new classification to determine extent of spinal arthrodesis. J. Bone Jt. Surg. Am. bf 83-A(8), 1169–1181 (2001)
11. van der Maaten, L.J., Postma, E.O., van den Herik, H.J.: Dimensionality reduction: A comparative review. J. Mach. Learn. Res. **10**(1–41), 66–71 (2009)
12. Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E.: Scikit-learn: Machine learning in Python. J. Mach. Learn. Res. **12**, 2825–2830 (2011)
13. Pomero, V., Mitton, D., Laporte, S., de Guise, J.A., Skalli, W.: Fast accurate stereoradiographic 3d-reconstruction of the spine using a combined geometric and statistic model. Clin. Biomech. **19**(3), 240–247 (2004)
14. Ray, S., Turi, R.H.: Determination of number of clusters in k-means clustering and application in colour image segmentation. In: Proceedings of the 4th international conference on advances in pattern recognition and digital techniques. pp. 137–143 (1999)
15. Sangole, A.P., Aubin, C., Labelle, H., Stokes, I.A.F., Lenke, L.G., Jackson, R., Newton, P.: Three-dimensional classification of thoracic scoliotic curves. Spine **34**(1), 91–99 (2009)
16. Stokes, I.A., Bigalow, L.C., Moreland, M.S.: Measurement of axial rotation of vertebrae in scoliosis. Spine **11**(3), 213–218 (1986)

17. Stokes, I.A., Sangole, A.P., Aubin, C.E.: Classification of scoliosis deformity 3-d spinal shape by cluster analysis. Spine **34**(6), 584–590 (2009)
18. Vincent, P., Larochelle, H., Bengio, Y., Manzagol, P.A.: Extracting and composing robust features with denoising autoencoders. In: Proceedings of the 25th international conference on machine learning. pp. 1096–1103. ACM (2008)

An Active Optical Flow Model for Dose Prediction in Spinal SBRT Plans

Jianfei Liu, Q. Jackie Wu, Fang-Fang Yin, John P. Kirkpatrick, Alvin Cabrera and Yaorong Ge

Abstract Accurate dose predication is critical to spinal stereotactic body radiation therapy (SBRT). It enables radiation oncologists and planners to design treatment plans that maximally protect spinal cord while effectively controlling surrounding tumors. Spinal cord dose distribution is primarily affected by the shapes of tumor boundaries near the organ. In this work, we estimate such boundary effects and predict dose distribution by exploring an active optical flow model (AOFM). To establish AOFM, we collect a sequence of dose sub-images and tumor contours near spinal cords from a database of clinically accepted spine SBRT plans. The data are classified into five groups according to the tumor location in relation to the spinal cords. In each group, we randomly choose a dose sub-image as the reference and register all other dose images to the reference using an optical flow method. AOFM is then constructed by importing optical flow vectors and dose values into the principal component analysis. To develop the predictive model for a group, we also build active shape model (ASM) of tumor contours near the spinal cords. The correlation between ASM and AOFM is estimated via the multiple regression model. When predicting dose distribution of a new case, the group was first determined based on the case's

J. Liu (✉) · Q. Jackie Wu · F.-F. Yin · J.P. Kirkpatrick · A. Cabrera
Department of Radiation Oncology, Duke University Medical Centre,
Durham, NC, USA
e-mail: jianfei.liu@duke.edu

Q. Jackie Wu
e-mail: jackie.wu@dm.duke.edu

F.-F. Yin
e-mail: fangfang.yin@dm.duke.edu

J.P. Kirkpatrick
e-mail: john.kirkpatrick@dm.duke.edu

A. Cabrera
e-mail: alvin.cabrera@dm.duke.edu

Y. Ge
Department of Software and Information Systems,
University of North Carolina at Charlotte, Charlotte, NC, USA
e-mail: yge@uncc.edu

© Springer International Publishing Switzerland 2015

J. Yao et al. (eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, Lecture Notes in Computational Vision and Biomechanics 20, DOI 10.1007/978-3-319-14148-0_3

tumor contour. Then the corresponding model for the group is used to map from the ASM space to the AOFM space. Finally, the parameters in the AOFM space are used to estimate dose distribution. This method was validated on 30 SBRT plans. Analysis of dose-volume histograms revealed that at the important 2 % volume mark, the dose difference between prediction and clinical plan is less than 4 %. These results suggest that the AOFM-based approach is a promising tool for predicting accurate spinal cord dose in clinical practice.

1 Introduction

Spinal tumors are neoplasm located at spinal cords, and most of them are metastases from primary cancers elsewhere [10]. The compression of spinal tumors causes patients to undergo severe pain, and radiation therapy is a primary procedure for pain relief and tumor removal [5]. In contrast to radiation therapy on lung [12] and bladder [11] in which parts of organ can be scarified to ensure tumors receiving maximum dosage, spinal cords are a sensitive serial organ that must be protected because they control the nerve system. Due to this challenge, knowledge-guided dose prediction on spinal cords is seldom addressed.

Dose levels at spinal cords are mainly caused by the tumor enclosure around the cord in Fig. 1. Dose values within the yellow curve in Fig. 1d are higher than in Fig. 1c because the tumor surrounds the cord more in Fig. 1b in comparison with Fig. 1a. One way to predict dose is therefore to estimate correlation between dose distributions and tumor boundary shapes. Statistical shape analysis can serve this purpose, and it has been extensively studied in the fields of computer vision and medical image analysis [4]. Cootes [2] pioneered this research by applying the principal component analysis (PCA) to a set of landmarks in face images. Active shape model (ASM) was constructed to guide face recognition. Later on, Cootes embedded intensity values and landmark positions into the PCA, which yielded the active appearance model (AAM) [3]. However, both ASM and AAM require extensive labor work to manually

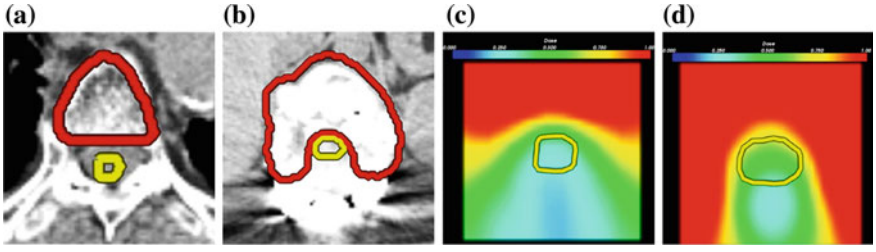

Fig. 1 The effects of tumor enclosure on dose distributions. **a**, **b** Spinal tumors (marked by *red curves*) and cords (*yellow curves*), and **c**, **d** corresponding dose distributions, where *blue* to *red* means small to large dose values (Color figure online)

select a set of points from each sample image. Rueckert [9] directly imported control points of image deformation fields into the PCA analysis to avoid manual labeling, and the resulting statistical model showed great accuracy in measuring shape variance of a brain image database. Such idea was also extended to 3D irregular heart models to estimate cardiac motion [6]. This method may not directly apply to spinal cord dose prediction because it focused on the statistical description of shape variance to assist organ segmentation and registration. Moreover, ASM established on a set of sparse points cannot meet the accuracy requirements of dose prediction in small spinal cords.

In this paper, we present an active optical flow model (AOFM) to predict dose in spinal cords by (1) applying optical flow to measure the variations at every point of a dose image, (2) importing dose values and optical flow vectors into the PCA analysis, (3) establishing ASM of the tumor contours, and (4) utilizing multiple regression method to compute correlation between AOFM and ASM.

2 Methods

This section elaborates on our framework with respect to data preprocessing, active optical flow model, correlation computation, and dose prediction. In this framework, we treat dose distribution as dose images with intensity at image point representing dose value.

2.1 Data Preprocessing

This step aims to extract a sequence of dose sub-images and tumor contours adjacent to spinal cords from SBRT plans in Fig. 2. AOFM is constructed in the sub-images because dose at spinal cords are only affected by local tumor boundaries.

The size of dose sub-image is 41 $\times$ 41 pixels indicated as a green square in Fig. 2a (approximately 41 mm if intra-spacing is 1.0 mm in the CT images) because the diameter of spinal cord is 10–15 mm and tumors have minor effect on spinal cord if their distance is larger than 10 mm. Thus, 41 pixels are sufficient to include all types of tumors located at different sides of the cord while still preserving the accuracy of measuring the boundary effects on dose prediction. We next extract tumor contours that are adjacent to spinal cords in Fig. 2c because tumors and cord contours are available in the SBRT plan. Finally, all dose sub-images and tumor contours are classified into five groups in terms of spatial relationships between tumors and spinal cords in Fig. 3.

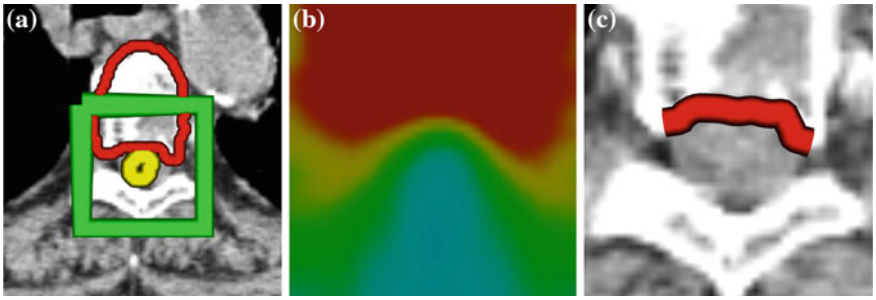

Fig. 2 Extracting dose sub-images and tumor contours to build AOFM and ASM. **a** *Green rectangle* at the cord center illustrates the spatial range of the sub-image, **b** dose sub-image, and **c** the *red curve* representing the tumor contours near the spinal cord (Color figure online)

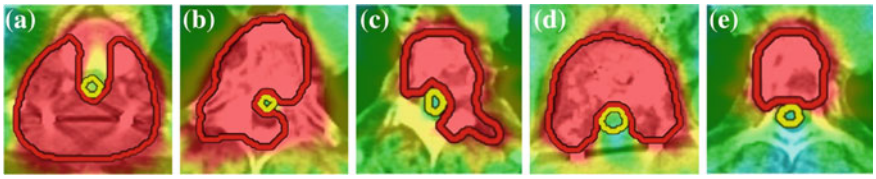

Fig. 3 Five types of spatial relationships between spinal cords and tumors, where dose images are overlaid on the CT images

2.2 Active Optical Flow Model (AOFM)

Dose sub-images are used to compute AOFM, which is essentially a statistical model to describe the dose distributions within a group. Estimating AOFM involves optical flow computation and principal component analysis.

In each group, we randomly choose a dose image $D_r(x, y)$ as the reference image (Fig. 4a). Rigid image registration [7] is performed to remove global motion between the reference image and the current dose image (Fig. 4b), and generate the registered

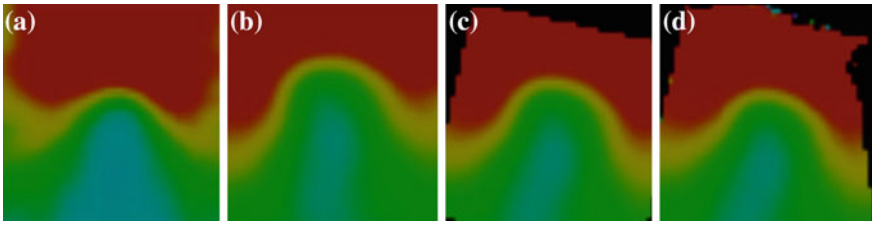

Fig. 4 Process of optical flow computation. **a** Reference dose image, **b** current dose image, **c** transformed current dose image after rigid registration, and **d** transformed image after optical flow computation

image $D_g(x, y)$ (Fig. 4c). $D_g(x, y)$ is represented in the reference image coordinate, and optical flow computation thus becomes meaningful to measure local deformation between the registered image and the reference image. Let $\mathbf{u} = (u_x, u_y)$ be the optical flow vector at a point $\mathbf{p} = (x, y)$ in the sub-image domain Ω. The optical flow computation is formulated as a global energy functional within a variational framework.

$$E(\mathbf{u}) = \int_{\Omega} \Big(\Psi(|D_g(\mathbf{p}+\mathbf{u}) - D_r(\mathbf{p})|^2) + \alpha\Psi(|\nabla D_g(\mathbf{p}+\mathbf{u}) - \nabla D_r(\mathbf{p})|^2) + \beta\Psi(|\nabla u_x|^2 + |\nabla u_y|^2) \Big) d\mathbf{p} \tag{1}$$

where $\Psi(s^2) = \sqrt{s^2 + 0.001^2}$ is a modified *L1* norm that allows for handling outliers. α and β are constant values to balance different terms. Minimizing Eq. 1 generates an optical flow field [1]. Figure 4d gives the result by using optical flow vectors to transform Fig. 4c.

Next, we perform PCA analysis on M optical flow fields to establish AOFM, where M is the number of sub-images in the current group and each sub-image defines a feature vector $\mathbf{x} = (u_x^1, \ldots, u_x^N, u_y^1, \ldots, u_y^N, d^1, \ldots, d^N)$. N is the number of pixels in Ω and d^i is the dose value at ith pixel. We can approximate any feature vector $\mathbf{x}$ using

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{\Phi}_f \mathbf{b}_f \tag{2}$$

Here, $\bar{\mathbf{x}}$ is the average feature vector, $\mathbf{\Phi}_f$ is formed by the eigenvectors of the covariance matrix, and $\mathbf{b}_f$ is the vector of principal component scores. Figure 5a shows the variance of AOFM corresponding to the group in Fig. 3e by setting the first parameter of $\mathbf{b}_f$ to $\pm 3\sqrt{\lambda_1}$, where $\lambda_1 \geq \cdots \geq \lambda_N$ are eigenvalues of the covariance matrix.

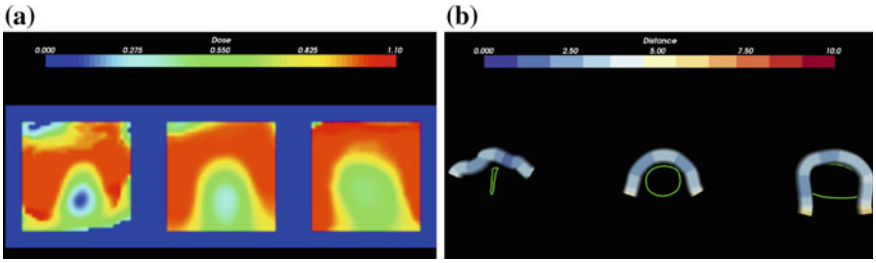

Fig. 5 Instances of **a** active optical flow model and **b** active shape model. The *center* image of (**a**) and (**b**) is the average model, and the *left* and *right* images are the first modes of variation equal to $-3\sqrt{\lambda_1}$ and $3\sqrt{\lambda_1}$, respectively

Similarly, we can adapt ASM [2] to measure the shape variance of the tumor contours in the current group.

$$\mathbf{y} = \bar{\mathbf{y}} + \mathbf{\Phi}_s \mathbf{b}_s \tag{3}$$

Each shape feature vector $\mathbf{y}$ includes tumor contour locations, cord contour locations, and distance between tumor contours and spinal cords. Figure 5b shows the resulting ASM. We can observe that dose values in the cord (the center of the images) increase in proportion to the extent of tumor enclosure on spinal cords.

2.3 Correlation Computation

This step quantitatively measures correlation between AOFM and ASM using the multiple regression method [8]. Principal component scores, $\mathbf{b}_f$ and $\mathbf{b}_s$ of AOFM and ASM are chosen for the estimation of correlation parameters since $\mathbf{b}_f = (\mathbf{x} - \bar{\mathbf{x}}) \cdot \mathbf{\Phi}_f^{\mathrm{T}}$ and $\mathbf{b}_s = (\mathbf{y} - \bar{\mathbf{y}}) \cdot \mathbf{\Phi}_s^{\mathrm{T}}$ can normalize the feature vectors in two models. The first 11 components of $\mathbf{b}_f$ and $\mathbf{b}_s$ are selected to estimate correlation parameters $\mathbf{r}$ satisfying $\mathbf{b}_f = \mathbf{r} \cdot \mathbf{b}_s$, because these components are found through experimentation sufficient to represent the characteristics of AOFM and ASM.

2.4 Dose Prediction

The final step is to predict dose in the new case. A sequence of image slices containing both spinal cords and tumors are first determined. In each image slice, we extract tumor contours and formulate a shape feature vector $\mathbf{y}$. A group is then selected by searching for the shortest distance between the current tumor contour and the average contour in the group. The principal component score of the current contour is computed as $\mathbf{b}_s = (\mathbf{y} - \bar{\mathbf{y}}) \cdot \mathbf{\Phi}_s^{\mathrm{T}}$. The corresponding principal component score of the AOFM is computed as $\mathbf{b}_f = \mathbf{r} \cdot \mathbf{b}_s$. The feature vector containing dose values and optical flow vectors is derived as $\mathbf{x} = \mathbf{r} \cdot \mathbf{b}_s \cdot \mathbf{\Phi}_f + \bar{\mathbf{x}}$. An initial dose shown in Fig. 6a can be reconstructed by using x. However the initial dose is represented in the reference image coordinate of the selected group, and we need to transform it into the current image space. Such image transformation is estimated by using iterative closest point algorithm [13] to match the tumor contours in the reference image (red curve in Fig. 6a) and contours in the current image (red curve in Fig. 6c). Figure 6b gives the transformed result, and as shown, it is comparable with the actual clinical plan in Fig. 6c. Finally, dose of the current case was predicted by iteratively applying the mentioned strategy to all other image slices that contain tumor and spinal cord.

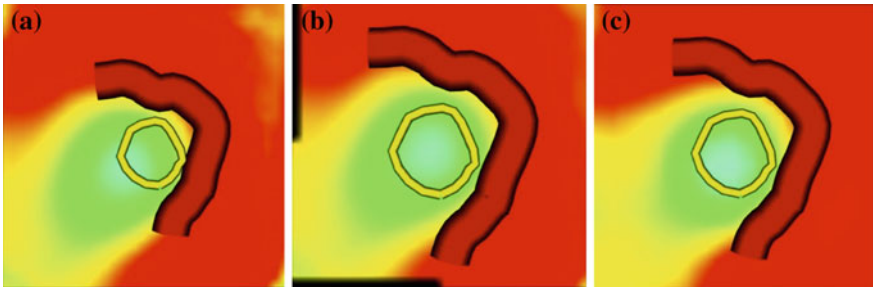

Fig. 6 Process of dose prediction. **a** Initial dose in the reference image coordinate, **b** final dose in the current patient image coordinate, and **c** clinical dose

3 Validation Datasets and Methods

Thirty clinically accepted spinal SBRT plans were evenly divided into training and testing datasets (15 patients each) in this study. The anatomic distributions were 4 C-spine, 6 L-spine and 20 T-spine SBRT plans, with tumor size ranging in 13.24–982.8 cm^3 (mean ± std.: 116.68 ± 175.34 cm^3), and the affected cord volume range in 0.62–16.04 cm^3 (mean ± std.: 4.49 ± 4.12 cm^3). The prescription dose

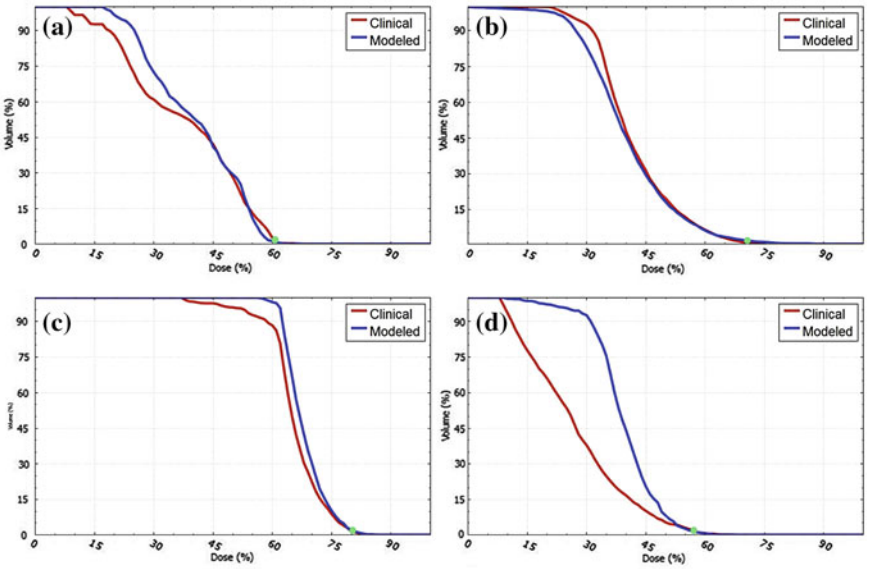

Fig. 7 DVH graphs of four cases, where *green points* represent D_2 (Color figure online). Note that DVH results are comparable in the first three cases (**a**-**c**), while there contain large errors in the low dose region in (**d**). However, the dose values are still similar at 2 % volume

range was 14.25–25 Gy (meanstd.: 18 ± 3 Gy) in 2–5 fractions. Dose-volume histogram (DVH) (Fig. 7) was chosen to evaluate the predicted dose in comparison with the clinical dose, which was considered as the ground-truth. The DVH was plotted with bin doses along the horizontal axis, and the column height of the bin represents the volume of spinal cords receiving greater than or equal to that dose. We measured the dose difference between the predicted and clinical dose at 2 % volume in the DVH (D_2). This is a common strategy to evaluate the quality of SBRT radiation plans in the clinical settings.

4 Experimental Results

Figure 7 shows the DVH results on L-spine (Fig. 7a), C-spine (Fig. 7b), and T-spine (Fig. 7c, d) SBRT plans in the testing dataset. One can find that the estimated DVH (blue curve) and clinical DVH (red curve) are very similar except in the low dose regions in Fig. 7d. In this case, the dose in the spinal cord is much lower than all samples in the training dataset. However, the clinical and predicted dose at D_2 (green points) are still comparable. We compute this value in the entire testing dataset, and the dose difference between prediction and clinical plan at D_2 is 3.3 ± 3.5 %.

In Fig. 8, we show four predicted dose corresponding to four plans in Fig. 7. In the left column, tumors are located at the top of the spinal cord, and AOFM can predict dose very well in comparison with the clinical dose. The second column illustrated a tumor in the left side, and AOFM still predicted it reasonably well. Similar results can be observed in the third column where tumors wrap around spinal cords from the bottom. The fourth column gives an example in which our algorithm over-predicted the dose values because the training dataset was considerably different from this

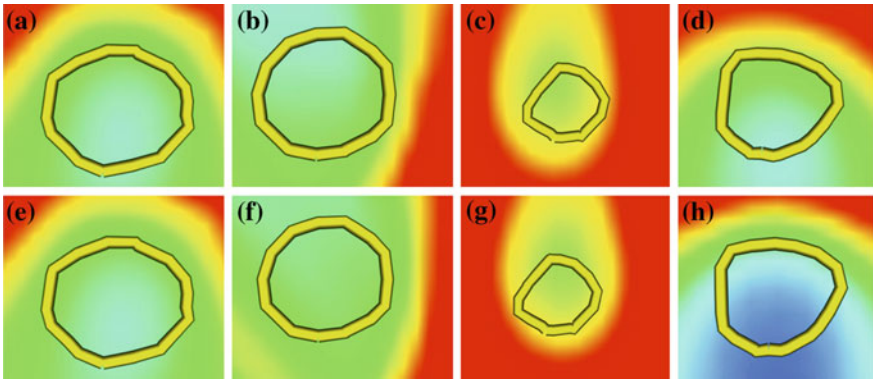

Fig. 8 Comparison of the predicted dose (*top row*) and clinical dose (*bottom row*). Each column corresponding to a SBRT plan in Fig. 7. Spinal cords are represented as the *yellow curves* (Color figure online)

extreme case. Thus, in Fig. 7d, the predicted DVH is higher than the clinical DVH in the low dose region. All these experimental results support our findings in Fig. 7.

5 Conclusion and Future Work

In this work, we developed an active optical flow model to represent dose distributions and active shape model to measure the shape variance of tumor contours near the spinal cords. We used multiple regression models to correlate parameters between AOFM and ASM of a clinically accepted SBRT plan database to predict dose distributions in the new cases. The early experiments are very encouraging; predicting dose with less than 4 % difference from the clinically planned dose at 2 % volume in the Dose-volume histogram. Future work will include training and testing our models on significantly larger datasets, fine tuning of model parameters, and further exploration of higher order models for more robustness and better accuracy.

References

1. Brox, T., Bruhn, A., Papenberg, N., Weickert, J.: High accuracy optical flow estimation based on a theory for warping. Proc. ECCV **3024**, 25–36 (2004)
2. Cootes, T., Cooper, D., Taylor, C., Graham, J.: Active shape models—their training and application. CVIU **61**, 38–59 (1995)
3. Cootes, T., Edwards, G., Taylor, C.: Active appearance models. IEEE PAMI **23**, 681–685 (2001)
4. Heimann, T., Meinzer, H.: Statistical shape models for 3d medical image segmentation: a review. Med. Image Anal. **13**(4), 543–563 (2009)
5. Holt, T., Hoskin, P., Maranzano, E., et al.: Malignant epidural spinal cord compression: the role of external beam radiotherapy. Curr. Opin. Supportive Palliat. Care **6**(1), 103–108 (2012)
6. Horkaew, P., Yang, G.: Optimal deformable surface models for 3d medical image analysis. IPMI **2732**, 13–24 (2003)
7. Ibanez, L., Schroeder, W., Ng, L., Cates, J.: The ITK Software Guide. Kitware, Clifton Park (2005)
8. Neter, J., Kutner, M., Nachtsheim, C., Wasserman, W.: Applied Linear Regression Models. Irwin, Homwood (1996)
9. Rueckert, D., Frangi, A., Schnabel, J.: Automatic construction of 3-d statistical deformation models of the brain using nonrigid registration. IEEE TMI **22**, 1014–1025 (2003)
10. Yasushi, H., Kataoka, M., Senba, T., et al.: Vertebral metastases with high risk of symptomatic malignant spinal cord compression. Jpn. J. Clin. Oncol. **39**(7), 431–434 (2009)
11. Yuan, L., Ge, Y., Lee, W., Yin, F., Kirpatrick, J., Wu, Q.: Quantitative analysis of the factors which affect the interpatient organ-at-risk dose sparing variation in IMRT plans. Med. Phys. **39**(11), 6868–6878 (2012)
12. Zhang, X., Li, X., Quan, E., Li, Y.: A methodology for automatic intensity-modulated radiation treatment planning for lung cancer. Phys. Med. Biol. **56**(13), 3873–3893 (2011)
13. Zhang, Z.: Iterative point matching for registration of free-form curves and surfaces. IJCV **13**(2), 119–152 (1994)

Portable Optically Tracked Ultrasound System for Scoliosis Measurement

Guillermo Carbajal, Álvaro Gómez, Gabor Fichtinger and Tamas Ungi

Abstract Monitoring spinal curvature in adolescent kyphoscoliosis requires regular radiographic examinations; however, the applied ionizing radiation increases the risk of cancer. Ultrasound imaging is favorable over X-ray because it does not emit ionizing radiation. It has been shown in the past that tracked ultrasound can be used to localize vertebral transverse processes as landmarks along the spine to measure curvature angles. Tests have been performed with spine phantoms, but scanning protocol, tracking system, data acquisition and processing time has not been considered in human subjects yet. In this paper, a portable optically tracked ultrasound system for scoliosis measurement is presented. It provides a simple way to acquire data in the clinical environment with the aim of comparing results to current X-ray-based measurement. The workflow of the procedure was tested on volunteers. The customized open-source software is shared with the community as part of our effort to make a clinically practical system.

Keywords Adolescent idiopathic kyphoscoliosis · Scoliosis · Kyphosis · Tracked sonography · Tracked ultrasound snapshot

G. Carbajal · Á. Gómez
Instituto de Ingeniería Eléctrica Facultad de Ingeniería, Universidad de la República, Montevideo, Uruguay
e-mail: carbajal@fing.edu.uy

Á. Gómez
e-mail: agomez@fing.edu.uy

G. Fichtinger (✉) · T. Ungi
Laboratory for Percutaneous Surgery, School of Computing, Queen's University, Kingston, ON, Canada
e-mail: gabor@cs.queensu.ca

T. Ungi
e-mail: ungi@queensu.ca

© Springer International Publishing Switzerland 2015

J. Yao et al. (eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, Lecture Notes in Computational Vision and Biomechanics 20, DOI 10.1007/978-3-319-14148-0_4

1 Introduction

1.1 Adolescent Idiopathic Scoliosis

The most frequent form of spinal deformities is adolescent idiopathic scoliosis. Studies conducted in different countries found its prevalence between 1 and 5 %, therefore, it is considered a common disease in children [1–3]. The etiology of this disease has not yet been fully discovered, however, genetic factors influence both the incidence and the severity of scoliosis [4]. It is most commonly diagnosed in early adolescence (9–12 years of age) and progresses until the spine reaches full development, around the age of 18–20 years.

Symptoms of adolescent idiopathic scoliosis include lateral curvature of the spine in straight standing position, kyphosis, and permanent vertebral rotation around the axis of the spine. The severity of scoliosis is characterized by the angle of curvature between vertebrae above and below the curvature (Cobb angle). Therapeutic protocols are based on Cobb angle: (i) curvatures less than 20° require X-ray monitoring every 3–6 months, depending on the rate of progression; (ii) scoliosis between 20°–40° degrees is treated with bracing; and (iii) scoliosis over 40°, or if chest deformation causes breathing difficulties, is treated surgically, permanently fusing the vertebrae in a straight position. Scoliosis monitoring continues after surgical treatment as well using regular spine imaging. About 10 % of adolescent scoliosis cases progress to a state when they require therapy.

Although screening children in the early adolescent age for scoliosis would be necessary for optimal treatment and optimal use of clinical resources, currently affordable screening methods are inaccurate [5]. Therefore, patients may be diagnosed when scoliosis progressed in an advanced stage.

Once diagnosis is confirmed, it is important to frequently monitor all cases to make sure progression is detected early and appropriate therapy is started. However, frequent X-ray imaging in children can lead to increased risk of cancer. Girls who undergo regular spine X-ray, have a nearly twofold risk of breast cancer as adults [6, 7]. Another study found that repeated X-ray exams in childhood significantly contribute to leukemia and prostate cancer [8]. Further studies may be needed to establish accurate estimates of the risks of X-ray, but a safe scoliosis monitoring method would improve the health of this young patient population.

1.2 Scoliosis Monitoring Techniques

Radiation-free scoliosis monitoring methods have been investigated in the past, but none of them have been successful in replacing X-ray in the clinical practice. The optimal method needs to be: free of ionizing radiation, accessible to the patient population, and accurate for therapeutic decision making.

Magnetic resonance imaging (MRI) is a safe and accurate alternative to X-ray. Open MRI machines permit scanning from a standing patient position, making these images suitable for scoliosis angle measurement. Unfortunately, MRI is less accessible than X-ray due to its high cost and the patient has to stand motionless for several minutes while the scanner captures the entire spinal column, which further limits its use for routine monitoring in an adolescent population [9].

Inspection will yield visual signs of scoliosis on the back of the patient. Visual signs of spinal curvatures have been measured using computerized topographic scans of the skin surface using laser scanners and stereo camera technology [10]. However, surface scans are not informative enough to support therapeutic decisions [11].

Ultrasound does not have a large enough field of view to directly visualize spinal curvatures. There have been attempts to use indirect ways of measuring spinal curvatures with ultrasound. A correlation between vertebral rotation and scoliosis angles in untreated patients was discovered [12] but this correlation is unreliable and completely lost in patients receiving therapy [13].

One of the recently emerging imaging modalities is tracked ultrasound: a combination of conventional ultrasound and position tracking technology. With tracked ultrasound it is possible to create a 3D reconstruction from 2D ultrasound images. Position tracking of the ultrasound transducer allows to display the whole spine region in 3D. Experimental results have confirmed that landmarks on reconstructed image volumes could be used to monitor spine curvature angles [14, 15]. Tracked ultrasound appears to be the only alternative to X-ray that fulfills all three requirements: safety, accessibility and accuracy.

A new method for scoliosis measurement has been developed previously, and tested on phantom models [16]. Tracked ultrasound (TUS) has been successfully used in medical imaging and image-guided interventions. In particular, it has been tested in spinal injection navigation [17] and vertebra localization for spine surgery [18].

1.3 Objectives

Our goal is to develop a clinically practical system for scoliosis monitoring. As part of this translational effort we build upon of the proof of concept phantom system proposed by [16] and present a portable TUS system based on a new ultrasound machine and optical tracking. We streamline the workflow of the procedure so it can be easily replicated by other researchers and operated without deep technical background. The workflow is implemented as an open-source module of the 3D Slicer[1] application, and can be conveniently installed from the 3D Slicer extension manager (app store). We evaluated the usability of the developed TUS system.

[1] www.slicer.org.

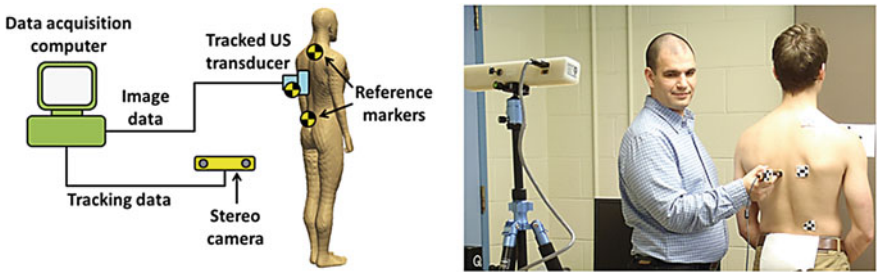

Fig. 1 Schematics of the TUS system (*left panel*), and the prototype system in use (*right panel*)

2 Materials and Methods

2.1 System Design

The TUS system has three main components. An optical tracker (Micron Tracker, Claron Technologies Inc., Toronto, Ontario, Canada), a portable ultrasound scanner with USB connection (General Purpose 99–5914 Probe, Interson Corporation, Pleasanton, CA, USA), and a laptop computer (Fig. 1).

Contrary to [16] we use an optical tracker instead of an electromagnetic tracker. Optical trackers can use wireless markers, are generally more accurate and are not affected by metallic or electronic objects in the environment. Sensors used as reference markers are made of paper, which is very cheap and also makes feasible the tracking of multiple reference markers. This can be used for patient motion detection and to be more robust to markers occlusions.

Hardware interfaces are implemented in the PLUS software[2] [19]. PLUS implements an abstraction layer over tracker and imaging devices, and provides calibration and synchronization methods for tracked ultrasound. It transmits tracked images to 3D Slicer through the OpenIGTLink protocol [20].

The interface for the portable Interson ultrasound probe was implemented in PLUS. Nevertheless, the system can be used with any of the ultrasound machines and trackers supported by PLUS, which makes reproducibility of the system convenient.

2.2 System Setup

The hardware described in previous section was integrated using the PLUS library. Configuration files to connect different acquisition systems are provided in PLUS website. In particular, sample configuration files used to connect the Interson Probe and the Micron Tracker can be found in the PLUS website.

[2] www.plustoolkit.org.

As with any TUS system, it requires a calibrated probe. If US imaging parameters are not changed, the calibration can be done once and saved in the configuration file. We found that 10 cm depth is adequate for this application.

When the patient comes in, at least one reference marker is attached to his back. This reference marker is used because the patient may move during the acquisition. The assumption is that the relative position between the transverse processes and the reference marker is the same during the whole scanning.

The tracker should be facing a wall where a reference marker must be attached. During the acquisition, the patient will be asked to place in front of the wall. The reason why a marker is attached to the wall is to be able to project the transverse processes positions to the wall. This makes the angle measurements compatible with the traditional procedure done on frontal X-ray radiographies.

2.3 Slicelet Implementation for Spinal Curvature Measurement

We designed and implemented a custom workflow and user interface, called a slicelet, based on the 3D Slicer application platform and its SlicerIGT extension.[3] Slicelets are custom user interfaces programmed for the 3D Slicer application that typically support a single clinical workflow without the complexity of the full 3D Slicer user interface. Our slicelet helps in the localization of vertebral transverse processes as landmarks along the spine to measure curvature angles. We build upon [16] and added new features and interfaces to streamline the procedure, reducing the lengthy analysis time.

The slicelet currently consists of the steps depicted in Fig. 2: communication with acquisition devices, tracked ultrasound snapshots and continuous tracked ultrasound video acquisition, localization of vertebral transverse processes in the acquired images and computation of angles between vertebrae.

2.3.1 Communication with Acquisition Devices

Images and tracker positions are provided by the PLUS server. Once the PLUS server is running, the status of the communication can be seen in the slicelet as shown in Fig. 3. A tool viewer, with a convenient color code, shows markers and tools visibility. This tool is very useful when working with optical trackers.

At this point of the procedure, it is important to check the working volume and also to make sure that the markers and the probe are visible in the region of interest. The wall reference marker does not need to be visible during the acquisition, but it is assumed that the relative position between the tracker and the wall does not change during patient scanning.

[3] www.slicerigt.org.

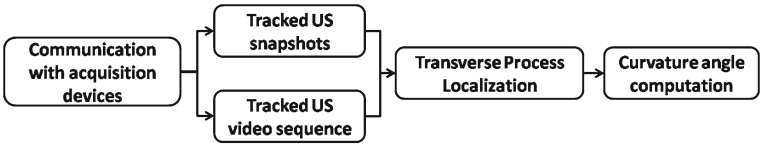

Fig. 2 Workflow of the slicelet

Fig. 3 The tool viewer shows the visibility of the markers. The spatial position of the US probe (Probe), three references markers (R1, R2 and R3) attached to the patient, a Stylus and a marker attached to the wall that faces the patient are acquired by the software. Currently, only the Probe, R1 and the Wall are used to compute the angle between the vertebrae but in the future we plan to explore the use of the others to improve the workflow

2.3.2 Tracked US Images Acquisition

The scoliosis monitoring slicelet can be used to measure vertebral angles using both: tracked ultrasound snapshots and video sequences. We decided to offer both possibilities because transverse process localization is easier with snapshots, while time occupied during patient scan is lower when video sequences are acquired. In both cases, data can be saved to be inspected afterwards and can also be used to build a database.

If the physician feels comfortable using snapshots, he must scan the patient until he finds a transverse process and then a snapshot is acquired with the aid of the table shown in Fig. 4. To reduce patient waiting time, a quick scan may be preferred. The software offers the possibility to acquire sequences for the left and right side of the spine. Those video sequences are stored, and used afterwards, to localize the images where the transverse processes are visible.

	Vertebrae	LTP Image	LTP Point	RTP Image	RTP Point
1	T1				
2	T2				
3	T3	Get Snapshot			
4	T4				
5	T5				

Fig. 4 The table is used to acquire the *Left Transverse Process* (LTP) and *Right Transverse Process* (RTP) images and to mark the LTP and RTP points

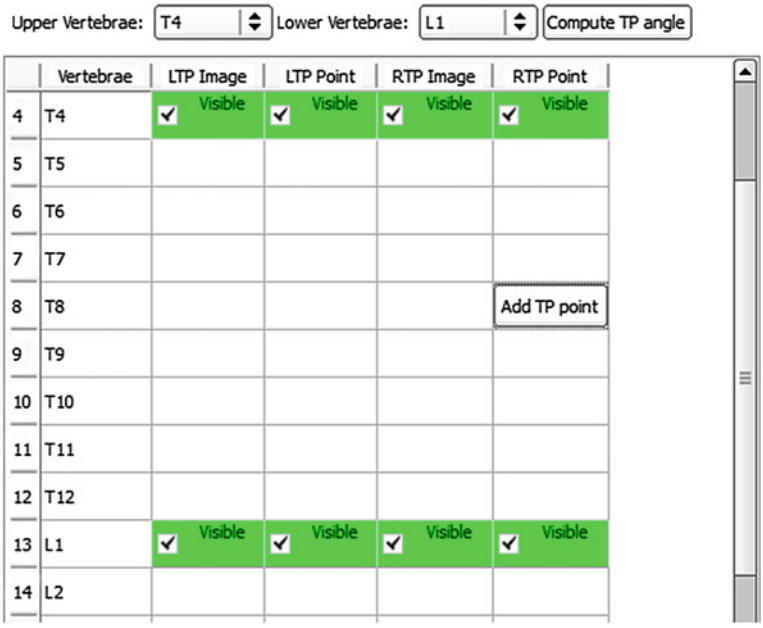

Fig. 5 When the transverse processes points that define a vertebra line are already acquired the corresponding row in the table is shown in *green*. When two lines are already defined is possible to compute the angle between them (Color figure online)

2.3.3 Transverse Processes Localization

Regardless of the method used to acquire the images, left and right transverse process must be manually marked in the corresponding images. When the left and right transverses processes of a vertebra are already located, the software automatically computes and shows the segment determined by the two transverse processes in a 3D Scene. When the transverse process points that define two vertebrae lines are already acquired, it is possible to compute the angle between them as shown in Fig. 5.

3 Curvature Angle Computation

The segments that determine the angle between two vertebrae are projected to the wall that faces the patient and then the angle between the projected segments is computed. The segments are projected to the wall to simulate the X-ray projection and to have an angle comparable with the Cobb angle.

4 Results

The workflow described in previous section has been implemented; it provides all of the required capabilities for scoliosis monitoring using tracked ultrasound snapshots and video sequences.

As an example, Fig. 6 shows the tracked ultrasound snapshots acquired during a procedure. Red points correspond to the transverse processes locations in the ultrasound images. The red lines connect two transverse process that belong to the same vertebra.

The software can be installed from the extension manager of the latest version of the 3D Slicer application. The extension manager is a plug-in mechanism that allows researchers and developers to add extensions to the core 3D application. Those extensions can be installed by the user after installing the main 3D application without any programming required. The name of the slicelet is Scoliosis Monitoring and is part of the Scoliosis extension. All software used in this study is open-source, freely available for research or commercial use without any restrictions. This allows incremental research, and focus work on unsolved problems like transverse process automatic segmentation or spine volume reconstruction.

The whole procedure can be tested with various ultrasound and tracker systems supported by PLUS. This includes the most popular tracking systems (Polaris, Micron, Aurora, Ascension) and the Interson USB and Ultrasonix ultrasound machine. Besides, Epiphan or Imaging Control frame grabber can be used to acquire the images of any ultrasound machines with video output.

The workflow was implemented with the portable USB ultrasound probe, which shows sufficient image quality for this application. The size and weight of the acqui-

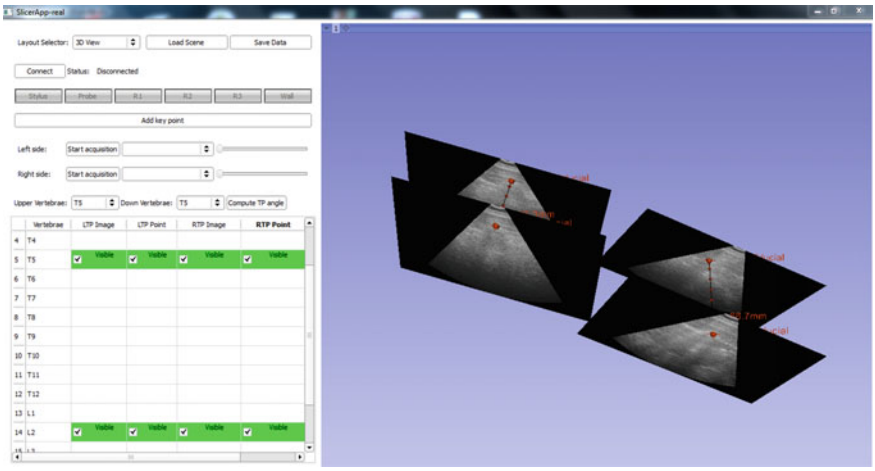

Fig. 6 The vertebral lines are shown in the 3D scene. In this case the transverse processes corresponding to L2 and T5 vertebrae were acquired

sition systems that were used allows to pack and move the system, which will make this a portable solution for scoliosis measurement in future clinical applications.

5 Discussion

Although some previous works have shown the feasibility of tracked ultrasound to localize vertebral processes to measure spinal curvature, the problem is still open. We believe we are sharing a valuable tool that can be very useful to show physicians the potential of the methods and also to identify the main challenges we need to tackle in order to bring scoliosis monitoring with ultrasound into routine clinical use.

The portable optically tracked ultrasound system can save ultrasound video sequences along with tracking data. This enables the acquisition of experimental studies from volunteers and patients looking forward to building the essential dataset for the research in the field. Currently, vertebral transverse processes are marked manually by the physician and the data is used to compute the curvature angle. This expert segmentation can also be saved in the application and this valuable information will enable testing future automatic segmentation algorithms.

The tool is integrated into the Slicer platform and the Slicer-IGT extension. This Slicer platform can be freely used by the research community and it enables an easy manipulation of the tracked ultrasound studies and the integration with other modalities such as X-ray, CT or MRI. Future developments in the field can also be rapidly integrated into Slicer as extensions (its plug-in mechanism) and shared through the Slicer app-store.

Our next step in this ongoing project will be to study the correlation between the radiography and sonographic curvature measuring methods in real patients. Another future research line will be related to the segmentation of the transverse processes. In some patients it is hard for the physician to correctly identify some transverse processes. We plan to explore the feasibility of using tracked ultrasound volume reconstructions of the spine in conjunctions with models of the spine. A nice 3D model of the spine might give the physician some spatial information that could greatly facilitate the transverse process localization procedure.

Acknowledgments This work was supported in part by the Canada-Latin America Research Exchange Grant (LACREG) program. Comision Sectorial de Investigacion Cientıfica (CSIC, Universidad de la Republica, Uruguay) supported the internship of G. Carbajal at Perklab.

References

1. Nery, L.S., Halpern, R., Nery, P.C., Nehme, K.P., Stein, A.T.: Prevalence of scoliosis among school students in a town in southern Brazil. Sao Paulo Med. J. **128**(2), 69–73 (2010)
2. Suh, S.W., Modi, H.N., Yang, J.H., Hong, J.Y.: Idiopathic scoliosis in Korean schoolchildren: a prospective screening study of over 1 million children. Eur. Spine J. **20**(7), 1087–1094 (2011)

3. Wong, H.K., Hui, J.H., Rajan, U., Chia, H.P.: Idiopathic scoliosis in Singapore schoolchildren: a prevalence study 15 years into the screening program. Spine (Phila Pa 1976). **30**(10), 1188–96 (2005)
4. Ward, K., Ogilvie, J.W., Singleton, M.V., Chettier, R., Engler, G., Nelson, L.M.: Validation of DNA-based prognostic testing to predict spinal curve progression in adolescent idiopathic scoliosis. Spine (Phila Pa 1976). **35**(25), E1455–64 (2010)
5. Aartun, E., Degerfalk, A., Kentsdotter, L., Hestbaek, L.: Screening of the spine in adolescents: inter- and intra-rater reliability and measurement error of commonly used clinical tests. BMC Musculoskelet. Disord. **10**(15), 37 (2014)
6. Doody, M.M., Lonstein, J.E., Stovall, M., Hacker, D.G., Luckyanov, N., Land, C.E.: Breast cancer mortality after diagnostic radiography: findings from the U.S. Scoliosis Cohort study. Spine (Phila Pa 1976). **25**(16), 2052–63 (2000)
7. Hoffman, D.A., Lonstein, J.E., Morin, M.M., Visscher, W., Harris 3rd, B.S., Boice Jr, J.D.: Breast cancer in women with scoliosis exposed to multiple diagnostic x rays. J. Natl. Cancer Inst. **81**(17), 1307–1312 (1989)
8. Schmitz-Feuerhake, I., Pflugbeil, S.: 'Lifestyle' and cancer rates in former East and West Germany: the possible contribution of diagnostic radiation exposures. Radiat. Prot. Dosim. **147**(1–2), 310–313 (2011)
9. Diefenbach, C., Lonner, B.S., Auerbach, J.D., Bharucha, N., Dean, L.E.: Is radiation-free diagnostic monitoring of adolescent idiopathic scoliosis feasible using upright positional magnetic resonance imaging? Spine (Phila Pa 1976). **38**(7), 576–80 (2013)
10. Goldberg, C.J., Kaliszer, M., Moore, D.P., Fogarty, E.E., Dowling, F.E.: Surface topography, Cobb angles, and cosmetic change in scoliosis. Spine (Phila Pa 1976). **26**(4), E55–63 (2001)
11. Frerich, J.M., Hertzler, K., Knott, P., Mardjetko, S.: Comparison of radiographic and surface topography measurements in adolescents with idiopathic scoliosis. Open Orthop. J. **6**, 261–265 (2012)
12. Suzuki, S., Yamamuro, T., Shikata, J., Shimizu, K., Iida, H.: Ultrasound measurement of vertebral rotation in idiopathic scoliosis. J. Bone Jt. Surg. Br. **71**(2), 252–255 (1989)
13. Li, M., Cheng, J., Ying, M., Ng, B., Zheng, Y.P., Lam, T.P., Wong, W.Y., Wong, M.S.: Application of 3-D ultrasound in assisting the fitting procedure of spinal orthosis to patients with adolescent idiopathic scoliosis. Stud. Health Technol. Inform. **158**, 34–37 (2010)
14. Chen, W., Lou, E.H., Le, L.H.: Using ultrasound imaging to identify landmarks in vertebra models to assess spinal deformity. Conf. Proc. IEEE Eng. Med. Biol. Soc. **2011**, 8495–8498 (2011)
15. Purnama, K.E., Wilkinson, M.H., Veldhuizen, A.G., van Ooijen, P.M., Lubbers, J., Burgerhof, J.G., Sardjono, T.A., Verkerke, G.J.: A framework for human spine imaging using a freehand 3D ultrasound system. Technol. Health Care **18**(1), 1–17 (2010)
16. Ungi, T., King, F., Kempston, M., Keri, Z., Lasso, A., Mousavi, P., Rudan, J., Borschneck, D.P., Fichtinger, G.: Spinal curvature measurement by tracked ultrasound snapshots. Ultrasound Med. Biol. **40**(2), 447–454 (2014)
17. Ungi, T., Abolmaesumi, P., Jalal, R., Welch, M., Ayukawa, I., Nagpal, S., Lasso, A., Jaeger, M., Borschneck, D.P., Fichtinger, G., Mousavi, P.: Spinal needle navigation by tracked ultrasound snapshots. IEEE Trans. Biomed. Eng. **59**(10), 2766–2772 (2012)
18. Ungi, T., Moult, E., Schwab, J.H., Fichtinger, G.: Tracked ultrasound snapshots in percutaneous pedicle screw placement navigation: a feasibility study. Clin. Orthop. Relat. Res. **471**(12), 4047–4055 (2013)
19. Lasso, A., Heffter, T., Rankin, A., Pinter, C., Ungi, T., Fichtinger, G.: PLUS: open-source toolkit for ultrasound-guided intervention systems. IEEE Trans. Biomed. Eng. (In press)
20. Tokuda, J., Fischer, G.S., Papademetris, X., Yaniv, Z., Ibanez, L., Cheng, P., Liu, H., Blevins, J., Arata, J., Golby, A.J., Kapur, T., Pieper, S., Burdette, E.C., Fichtinger, G., Tempany, C.M., Hata, N.: OpenIGTLink: an open network protocol for image-guided therapy environment. Int. J. Med. Robot. **5**(4), 423–434 (2009)

Part II
Spine Segmentation

Atlas-Based Registration for Accurate Segmentation of Thoracic and Lumbar Vertebrae in CT Data

Daniel Forsberg

Abstract Segmentation of the vertebrae in the spine is of relevance to many medical applications related to the spine. This paper describes a method based upon atlas-based registration for achieving an accurate segmentation of the thoracic and the lumbar vertebrae in the spine as imaged by computed tomography. The method has been evaluated on ten data sets provided as a part of the segmentation challenge hosted by the 2nd MICCAI workshop on Computational Methods and Clinical Applications for Spine Imaging. An average point-to-surface error of 1.05 ± 0.65 mm and a mean DICE coefficient of 0.94 ± 0.03 were obtained when comparing the computed segmentations with ground truth segmentations. These results are highly competitive when compared to the results of earlier presented methods.

1 Introduction

Segmentation of organs in the human body is one of the most important problems in medical image computing. The challenges associated with medical image segmentation, and what is causing further research to be motivated within the domain, include low contrast between organ of interest and surrounding tissues, imaging artifacts, image noise and anatomical variation. There exists many different approaches for how a segmentation problem can be solved. Proposed methods vary from simple thresholding to more elaborate methods using e.g. probabilistic shape models. The relevance of the different approaches depend on the specific organ to segment along with the employed imaging modality. This work is focused on the segmentation of the vertebrae in the spine as imaged by computed tomography (CT).

The spinal column forms an important support structure in the human body and consists of the vertebral bones, seven cervical, twelve thoracic and five lumbar ver-

D. Forsberg (✉)
Center for Medical Image Science and Visualization (CMIV),
Linköping University, Linköping University Hospital, SE-581 85, Linköping, Sweden
e-mail: daniel.forsberg@sectra.se

Sectra, Teknikringen 20, SE-583 30, Linköping, Sweden

© Springer International Publishing Switzerland 2015

J. Yao et al. (eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, Lecture Notes in Computational Vision and Biomechanics 20, DOI 10.1007/978-3-319-14148-0_5

tebrae. As such, the vertebrae form an important part of the diagnosis, treatment planing and the understanding of various conditions affecting the spine. Thus, an accurate segmentation of the vertebrae is of relevance in several medical applications. Segmentation of the vertebrae is challenging, mainly due to shape variation and neighboring structures of similar intensity (e.g. other vertebrae and/or tissues).

There already exists a number of proposed methods for segmentation of the vertebrae in the spine from CT data. For example, in [9], Kim and Kim proposed a fully automatic method based upon deformable fences for separating lumbar vertebra in the spinal column. Klinder et al. presented in [10] a method, commonly referred to as the current state-of-the-art for both identification and segmentation of an arbitrary set of vertebrae. The method is based upon spinal curve extraction with vertebra detection followed by vertebra identification and segmentation based upon shape models. Ma et al. [13] also proposed a method capable of both segmentation and identification, but only for the thoracic vertebrae. Their method is based upon training bone-structure edge detectors along with coarse-to-fine registration of a deformable model. Other and more recently proposed methods are found in [8, 12, 14], where the authors of the first work propose an improved level set method based upon edge- and region-information. The authors of the second work also propose to use a level set method but including the Willmore flow and where the third work presents a method based on a multi-vertebrae anatomical shape and pose model. Note that all of these three methods were only evaluated segmenting the lumbar vertebrae.

In this paper, a method based upon atlas-based segmentation is proposed and evaluated. Atlas-based segmentation is a method where image registration between an atlas image and a target image is performed and where the predefined labels of the atlas are used for segmenting the region of interest in the target image. Atlas-based segmentation has found its way into many application domains within medical imaging and in particular for segmentation of various structures in the brain in magnetic resonance imaging [2]. In atlas-based segmentation, it is common to use a number of atlases that are registered towards a target image and where label fusion is used to derive a single segmentation from the labels of the registered atlases.

The method, described in this paper, benefits from the knowledge of previous work in term of preprocessing (vertebral pose estimation), spine registration (non-rigid registration based on phase-difference and the use of graphics processing units) and label fusion (majority voting), and combines these into an effective pipeline providing automatic and accurate segmentation of both lumbar and thoracic vertebrae in the spinal column.

2 Methods

The method used in this work for vertebrae segmentation is inspired and to a large extent based upon the work presented in [6, 7], although some components have changed and others have been added. This has been done in order to improve

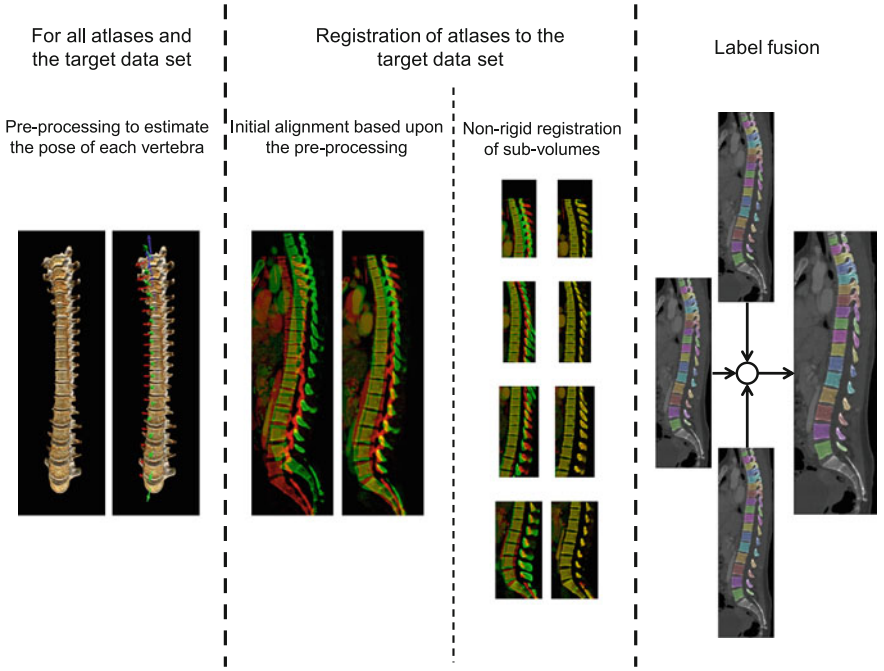

Fig. 1 An overview of the method used for segmentation of the thoracic and lumbar vertebrae in the spine. In the middle section of this figure, the *green-colored* spine refers to the atlas and the *red-colored* to the target data set (Color figure online)

the performance but mainly since the work in [6, 7] was targeted at scoliotic spines. The most notable differences are the use of multiple gray-level atlases, instead of a single binary model in the registration step, and the subsequent use of label fusion. The employed method consists of a preprocessing step, where an approximate position and rotation (pose) of each vertebra in the target data set and in the atlases is estimated. The results from the preprocessing are used to obtain an initial alignment between the atlases and the target data set. The preprocessing step is followed by a registration step, where each atlas is registered to the target data set. The labels of the registered atlases are combined to a single label volume using label fusion to form the segmentation of the spine as imaged in the target data set. A graphical overview of the method is provided in Fig. 1.

2.1 Preprocessing

The preprocessing consists of the following sub-steps:

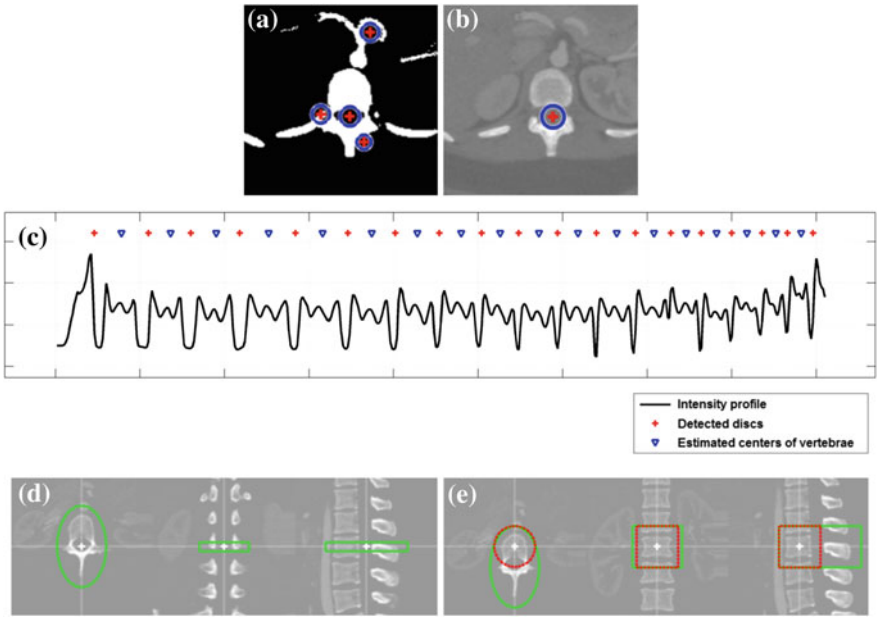

Fig. 2 The preprocessing steps starts with a seed point detection (**a**) for the tracking of the spinal canal (**b**). Given the spinal canal, an intensity profile running through the vertebrae is sampled and analyzed to detect the position of the discs and implicitly the mid-axial slices of the vertebra (**c**). The mid-axial slices with initial estimates of the axial vertebral rotation (**d**) provide input to the final pose estimations (**e**). Note that the marked regions in (**d**) and (**e**) correspond to the volumes used for symmetry assessment

1. **Spinal canal tracking**—Seed points for the spinal canal are detected using the Hough transform on a thresholded axial image in the middle of the image volume. A growing and moving circle is used to detect the center of the spinal canal, i.e. a small circle is initialized at a given seed point and the circle grows until it hits bone upon which it tries to move away from the bone, in order to grow further. This growing and moving process is iterated until the circle either cannot grow or move anymore without hitting bone. The growing and moving circle process is repeated for each image as the spinal canal is tracked in both the cranial and the caudal direction. A minimum and a maximum size of the circle is defined along with exit criteria to ensure that the tracking stops as the sacrum or the top cervical vertebrae are reached. A spline curve is fitted to the extracted centerline in order to obtain a smooth curve. Seed point detection and tracking of the spinal canal are depicted in Fig. 2a, b.
2. **Disc detection**—Given that the vertebrae are located anterior to the spinal canal, an intensity profile, running through the vertebrae, is sampled and filtered with quadrature filters of varying center frequencies. Given the distinctive pattern of the intensity profile, the filtered intensity profiles can be used to detect the position of the discs, simply by searching for peaks. Some heuristics have been added to

remove false positive peaks. An example of the characteristic intensity profile is given in Fig. 2c.

3. **Initial vertebral rotation estimation**—In an image slice located between the detected discs, an initial axial vertebral rotation is estimated based upon minimizing an error measure for assessing the symmetry between two halves of an image (left and right halves). The center of rotation is set to the center-point of the spinal canal and the orientation of the image plane is given by the tangent of the spinal canal centerline. For a more robust rotation estimation, not just a single image is used but a number of images. The volume enclosed by the green marking in Fig. 2d, refers to region for which the lateral symmetry is maximized.
4. **Vertebra position and rotation estimation**—The two previous steps provide an initial estimate of the position and the orientation of each vertebra. To improve this estimate, an error measure is defined with six parameters $[x, y, z]$ and $[\theta_X, \theta_Y, \theta_Z]$, defining the vertebra center-point and the rotation of the vertebra. The error measure is defined to assess the symmetry across various half-planes. The optimal parameters are found using Powell's method, as provided by the MATLAB package iFit [4]. This step is similar to the method used in [15] for estimating the center-point and the orientation of each vertebra in the spine. The green marked region, in Fig. 2e, is used to assess the lateral symmetry, and where the red marked region is used to assess anterior-posterior symmetry and the inferior-superior symmetry.

Note that all atlases and the target data set are processed with this preprocessing step. Parameters used in the preprocessing step were fixed for the processing of all data sets but were empirically determined based upon initial processing of one of the data sets.

2.2 Initial Alignment

The estimated poses are used for establishing an initial alignment between the vertebrae of the atlases and the vertebrae of the spine in the target data set. The alignment is achieved by translating one atlas according to the mean displacement between the respective vertebrae in the data sets, i.e. the displacement between vertebrae T1 in the target and in the atlas data sets, and between T2 in the target and in the atlas data sets and so forth. Furthermore, a scaling of the atlas data sets is included by using the ratio between the two T1-L5 distances in the target and the atlas data sets.

2.3 Atlas-Based Registration

Given the initial alignment between the vertebrae of the target data set and an atlas, the vertebrae are then processed in groups of five, starting from the caudal end of the

spine, i.e. L5-L1, L1-T9, T9-T5, T5-T1. A sub-volume is extracted from both the target data set and the atlas at hand, containing the vertebrae to be registered along with any surrounding vertebrae. After the initial alignment, a non-rigid registration step is applied, minimizing the voxel-wise local phase-difference [11], according to

$$\mathbf{u} = \arg\min_{\mathbf{d}} \sum_k \frac{1}{2} \| c_k \mathbf{T} \left(d_k \hat{\mathbf{n}}_k - \mathbf{d} \right) \|^2, \tag{1}$$

where $\mathbf{T}$ is the local structure tensor and d_k the local phase-difference in orientation $\hat{\mathbf{n}}_k$. For computational performance, an implementation of the non-rigid registration algorithm on graphics processing units has been used [5]. The finally computed transformation is used to deform the atlas onto the target data set, and, thus, the labels of the deformed atlas provide a segmentation of the vertebrae in the target data set. This process is repeated for each available atlas.

2.4 Label Fusion

The final step is to merge the labels of the deformed atlases into a single label volume. In this case, a straight forward majority voting has been employed for label fusion, although more refined methods for label fusion are available, e.g. STAPLE [16].

2.5 Implementation

The method has been implemented in MATLAB and CUDA, and is available at https://bitbucket.org/dforsberg/ for download.

3 Data

The data used for evaluation consists of ten CT data sets acquired during daily clinical routine work at a trauma center, and is provided as training data for the segmentation challenge at the 2nd MICCAI Workshop on Computational Methods and Clinical Applications for Spine Imaging. The included patients are 16–35 years old and the scans covers all thoracic and lumbar vertebrae. The resolution of the data ranges between 0.31 and 0.45 mm for the in-plane resolution and the slice thickness is 1 mm. Along with the image data there also exists ground truth data consisting of semi-automatically segmented vertebrae. More information about the image data is found in [17]. Before processing, the image data was resampled to an isotropic resolution of $1 \times 1 \times 1\,\text{mm}^3$ using linear interpolation, in order to allow the usage of quadrature filters with isotropic resolution in the non-rigid registration step.

Since ten data sets were available, each data set was segmented using the remaining nine as atlases.

The ground truth data was compared with the segmentations obtained from the atlas-based segmentation using the point-to-surface error (PSE) and the DICE coefficient. The PSE is defined as the smallest distance between a node on the computed segmentation and the closest surface of the ground truth segmentation. The meshes defining the vertebrae were acquired from their corresponding binary volumes using `iso2mesh` [3]. The DICE coefficient is defined as

$$DICE = \frac{2 * |GT \cap S|}{|GT| + |S|} \tag{2}$$

where GT and S refer to the ground truth and the computed segmentations respectively, and where $|\ldots|$ denotes the volume in voxels.

4 Results

The results from the atlas-based segmentation of the ten CT data sets are provided in Table 1. In the results, it can be noted that the lumbar vertebrae and the lower thoracic vertebrae in general obtain a PSE < 0.90 and a DICE coefficient of ≥ 0.95, which is

Table 1 Average PSE and average DICE coefficient as estimated from the atlas-based segmentation of the ten data sets

Vertebra	PSE	DICE
T1	1.37 ± 1.15	0.86 ± 0.09
T2	0.89 ± 0.35	0.94 ± 0.00
T3	0.90 ± 0.37	0.93 ± 0.03
T4	1.27 ± 1.07	0.90 ± 0.09
T5	1.66 ± 1.53	0.88 ± 0.11
T6	1.78 ± 1.61	0.86 ± 0.17
T7	1.57 ± 1.57	0.88 ± 0.17
T8	1.10 ± 0.74	0.91 ± 0.10
T9	0.96 ± 0.52	0.94 ± 0.05
T10	0.89 ± 0.35	0.95 ± 0.02
T11	0.87 ± 0.32	0.95 ± 0.01
T12	0.87 ± 0.34	0.95 ± 0.00
L1	0.86 ± 0.32	0.96 ± 0.00
L2	0.86 ± 0.36	0.96 ± 0.01
L3	0.87 ± 0.37	0.96 ± 0.01
L4	0.89 ± 0.46	0.96 ± 0.00
L5	0.93 ± 0.53	0.95 ± 0.01
All	1.05 ± 0.65	0.94 ± 0.03

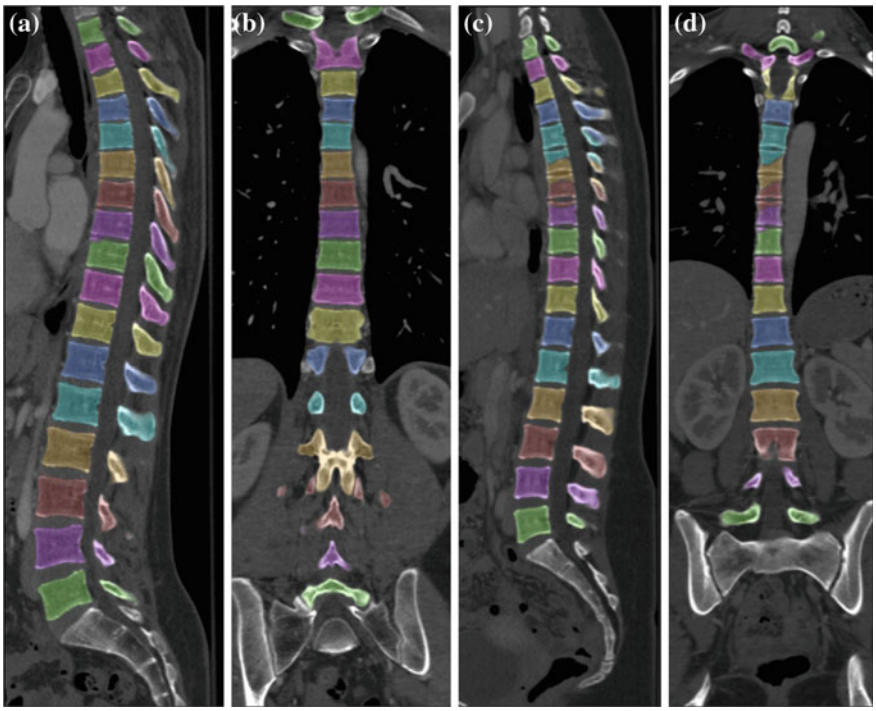

Fig. 3 Example results of the atlas-based registration, where (**a**) and (**b**) depict the results for data set #6 and (**c**) and (**d**) for data set #8, i.e. the data sets with the best respectively the worst results for the DICE coefficient

to be considered as excellent. Good scores are also achieved for the remaining thoracic vertebrae but where there are some vertebrae that obtain notably worse results. When observing the results data set by data set, it is apparent that the segmentation failed for for vertebrae T5-T8 in data sets #4 and #8. In addition, vertebra T1 obtains, almost consistently, lower scores than the other vertebrae. Example visualizations of the results are provided in Fig. 3. Note the failed segmentations of vertebrae T5-T8 in Fig. 3c, d, which correspond to the segmentation results for data set #8.

A run time analysis was also conducted for the proposed method. The results of this analysis are given in Table 2 and show that a complete segmentation of all

Table 2 Average run time for the proposed method when segmenting a single data set, i.e. the results correspond to the average time of preprocessing a single target data set, registering nine atlases to a target data set and fusing nine label volumes to a single label volume

Preprocessing (s)	Non-rigid registration (s)	Label fusion (s)	Total (s)
191 ± 16	509 ± 25	19 ± 5	718 ± 40

thoracic and lumbar vertebrae can be achieved within 12 min. Note that the timing results for the preprocessing step correspond to the preprocessing of a single data set. In addition, time needed for I/O and similar activities along with the resampling of the original data to a $1 \times 1 \times 1\,\text{mm}^3$ resolution has not been included in the results.

5 Discussion

In this paper, a method for segmentation of the thoracic and the lumbar vertebrae has been described and evaluated. The method is based on atlas-based segmentation and makes use of a preprocessing step for estimating the local pose of each vertebrae. The preprocessing step is important for the initial alignment of the atlases before the registration of the atlases to the target data set. This is because the repetitive structure of the vertebrae in the inferior-superior direction easily can cause a mis-alignment of adjacent vertebrae.

The evaluation of the described method on ten data sets achieved an average PSE of 1.05 ± 0.65 mm and a mean DICE coefficient of 0.94 ± 0.03. Note that both scores could be expected to be better if not the method had failed for a few of the thoracic vertebrae in two of the data sets. As noted earlier, the registration accuracy for vertebra T1 is almost consistently lower than the other vertebrae. This is most likely due to the fact that in some data sets vertebra T1 is the top vertebra being imaged, while some other data sets contain some of the cervical vertebrae as well. Because of this, the registration process sometimes extends vertebra T1 to vertebra C7 or similar. This problem is not present at the other end of the spinal column, i.e. for vertebra L5, since the sacrum is included in all data sets and, thus, creates an effective barrier for making sure that vertebra L5 is not mis-aligned with S5 in the sacrum.

In comparison with related work, the presented results can be considered as good to excellent. For example, Klinder et al. [10] achieves a mean PSE of 1.12 ± 1.04 mm for their method. Ma et al. [13] obtain a slightly better mean PSE of 0.95 ± 0.91 mm for their algorithm, and where their method was only evaluated for the thoracic vertebrae. Rasoulian et al. [14], on the other hand, only segmented lumbar vertebrae and obtained an average error distance of 1.38 ± 0.56 between the fitted model and the manual segmentations. Huang et al. [8] and Lim et al. [12] are the only that report the DICE coefficient in their results, and where Huang et al. achieved an average DICE score of 0.94 ± 0.02 and Lim et al. an average DICE score of 0.89 ± 0.02. Note that both methods are only evaluated on lumbar vertebrae. In addition, the method of Huang et al. segments the vertebrae slice by slice and then fuses the separate 2D segmentations to form a 3D segmentation. Hence, they only segment out a single 3D object and are unable to separate adjacent vertebrae. Kim and Kim [9] does not provide any results useful for comparison, since their assessment is only based on a visual grading of how well the lumbar vertebrae are separated.

In terms of run time, it is to be noted that the time required for the registration step is directly related to the number of atlases that are used, i.e. decreasing or increasing

the number of atlases will correspondingly decrease or increase the time needed for the registration. Comparing run times of different methods is difficult given that hardware architectures, implementation methods and data sizes differ too much. In addition, some report the run times per slice, some per vertebra and some per data set. Hence, here it is simply noted that the run times reported by others correspond to 1–6 s per slice [8], 1 min per vertebra [14] or approximately 40 min for all thoracic vertebrae [10].

A limitation of the results presented here refers to the fact that the segmented data sets did not include any pathological spines or any notable imaging artifacts. In terms of larger spinal deformities, e.g. as such induced by idiopathic scoliosis, it was shown in [6] that a similar approach as described in this paper was able to achieve an average PSE of 0.9 ± 0.9 mm. Thus, for larger spinal deformities the segmentation accuracy of the proposed method is likely to be similar as presented here. How other issues, such as imaging artifacts from e.g. metal implants, will affect the results are difficult to foresee and further speculations are, thus, refrained from.

Future work for improving the performance of the method includes making better use of the estimated poses for each vertebra, since the current poses are only used to create a translation and a scaling between the target data set and the atlases. One approach would be to use the estimated poses to compute a linear transformation for each vertebra between the target data set and an atlas, and to combine these into a single transformation using the Log-Euclidean framework as presented by Arsigny et al. [1]. In addition, it would be relevant to investigate other methods for label fusion, e.g. STAPLE as already mentioned, and to add a step of selecting which atlases to include in the label fusion, i.e. a best atlas selection step.

References

1. Arsigny, V., Commowick, O., Ayache, N., Pennec, X.: A fast and Log-Euclidean polyaffine framework for locally linear registration. J. Math. Image. Vis. **33**(2), 222–238 (2009)
2. Cabezas, M., Oliver, A., Lladó, X., Freixenet, J., Bach Cuadra, M.: A review of Atlas-based segmentation for magnetic resonance brain images. Comput. Methods Progr. Biomed. **104**(3), e158–e177 (2011)
3. Fang, Q., Boas, D.A.: Tetrahedral mesh generation from volumetric binary and grayscale images. In: Biomedical Imaging: From Nano to Macro, 2009. ISBI'09. IEEE International Symposium on, pp. 1142–1145. IEEE (2009). doi:10.1109/ISBI.2009.5193259
4. Farhi, E., Debab, Y., Willendrup, P.: iFit: A new data analysis framework, applications for data reduction and optimization of neutron scattering instrument simulations with mcstas. J. Neutron Res. **17**(1), 5–18 (2014)
5. Forsberg, D., Eklund, A., Andersson, M., Knutsson, H.: Phase-based non-rigid 3D image registration—from minutes to seconds using CUDA. In: HP-MICCAI/MICCAI-DCI 2011 (2011)
6. Forsberg, D., Lundström, C., Andersson, M., Knutsson, H.: Model-based registration for assessment of spinal deformities in idiopathic scoliosis. Phys. Med. Biol. **59**(2), 311–326 (2014)
7. Forsberg, D., Lundström, C., Andersson, M., Vavruch, L., Tropp, H., Knutsson, H.: Fully automatic measurements of axial vertebral rotation for assessment of spinal deformity in idiopathic scoliosis. Phys. Med. Biol. **58**(6), 1775–1787 (2013)

8. Huang, J., Jian, F., Wu, H., Li, H.: An improved level set method for vertebra CT image segmentation. Biomed. Eng. Online **12**(1), 48 (2013)
9. Kim, Y., Kim, D.: A fully automatic vertebra segmentation method using 3D deformable fences. Comp. Med. Imag. Graph. **33**(5), 343–352 (2009)
10. Klinder, T., Ostermann, J., Ehm, M., Franz, A., Kneser, R., Lorenz, C.: Automated model-based vertebra detection, identification, and segmentation in CT images. Med. Image Anal. **13**(3), 471–482 (2009)
11. Knutsson, H., Andersson, M.: Morphons: Segmentation using elastic canvas and paint on priors. In: Image Processing (ICIP), 2005 IEEE International Conference on, pp. II-1226-9. IEEE (2005). doi:10.1109/ICIP.2005.1530283
12. Lim, P.H., Bagci, U., Bai, L.: Introducing Willmore flow into level set segmentation of spinal vertebrae. Biomed. Eng. IEEE Trans. **60**(1), 115–122 (2013)
13. Ma, J., Lu, L., Zhan, Y., Zhou, X., Salganicoff, M., Krishnan, A.: 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, Lecture Notes in Computer Science, vol. 6361, pp. 19–27. Springer (2010)
14. Rasoulian, A., Rohling, R., Abolmaesumi, P.: Lumbar spine segmentation using a statistical multi-vertebrae anatomical shape+pose model. Med. Imag. IEEE Trans. **32**(10), 1890–1900 (2013)
15. Vrtovec, T.: Modality-independent determination of vertebral position and rotation in 3D. In: Medical Imaging and Augmented Reality, Lecture Notes in Computer Science, vol. 5128, pp. 89–97. Springer (2008)
16. Warfield, S.K., Zou, K.H., Wells, W.M.: Simultaneous truth and performance level estimation (STAPLE): an algorithm for the validation of image segmentation. Med. Imag. IEEE Trans. **23**(7), 903–921 (2004)
17. Yao, J., Burns, J., Munoz, H., Summers, R.: Detection of vertebral body fractures based on cortical shell unwrapping. In: Medical Image Computing and Computer-Assisted Intervention MICCAI 2012, Lecture Notes in Computer Science, vol. 7512, pp. 509–516. Springer (2012)

Segmentation of Lumbar Vertebrae Slices from CT Images

Hugo Hutt, Richard Everson and Judith Meakin

Abstract We describe a fully automated approach to vertebrae segmentation from CT images which operates on superpixels. The method is based on a conditional random field model incorporating constraints learned from labelled superpixel features. The method is shown to provide consistently accurate segmentations of different vertebrae from a variety of subjects.

1 Introduction

Automatic segmentation of vertebrae from CT images is a challenging problem due to the complex and varied shape of the vertebrae, in addition to the various artefacts which may result from the acquisition process. However, segmenting the vertebrae by hand is a difficult and time consuming process. Automated segmentation is therefore desired to obtain reliable and accurate segmentations on any large scale.

Much of the previous work in this area has concentrated on sagittal views to provide segmentation of many vertebrae and intervening discs. With this view the pedicles and posterior elements of the vertebrae are frequently not visible, so segmentation has focussed on the vertebral bodies and employed tools such as statistical shape models and appearance models combined with probabilistic graphical models; e.g., [7, 11]. Huang et al. [8] have recently described a level set method for vertebrae segmentation using transverse (axial) CT slices.

In this paper, we describe a fully automated segmentation method that effectively segments the whole vertebra structure (including pedicles and posterior elements)

H. Hutt (✉) · R. Everson · J. Meakin
University of Exeter, Exeter, UK
e-mail: hwh202@exeter.ac.uk

R. Everson
e-mail: R.M.Everson@exeter.ac.uk

J. Meakin
e-mail: J.R.Meakin@exeter.ac.uk

© Springer International Publishing Switzerland 2015

J. Yao et al. (eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, Lecture Notes in Computational Vision and Biomechanics 20, DOI 10.1007/978-3-319-14148-0_6

from transverse (axial) views. Our method for CT images is adapted from a method developed for MR images, described in [9]. At heart, the method uses a conditional random field (CRF) model on superpixels. Operating on superpixels reduces computational complexity and enables more descriptive features to be extracted to characterise the vertebra (foreground) and non-vertebra (background) classes, while the CRF relates the underlying class labels of the superpixels to the observed features and promotes coherence. We use supervised learning to train a classifier on labelled superpixel features and obtain probability estimates expressing the likelihood of belonging to either the vertebra or background class. Distance metric learning [17] is also used to find an appropriate dissimilarity measure between superpixel pairs. The probability estimates and learned distance metric are incorporated into the CRF model in the form of first- and second-order clique potentials of the CRF energy function. This formulation enables minimisation of the energy function to be carried out efficiently using graph cuts [3].

We evaluate the performance of the method on CT data from a range of subjects collected for the Computational Methods and Clinical Applications for Spine Imaging (CSI 2014) segmentation competition. We show that consistently accurate segmentations can be obtained for each of the different lumbar vertebrae.

2 Segmentation Model

Our method is based on a conditional random field (CRF) [2] model which operates on the superpixels of an image; we denote the set of superpixels by $\mathcal{S}$. The energy function of the CRF defines a posterior probability distribution $P(\mathbf{x} \mid \mathbf{y})$ for a set of class labels $\mathbf{x}$ for the superpixels, given a set of features $\mathbf{y}$ describing the superpixels. The energy function can be written as a sum of first- and second-order potential functions in the form

$$E(\mathbf{x}, \mathbf{y}) = \sum_{i \in \mathcal{S}} \underbrace{\psi(\mathbf{y}_i \mid x_i)}_{\text{Data term}} + \lambda \sum_{i \in \mathcal{S}} \sum_{j \in \mathcal{N}_i} \underbrace{\phi(\mathbf{y}_i, \mathbf{y}_j \mid x_i, x_j)}_{\text{Smoothness term}} \tag{1}$$

where $\mathcal{N}_i$ is the set of neighbours of superpixel i. The constant λ controls the relative importance of the data and smoothness terms. The CRF formulation enables maximum a posteriori (MAP) inference of the labels $\mathbf{x}$ to be carried out efficiently using graph cuts. We use the min-cut/max-flow algorithm of [4] to find the optimal solution.

We define the potential functions of (1) by using supervised learning on labelled superpixel features and deriving constraints using the resulting trained models. Sections 3 and 4 describe the superpixel features used to learn the constraints and how they are incorporated into the CRF potential functions.

3 Superpixels

We use the Simple Linear Iterative Clustering (SLIC) [1, 16] algorithm to partition the image into superpixels. As shown in Fig. 1, boundaries of superpixels tend to coincide with boundaries of anatomical objects, enabling an accurate pixel-level segmentation to be recovered from the classified superpixels. The primary advantages of using superpixels are twofold: firstly, as the number of nodes in the graph decreases significantly from a pixel-level graph, there is a corresponding reduction in computational complexity. Secondly, multiple features can be extracted from the superpixel regions which can help to discriminate between the classes more effectively.

We aim to characterise the superpixels by extracting multiple features from them that incorporate information about intensity, texture, location and edge response. As described in the next section, these features are used to discriminate between the vertebra and background superpixels by learning a classifier and distance metric on a set of ground truth images. We emphasise that this training occurs only once, after which the trained models can be used in the CRF potential functions for any further images.

The superpixel features are summarised in Table 1. The feature vector for a superpixel i is a concatenation of the individual features:

$$\mathbf{y}_i = [\mathbf{y}_i^T, \mathbf{y}_i^L, \mathbf{y}_i^E]^\top. \tag{2}$$

We exhaustively tested different subsets of the features, but found that the best performance was obtained by combining all features. The features were chosen in part

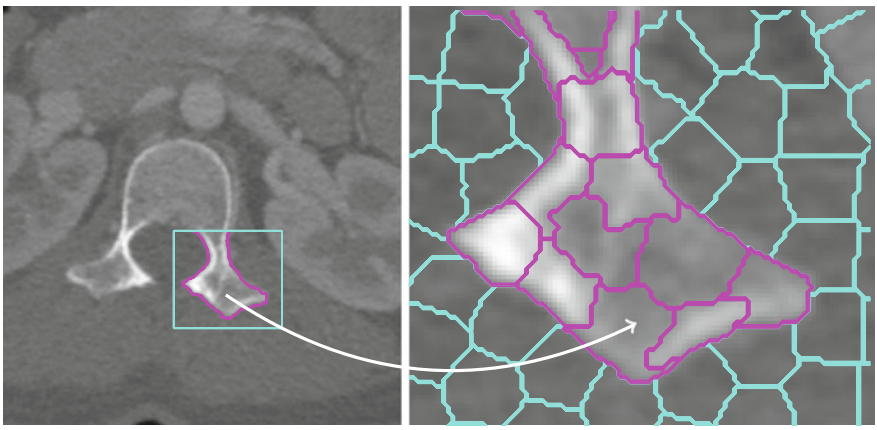

Fig. 1 The *left* figure shows a CT slice with ground truth contour (*magenta*) for a section of the vertebra. The *right* figure shows boundaries for superpixels assigned to the vertebra class (*magenta*) and background class (*cyan*). The superpixels preserve the boundary detail of the vertebrae (color figure online)

Table 1 Superpixel features (p_n denotes the nth percentile)

Feature	Description	Dimension
$\mathbf{y}_i^{T_1}$	Concatenation of intensity histogram from superpixel i and average histogram from neighbours $\mathcal{N}_i$	20
$\mathbf{y}_i^{T_2}$	SIFT descriptor calculated at the centroid of superpixel i	128
$\mathbf{y}_i^{L_1}$	Mean, p_{10} and p_{90} of the row and column pixel coordinates in the superpixel, centred on the matched contour region	6
$\mathbf{y}_i^{L_2}$	Mean, p_{10} and p_{90} of the matched contour distance transform gradient in the superpixel, in both the horizontal and vertical direction	6
$\mathbf{y}_i^{E_1}$	Mean, p_{10} and p_{90} of the LoG response within the superpixel, taken over 4 scales	12
$\mathbf{y}_i^{E_2}$	Mean, p_{10} and p_{90} of the structure tensor eigenvalues of the superpixel, taken over 4 scales	24

for their generality and as a consequence are directly applicable to different imaging modalities such as MRI [9].

The first set of features $\mathbf{y}_i^T$ characterise the intensity and textural properties of the superpixels. They take the form of normalised intensity histograms over the pixels within each superpixel and SIFT [14, 16] descriptors of a fixed size calculated at the superpixel centroids.

The location features are based on a local coordinate system for each vertebra. This helps segmentation by providing features that describe the superpixel's relative location. The local coordinates are obtained by matching a contour to the top of the vertebral body. We do this by first (automatically) cropping the ground truth segmentation contours above their centroids, so that the resulting contour set $\mathcal{C}$ corresponds to the upper, roughly semi-circular, boundary of each vertebral body in the ground truth set. Each ground truth image is therefore associated with a single contour $C \in \mathcal{C}$ and our goal is to find the best matching contour of the set for a new image. We use a Laplacian of Gaussian (LoG) filter to detect the outer boundary of the vertebra and search over the image to find the point where the average LoG response along the contour is greatest. The best match is the contour with the maximum response of the set. Features are derived from the matched contour region by centring the pixel coordinates at the region's centroid and computing the gradient of the distance transform [9]. While the matching process depends on the presence of an adequate number of ground truth contours, in practice only an approximate match to the vertebra is required to derive the location features. Using a set of generated synthetic contours is a possibility in cases where the ground truth data is very limited.

Finally, the features in $\mathbf{y}_i^E$ are distinctive of superpixels at the edges and corners of the vertebrae and help to separate the vertebra and background classes around the boundary. We take the LoG response within the superpixel over 4 different scales to form the first feature vector. The second feature vector is formed from the eigenvalues of the structure tensor [12] within the superpixel, taken over 4 scales.

4 Potential Functions

We next describe the potential functions used in (1). Both the data and smoothness terms of the CRF are based on the characteristics learned from superpixel training examples.

We first convert the pixel-level ground truth labels into superpixel-level labels by assigning each superpixel to the class with the majority vote; as Fig. 1 illustrates, there is little ambiguity in this assignment. We then use the superpixel feature/label examples to train a support vector machine (SVM) [5] using an RBF kernel, given by

$$K(\mathbf{y}_i, \mathbf{y}_j) = \exp\left(-\gamma ||\mathbf{y}_i - \mathbf{y}_j||_2^2\right) \tag{3}$$

where γ is a kernel width parameter found using cross-validation on the training data. Probability estimates for the vertebra and background classes are obtained from the SVM using the method of [18] and incorporated into the data term of the CRF. To do this we define the data term as the negative log likelihood of an observation (feature vector) given the class label (i.e. vertebra or background):

$$\psi(\mathbf{y}_i \mid x_i) = -\log\left(P(\mathbf{y}_i \mid x_i)\right) \tag{4}$$

where the likelihood term $P(\mathbf{y}_i \mid x_i)$ for each superpixel is given by the SVM posterior probability. The superpixel likelihoods given by the data term are highly discriminative and localised to the vertebrae regions, as can be seen in the examples shown in Fig. 2b. Note that all pixels within a given superpixel are assigned the same probability, so the figure shows the superpixel-wise probability estimates.

For the second-order potential of our CRF model, we use *distance metric learning* to learn an appropriate distance metric between the superpixel features. While second-order penalties based on standard Euclidean distance measures are often used in graph cut formulations, metric learning tailors the distance measure to the data itself, rather than being chosen *ad hoc*. In particular, we use the Large Margin Nearest Neighbour (LMNN) [17] algorithm to learn a pseudometric of the form

$$D_{\mathbf{M}}(\mathbf{y}_i, \mathbf{y}_j) = (\mathbf{y}_i - \mathbf{y}_j)^\top \mathbf{M} (\mathbf{y}_i - \mathbf{y}_j). \tag{5}$$

The metric is estimated by learning a linear transformation of the data $\mathbf{L}$ such that $\mathbf{L}^\top \mathbf{L} = \mathbf{M}$. The goal is that the k-nearest neighbours of examples in the transformed space (determined by $\mathbf{L}$) should belong to the same class while those belonging to different classes should be separated by a large margin.

We incorporate the learned metric into the second-order potential function as follows

$$\phi(\mathbf{y}_i, \mathbf{y}_j \mid x_i, x_j) = \begin{cases} \exp\left(-D_{\mathbf{M}}(\mathbf{y}_i, \mathbf{y}_j)\right) & \text{if } \; x_i \neq x_j \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

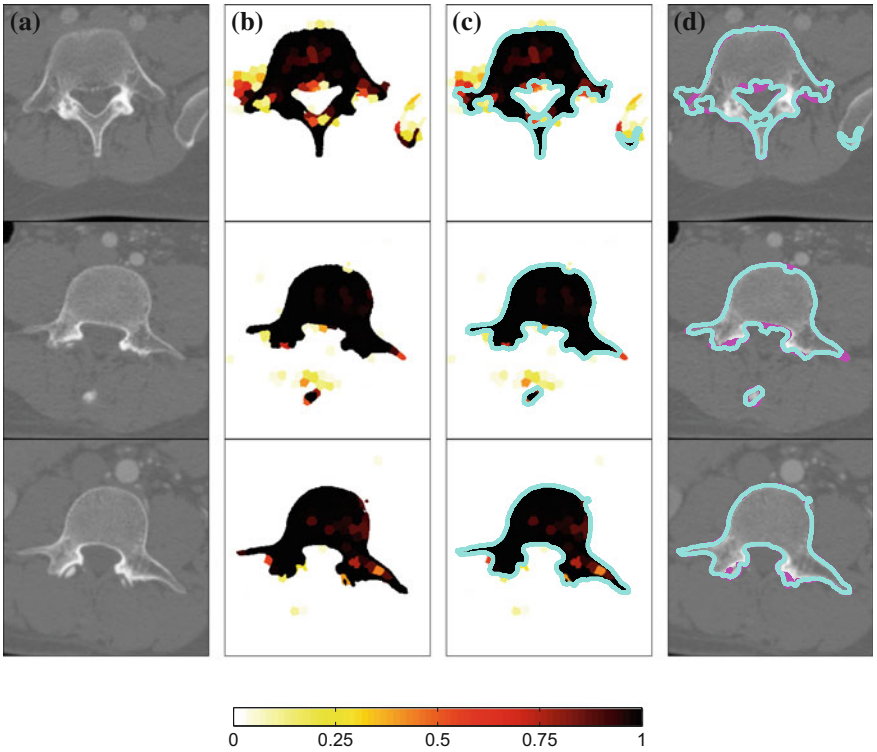

Fig. 2 **a** Shown *top* to *bottom* are CT images corresponding to the minimum, median and maximum Dice similarity score (0.88, 0.97 and 0.98), respectively. **b** SVM probability estimates for the images in the *left* hand column. *Darker* regions indicate higher probability of belonging to the vertebra class. **c** Final segmentation contours from the CRF shown overlaid with the probability estimates (*cyan*). **d** Segmentation contours shown for both the ground truth annotations (*magenta*) and CRF model (*cyan*) (color figure online)

which penalises neighbouring superpixels which have similar feature vectors and are assigned to different classes. The final segmentations using the CRF are compared with the probability estimates from the data term in Fig. 2c.

5 Experiments

We next assess the performance of the method. We first describe the data used for the experiments and the training procedure for the CRF. The segmentation results are then discussed.

5.1 Experimental Setup

The CT data consists of 2D axial slices of lumbar vertebrae from 10 different subjects, each of which has been manually annotated.[1] The images were acquired with Philips or Siemens multi detector CT scanners using an in plane resolution of between 0.31 and 0.45 mm with a slice thickness of 1 mm [19]. We used a total of 50 ground truth images by selecting the middle vertebral slice from each of the 5 lumbar vertebrae of each manually annotated subject. The 512×512 pixel images were cropped to 391×371 using a bounding box around the vertebrae regions.

The experiments were carried out on a 4-core Intel i5 2.50 GHz machine with 8 GB of RAM. The implementation is written in MATLAB with outside C++ code for certain tasks including superpixel extraction, SVM optimisation and CRF minimisation using graph cuts.

5.2 Model Training

To train the SVMs, leave-one-out (LOO) cross-validation was performed by leaving out one subject (i.e. 5 images) on each iteration and training on the remaining 45 images. The model was then tested on the 5 images from the held out subject and the process was repeated for all 10 subjects. Thus the training and test images were always from separate subjects. The SVM cost parameter $C = 4$ and the kernel width parameter $\gamma = 0.25$ were determined by cross-validation and used for all training runs.

Note that the training data is unbalanced, as there are many more negative (background) examples than positive (foreground) examples. We addressed this by training on a fixed proportion of randomly sampled positive and negative examples. The same LOO approach was used for the LMNN algorithm, with the distance metric learned on the training images for each LOO iteration and applied on the 5 held out images.

5.3 Segmentation Results

To evaluate the degree of overlap between the automatic segmentation and the ground truth, the Dice similarity coefficient (DSC) was used. Given two segmentations $\mathbf{x}$ and $\mathbf{x}'$, the DSC score is defined as

$$\mathrm{DSC}(\mathbf{x}, \mathbf{x}') = \frac{2|\mathbf{x} \cap \mathbf{x}'|}{|\mathbf{x}| + |\mathbf{x}'|}. \tag{7}$$

[1] Data from the CSI2014 segmentation competition is available from the SpineWeb initiative: http://spineweb.digitalimaginggroup.ca.

The score is in the range [0, 1] with 0 indicating no overlap and 1 indicating maximum overlap. LOO testing was used to evaluate the segmentation performance of the method, with the scores taken over all LOO runs.

The segmentations were also evaluated using three distance measures. The mean symmetric absolute surface distance (MSD) score is determined by finding for each set of boundary pixels of both the segmentation and corresponding ground truth, the closest boundary pixels of the other set. The mean of the Euclidean distances to the closest points gives the score for the image, with 0 indicating a perfect segmentation. The RMS symmetric surface distance takes the squared distances between the two sets of boundary pixels, with the final score defined as the root of the average squared distances. Finally, the maximum symmetric absolute surface distance is similar to the MSD score but takes the maximum of the distances instead of the mean. Further discussion of these metrics is provided in [6].

The average processing time for segmentation of a single image was approximately 50 s. The average DSC score was 0.97 with standard deviation 0.01 and the average MSD score was 1.83 with standard deviation 2.54. Table 2 summarises the results obtained on each lumbar vertebra using the evaluation metrics. Figure 2d shows example segmentation contours for both the ground truth and CRF model, corresponding to the minimum, median and maximum DSC score (0.88, 0.97 and 0.98). As the figure suggests, in most cases the automatic segmentation is very close to the manually determined region.

The results obtained by our method compare favourably with those recently presented in [8], who reported an average DSC score of 0.94 ± 0.02. In the same work, the authors showed that their method obtained superior results compared with two other recent approaches to vertebra segmentation [10, 13].

Table 2 Minimum, median and maximum values of the evaluation metrics for each lumbar vertebra

Metric		L1	L2	L3	L4	L5
Dice score	Min	0.92	0.96	0.95	0.94	0.88
	Median	0.97	0.97	0.97	0.97	0.97
	Max	0.98	0.98	0.98	0.98	0.98
Mean surf. dist.	Min	0.91	0.61	0.88	0.85	0.85
	Median	1.20	1.09	1.34	1.37	1.40
	Max	5.29	1.96	1.77	1.99	7.99
RMS surf. dist.	Min	1.38	0.89	1.65	1.30	1.15
	Median	2.00	2.11	2.13	2.52	2.79
	Max	14.87	8.29	4.30	5.47	22.25
Max surf. dist.	Min	5.00	3.17	9.00	6.71	6.00
	Median	14.02	12.39	15.51	14.53	15.62
	Max	91.76	71.87	32.56	42.30	101.55

5.4 3D Segmentation of Vertebrae

The method we have described can also be used to obtain 3D segmentations of vertebrae from individually segmented slices by modifying the way the location features are derived. To do this, the contour matching is first carried out on each slice of the image stack. We then use the M-estimator sample consensus (MSAC) [15] algorithm to remove poor contour matches by detecting and eliminating outliers. Outliers are determined based on the distance to their k-nearest neighbours in the set of matched contours and removed by fitting a polynomial curve through the set of inliers. Location features analogous to the 2D case can then be derived from the correctly matched contours by computing the distance transform in 3D. Figure 3 shows an example 3D vertebra segmentation constructed from segmentations of the constituent slices.

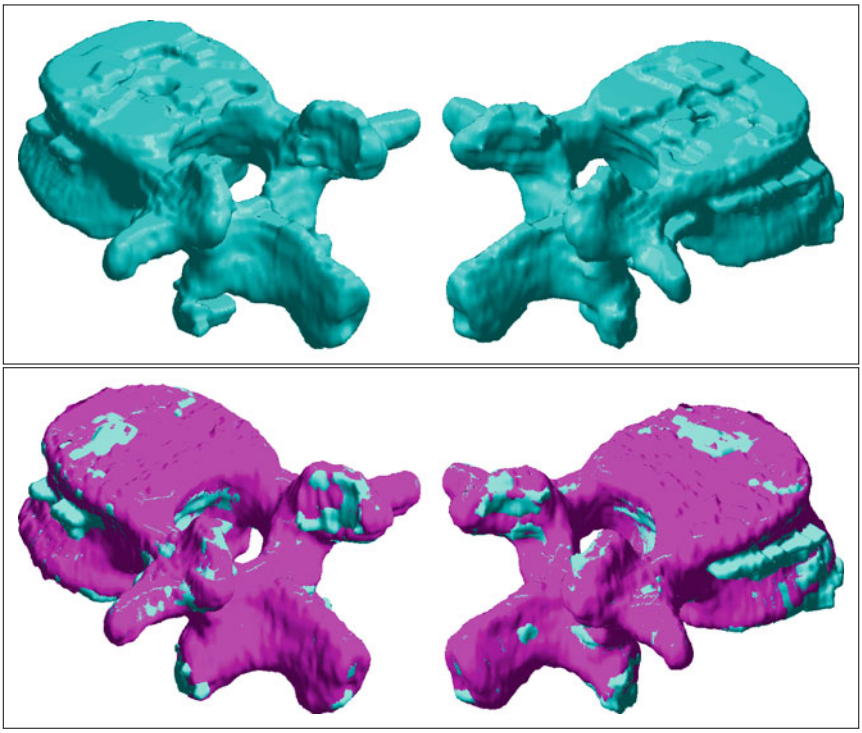

Fig. 3 The *top* figure shows a 3D segmentation of a lumbar vertebra (L2) constructed from segmentations of the constituent slices. The *bottom* figure shows the overlap between the CRF segmentation (*cyan*) and ground truth (*magenta*) (color figure online)

6 Conclusion

We presented an automatic approach for segmentation of vertebra slices from CT images. Our method avoids the requirement of explicit prior shape information and can therefore deal with a wide range of anatomical variation. The results demonstrate that consistently accurate segmentations can be obtained on each of the different lumbar vertebrae from a variety of subjects. Key to the effectiveness of this method is the learning of superpixel features from ground truth data for incorporation into the conditional random field, which in turn ensures spatial coherence. We note that much poorer performance is obtained with traditional features such as just intensity histograms. Finally, we note that this method may be extended to 3D segmentation in a straightforward way. Future work will aim to improve the results in 3D by operating on supervoxels rather than superpixels and by generalising the set of features to characterise the supervoxel regions.

Acknowledgments H. Hutt was funded by the EPSRC. We are grateful to the SpineWeb initiative for making the data available and to the organisers of the CSI2014 competition.

References

1. Achanta, R., Shaji, A., Smith, K., Lucchi, A., Fua, P., Susstrunk, S.: SLIC superpixels compared to State-of-the-Art superpixel methods. IEEE Trans. Pattern Anal. Mach. Intell. **34**(11), 2274–2282 (2012)
2. Blake, A., Kohli, P., Rother, C. (eds.): Markov Random Fields for Vision and Image Processing. The MIT Press, Cambridge (2011)
3. Boykov, Y., Funka-Lea, G.: Graph cuts and efficient N-D image segmentation. Int. J. Comput. Vis. **70**(2), 109–131 (2006)
4. Boykov, Y., Kolmogorov, V.: An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. IEEE Trans. Pattern Anal. Mach. Intell. **26**(9), 1124–1137 (2004)
5. Chang, C.-C., Lin, C.-J.: LIBSVM: A library for support vector machines. ACM Trans. Intell. Syst. Tech. 2:27:1–27:27 (2011). Software available at http://www.csie.ntu.edu.tw/~cjlin/libsvm
6. Gerig, G., Jomier, M., Chakos, M.: Valmet: a new validation Tool for assessing and improving 3D object segmentation. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI), vol. 2208, pp. 516–523. Springer (2001)
7. Ghosh, S., Alomari, R., Chaudhary, V., Dhillon, G.: Automatic lumbar vertebra segmentation from clinical CT for wedge compression fracture diagnosis. In: SPIE Conference Series, vol. 7963 (2011)
8. Huang, J., Jian, F., Wu, H., Li, H.: An improved level set method for vertebra CT image segmentation. BioMed. Eng. OnLine **12**(48) (2013)
9. Hutt, H. W., Everson, R. M., and Meakin, J. R.: Automatic segmentation of vertebrae from MR images. Technical Report, 2014, School of Physics, University of Exeter (2014)
10. Kim, Y., Kim, D.: A fully automatic vertebra segmentation method using 3D deformable fences. Comput. Med. Imaging Graph. **33**, 343–352 (2009)
11. Klinder, T., Ostermann, J., Ehm, M., Franz, A., Kneser, R., Lorenz, C.: Automated model-based vertebra detection, identification, and segmentation in CT images. Med. Image Anal. **13**(3), 471–482 (2009)

12. Knutsson, H.: Representing Local Structure Using Tensors. In: The 6th Scandinavian Conference on Image Analysis, Oulu, pp. 244–251 (1989)
13. Lim, P.H., Bagci, U., Bai, L.: Introducing willmore flow into level set segmentation of spinal vertebrae. IEEE Trans. Biomed. Eng. **60**(1), 115–122 (2013)
14. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. **60**(2), 91–110 (2004)
15. Torr, P.H.S., Zisserman, A.: MLESAC: a new robust estimator with application to estimating image geometry. Comput. Vis. Image Underst. **78**, 138–156 (2000)
16. Vedaldi, A., Fulkerson, B.: VLFeat: an open and portable library of computer vision algorithms (2008). http://www.vlfeat.org/
17. Weinberger, K.Q., Saul, L.K.: Distance metric learning for large margin nearest neighbor classification. J. Mach. Learn. Res. **10**, 207–244 (2009)
18. Wu, T.-F., Lin, C.-J., Weng, R.C.: Probability estimates for multi-class classification by pair wise coupling. J. Mach. Learn. Res. **5**, 975–1005 (2004)
19. Yao, J., Burns, J. E., Munoz, H., Summers, R.M.: Detection of Vertebral Body Fractures Based on Cortical Shell Unwrapping. In: Medical Image Computing and Computer-Assisted Intervention (MICCAI), vol. 7512, pp. 509–516. Springer (2012)

Interpolation-Based Detection of Lumbar Vertebrae in CT Spine Images

Bulat Ibragimov, Robert Korez, Boštjan Likar, Franjo Pernuš and Tomaž Vrtovec

Abstract Detection of an object of interest can be represented as an optimization problem that can be solved by brute force or heuristic algorithms. However, the globally optimal solution may not represent the optimal detection result, which can be especially observed in the case of vertebra detection, where neighboring vertebrae are of similar appearance and shape. An adequate optimizer has to therefore consider not only the global optimum but also local optima that represent candidate locations for each vertebra. In this paper, we describe a novel framework for automated spine and vertebra detection in three-dimensional (3D) images of the lumbar spine, where we apply a novel optimization technique based on interpolation theory to detect the location of the whole spine in the 3D image and to detect the location of individual vertebrae within the spinal column. The performance of the proposed framework was evaluated on 10 computed tomography (CT) images of the lumbar spine. The resulting mean symmetric absolute surface distance of 1.25 ± 0.41 mm and Dice coefficient of $83.67 \pm 4.44\%$, computed from the final vertebra detection results against corresponding reference vertebra segmentations, indicate that the proposed framework can successfully detected vertebrae in CT images of the lumbar spine.

B. Ibragimov (✉) · R. Korez · B. Likar · F. Pernuš · T. Vrtovec
Faculty of Electrical Engineering, University of Ljubljana, Tržaska 25,
1000 Ljubljana, Slovenia
e-mail: bulat.ibragimov@fe.uni-lj.si

R. Korez
e-mail: robert.korez@fe.uni-lj.si

B. Likar
e-mail: bostjan.likar@fe.uni-lj.si

F. Pernuš
e-mail: franjo.pernus@fe.uni-lj.si

T. Vrtovec
e-mail: tomaz.vrtovec@fe.uni-lj.si

© Springer International Publishing Switzerland 2015

J. Yao et al. (eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, Lecture Notes in Computational Vision and Biomechanics 20, DOI 10.1007/978-3-319-14148-0_7

1 Introduction

Detection and segmentation of an object of interest can always be represented as a specific optimization problem. As optimization problems are usually non-deterministic polynomial-time hard (NP-hard) or non-analytically defined, they can be solved by brute force algorithms that are computationally demanding or by heuristic approaches that may not find the globally optimal solution. Moreover, if the optimization problem is ill-posed, there is no guarantee that the optimal detection or segmentation result is associated with the global optimum of the problem. Such lack of correspondence is observed in the case of vertebra detection in medical images of the spine, where neighboring vertebrae are of similar appearance and shape, which often causes misdetections. We can therefore conclude that there is a need for an optimizer that will consider not only the global optimum but also local optima of the problem that represent candidate locations for each vertebra, and that can be further used for accurate and robust spine and vertebra detection. As standard optimization techniques, such as the Powell's optimizer [1] or covariance matrix adaptation evolution strategy (CMA-ES) [2], detect only a single local optimum, we propose to use interpolation theory to overcome the above mentioned limitations and ensure the detection of all local optima. Interpolation theory estimates the behavior of smooth functions according to a small set of points, for which the behavior is known. In the case of object detection, interpolation theory can predict the position of the object of interest by computing a specified detector response for a sparse set of points, i.e. possible transformations of the object of interest. In this paper, we describe a novel framework for automated spine and vertebra detection in three-dimensional (3D) computed tomography (CT) images of the lumbar spine, where we first apply interpolation-based optimization to detect the location of the whole lumbar spine in the 3D image, and then detect the location of individual lumbar vertebrae within the spinal column.

2 Methodology

Detection of the object of interest (e.g. spine, vertebra) in an unknown 3D image I can be, in general, performed by optimizing an objective function $f = f(\boldsymbol{\rho}, I)$, which maps the model of the object of interest, represented by a set $\boldsymbol{\rho}$ of n parameters describing its geometrical properties (e.g. position, rotation, scaling etc.), to image I in n-dimensional parameter space. The most straightforward but computationally demanding approach for optimizing f is to apply brute force to compute its response for all $\boldsymbol{\rho} \in \Omega^n$, where Ω^n is a bounded domain in the parameter space. Alternative approaches [1, 2] often compute function responses in a neighborhood of an initial set $\boldsymbol{\rho}'$; $\boldsymbol{\rho}' \in \Omega^n$, predict the optimal $\boldsymbol{\rho}^*$ and iteratively move towards $\boldsymbol{\rho}^*$. Although such optimization is computationally not demanding and may, especially if $\boldsymbol{\rho}'$ is relatively close to $\boldsymbol{\rho}^*$, lead to the globally optimal solution, it may still converge to local optima. Therefore, there is a need to observe the whole Ω^n and, at the same

time, minimize actual computations of f. To overcome these problems, we propose a novel optimization scheme that is based on interpolation theory [3]. The proposed optimization scheme does not depend on initialization and observes the whole Ω^n, and we apply it for spine and vertebra detection in 3D spine images. Depending on the objective function f, the scheme searches for optima in the form of either minima or maxima, however, we will refer to them as maxima in the following text.

2.1 Interpolation Theory

By assuming that the objective function f is smooth and its values are similar for similar sets $\boldsymbol{\rho}$ of parameters, f can be approximated by function $\bar{f}$ if the response of f is known for a limited set of nodes $\mathscr{P}$; $\mathscr{P} \subset \Omega^n$ (Fig. 1). If nodes in $\mathscr{P}$ are selected properly, the difference between f and its approximation $\bar{f}$, i.e. the interpolation error, is relatively small, and the maximum of $\bar{f}$ corresponds to the maximum of f at the same point $\boldsymbol{\rho}^*$. The main concept of interpolation theory [3] is to select the interpolation function, in our case $\bar{f}$, from a given class of functions so that its response passes through the nodes in $\mathscr{P}$, which form the interpolation grid over the observed domain Ω^n. However, interpolation of high-dimensional functions is computationally challenging, as the number of nodes grows exponentially with the increasing number of dimensions n. To reduce the computational complexity, various strategies can be applied [4, 5], but in our case we take into account the properties of spine detection. The standard spine image acquisition procedure restricts the orientation of the imaged subject, and therefore the relative incli-

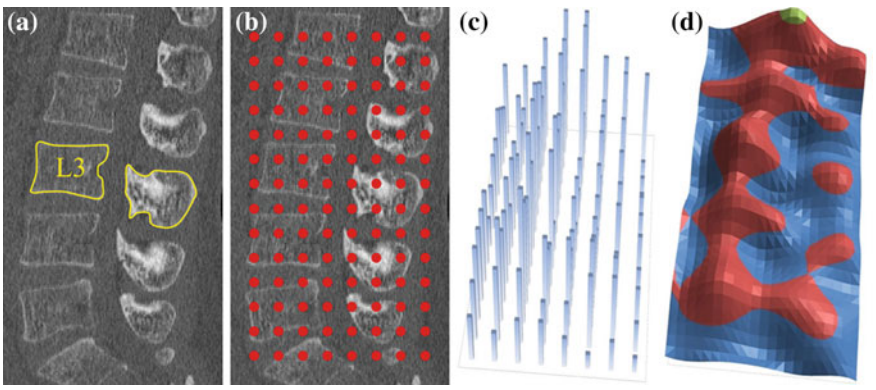

Fig. 1 An illustration of the L3 vertebra detection. **a** Reference segmentation binary mask, overlayed onto a selected sagittal cross-section of the CT image. **b** The interpolation grid. **c** The objective function f is computed only for nodes forming the interpolation grid. **d** The local maxima of the interpolation function $\bar{f}$ indicate candidate locations of the observed vertebra. In this case, the global maximum (in *green*) does not correspond to the location of L3 (Color figure online)

nation of the spinal column in the image is more constant than its position, and similarly the orientation of a vertebra is restricted by the position of neighboring vertebrae. It can be concluded that dimensions related to object translation have to be associated with a larger number of interpolation nodes than dimensions related to object rotations. To optimize the objective function f, defined on domain Ω^n of n dimensions, we therefore propose the following algorithm:

1. Initialize set σ of optimized dimensions as $\sigma = \{\}$, and set $\bar{\sigma}$ of the remaining dimensions as $\bar{\sigma} = \{1, 2, \ldots, n\}$.
2. Interpolate f with $\bar{f}$ against the first $k \ll n$ dimensions.
3. Find t local maxima of $\bar{f}$, then insert the observed dimensions into σ; $\sigma \leftarrow \sigma \cup \{1, 2, \ldots, k\}$, and remove them from $\bar{\sigma}$; $\bar{\sigma} \leftarrow \bar{\sigma} \setminus \{1, 2, \ldots, k\}$.
4. For each local maximum among t, perform the following steps:
   4a. If $\bar{\sigma} \neq \{\}$, take the next k available dimensions and interpolate f with $\bar{f}$ by considering dimensions from σ to be optimized.
   4b. Find the global maximum of $\bar{f}$, insert the observed dimensions into σ and remove them from $\bar{\sigma}$. Return to step 4a.
5. The resulting global maxima represent the locations of t local maxima of f.

In the proposed algorithm, the optimization dimensions have to be ordered according to the decreasing number of corresponding interpolation nodes, while the number of dimensions k considered in each interpolation step may vary for different steps. Nevertheless, by taking into account only k dimensions at once, the computational complexity of the high-dimensional optimization problem is reduced.

2.2 Mean Shape Model of the Lumbar Spine

Let set $\mathscr{T}$ contain 3D images of the lumbar spine, where each image is assigned a series of binary masks representing reference segmentations of each individual lumbar vertebra from level L1 to L5. To extract a shape model of each vertebra from each image in $\mathscr{T}$, the marching cubes algorithm [6] is applied to each corresponding binary mask, resulting in a 3D face-vertex mesh consisting of vertices with triangle connectivity information. The dependency of the number of vertices in each mesh on the size of the image voxel and of the observed vertebra is removed by isotropic remeshing [7], and the coherent point drift algorithm [8] is used to recover the nonrigid transformation among sets of vertices and establish point wise correspondences among vertices of the same vertebral level. Finally, the generalized Procrustes alignment [9] is used to remove translation, rotation and scaling from corresponding meshes, yielding the mean shape model of each vertebra, represented by a 3D face-vertex mesh $\mathscr{M} = \{V, F\}$ of $|V|$ vertices and $|F|$ faces (i.e. triangles). The mean shape model of the whole lumbar spine, i.e. a chain of mean shape models of individual vertebrae, is further used for spine detection, while mean shape models of individual vertebrae are used for vertebra detection in an unknown 3D image I.

2.3 Spine Detection

Let the 3D mesh $\mathscr{M} = \{V, F\}$ represent the mean shape model of the lumbar spine (Sect. 2.2), i.e. a chain of meshes representing individual vertebrae from L1 to L5. The objective function f that measures the agreement between mesh $\mathscr{M}$ and image I is defined as:

$$f(\mathscr{M}, I) = \sum_{i=1}^{|V|} \langle \mathbf{g}_H(\mathbf{v}_i, I), \mathbf{n}(\mathbf{v}_i) \rangle, \tag{1}$$

where $\langle \cdot, \cdot \rangle$ denotes the dot product, $|V|$ is the number of mesh vertices, $\mathbf{v}_i \in V$ is a mesh vertex, $\mathbf{g}_H(\mathbf{v}_i, I)$ is the Haar-like gradient [10] of image I at the location of vertex $\mathbf{v}_i$, and $\mathbf{n}(\mathbf{v}_i)$ is the mesh normal at vertex $\mathbf{v}_i$. If mesh $\mathscr{M}$ is correctly aligned with the spine in image I, then mesh normals pointing outwards of the mesh are in maximal agreement with the Haar-like gradients pointing outwards of the spine, resulting in the maximum of the objective function $f(\mathscr{M}, I)$. As Haar-like gradients are described as a difference between two neighboring cuboids instead of two neighboring voxels, they are not sensitive to local intensity fluctuations and therefore make spine detection more robust.

To detect the spine in image I, we optimize the pose of mesh $\mathscr{M}$ only against three translations (i.e. coordinates x, y and z representing sagittal, coronal and axial anatomical directions, respectively) according to the interpolation function $\bar{f}$. The dimensionality of the optimization problem is therefore $n = 3$ on the corresponding domain Ω_s^3, meaning that complexity reduction (Sect. 2.1) is in this case not required (i.e. $k = n = 3$). The global maximum of the resulting interpolation represents the location of the spine in the 3D image, and is further used to initialize the detection of individual vertebrae.

2.4 Vertebra Detection

Spine detection based on translation of the mean shape model of the lumbar spine is not adequate for robust detection of individual vertebrae due to natural anatomical differences in the size and/or curvature of the spine. If the detection error is usually low in sagittal and coronal directions, the detection error in the axial direction can be considerable, as neighboring vertebrae are of similar shape and therefore the mean shape model of the lumbar spine can be axially shifted for one or even two vertebral levels.

Let now the 3D mesh $\mathscr{M} = \{V, F\}$ represent the mean shape model of the observed vertebra of the lumbar spine (Sect. 2.2). To detect individual vertebrae after spine detection (Sect. 2.3), a domain Ω_v^3 that represents again three translations (i.e. coordinates x, y and z) is introduced, however, it is defined separately for each observed vertebra, moreover, it captures several neighboring vertebrae to take into account the eventual longitudinal shift of the mean shape model. As a result, by

aligning mesh $\mathscr{M}$ to the observed vertebra in image I, several local maxima of the objective function $f(\mathscr{M}, I)$ (Eq. 1) exist and correspond to vertebrae within Ω_v^3, moreover, the global maximum of f may not correspond to the correct location of the observed vertebra. We therefore optimize the pose of mesh $\mathscr{M}$ against three translations (i.e. coordinates x, y and z) according to the interpolation function $\bar{f}$. The dimensionality of the optimization problem is again $n = 3$ on the corresponding domain Ω_v^3, meaning that complexity reduction (Sect. 2.1) is in this case also not required (i.e. $k = n = 3$). However, in contrast to spine detection, in this case all local maxima of the resulting interpolation are extracted.

Let set $M = \{M_1, M_2, \ldots, M_R\}$ then represent local maxima of the resulting interpolation for all R vertebrae, where $M_r = \{m_{r,1}, m_{r,2}, \ldots, m_{r,t}\}$ is the set of t local maxima for the rth vertebra. To detect the location of each vertebra, we find the optimal path M^* that passes through R locations, where each location corresponds to a local maxima at a different vertebral level:

$$M^* = \arg\max_p \left(\sum_{r=1}^{R-1} g\left(m_{r,p_r}, m_{r+1,p_{r+1}}, r\right) \right), \tag{2}$$

where set $p = \{p_1, p_2, \ldots, p_R\}$ represents a combination of indices of local maxima belonging to different vertebral levels (i.e. exactly one from each set M_r), and function $g\left(m_{r,p_r}, m_{r+1,p_{r+1}}, r\right)$ measures the agreement of relative locations m_{r,p_r} and $m_{r+1,p_{r+1}}$ belonging to neighboring rth and $(r+1)$th vertebrae, respectively:

$$g\left(m_{r,p_r}, m_{r+1,p_{r+1}}, r\right) = \prod_{k=1}^{3} G\left(\Delta^k(m_{r,p_r}, m_{r+1,p_{r+1}}), r\right), \tag{3}$$

where $\Delta^k(m_{r,p_r}, m_{r+1,p_{r+1}})$ is the difference in kth dimension (i.e. one of the coordinates x, y or z) between locations m_{r,p_r} and $m_{r+1,p_{r+1}}$, and $G\left(\Delta^k(\cdot,\cdot), k\right)$ is the Gaussian kernel estimation of the agreement between $\Delta^k(\cdot,\cdot)$ and the difference in kth dimension for neighboring rth and $(r+1)$th vertebrae. The obtained optimal path M^* is a set that contains the optimal location of each observed vertebra.

2.5 Vertebra Alignment

After detecting all vertebrae, a more accurate alignment of the mean shape model of each observed vertebra is performed. As the length of the lumbar spine and relative orientation of each vertebra can be obtained from optimal path M^* (Eq. 2), each mean shape model is first scaled according to the ratio between the obtained length and the length of the mean shape model of the lumbar spine (Sect. 2.3), and oriented according to the location of its neighboring mean shape models in the optimal path. Next, a domain Ω_v^7 of three translations (i.e. coordinates x, y and z), one scaling

(i.e. factor s) and three rotations (i.e. angles φ_x, φ_y and φ_z about coordinate axes x, y and z, respectively) is initialized and used to interpolate the objective function $f(\mathscr{M}, I)$ (Eq. 1). However, as an exhaustive search through all $n = 7$ dimensions is computationally expensive, the complexity is reduced dimension-wise by first optimizing $k = 4$ dimensions consisting of translations and scaling, and then the remaining $k = 3$ dimensions consisting of rotations (Sect. 2.1). The resulting alignment represents the final vertebra detection result.

3 Experiments and Results

The performance of the described spine and vertebra detection framework was tested on a database of 10 CT images of the lumbar spine [11][1] (i.e. a total of 50 lumbar vertebrae, with axial in-plane pixel size of 0.3–0.8 mm and cross-sectional thickness of 0.7–1.5 mm) by applying a leave-one-out evaluation scheme. A reference segmentation binary mask was available for each vertebra in the database. The framework was implemented in C#, and executed on a personal computer with Intel Core i7 processor at 2.8 GHz and 8 GB of memory without graphics processing unit.

3.1 Experimental Details

The mean shape model of the lumbar spine (Sect. 2.2) was obtained by applying the marching cubes algorithm [6] to binary masks representing reference segmentations of each vertebra, resulting in 3D face-vertex meshes $\mathscr{M} = \{V, F\}$ of genus 1 (i.e. the number of holes is 1, as expected for lumbar vertebrae). The corresponding number of vertices $|V|$ =31.000–161.000 (the number of faces was $|F| = 2|V|$) was further reduced to $|V|$ =3.228–5.642 by isotropic remeshing with mean edge length of 2.25 mm [7]. After establishing correspondences among meshes of the same vertebral level by the coherent point drift algorithm [8] and applying the generalized Procrustes alignment [9], the mean shape model of each lumbar vertebra was obtained. To detect the spine and vertebrae in each 3D image, the interpolation-based approach was applied by using splines for the interpolation function $\bar{f}$ on an equidistant interpolation grid, which usually results in a low interpolation error [3]. For spine detection (Sect. 2.3), the interpolation domain Ω_s^3 was defined for each translation x, y and z on the [−30 mm, +30 mm] interval (6 mm step), with the resulting number of nodes on the interpolation grid equal to 1331. For vertebra detection (Sect. 2.4), the interpolation domain Ω_v^3 was defined for translations x and y on the [−25 mm, +25 mm] interval (5 mm step), and for the translation z on the [−70 mm, +70 mm] interval (5 mm step), with the resulting number of nodes on the interpolation grid equal to

[1] Publicly available through http://lit.fe.uni-lj.si.

3, 509. For the subsequent more accurate alignment of the mean shape model of each lumbar vertebra (Sect. 2.5), the interpolation domain Ω_v^7 was defined for each translation x, y and z on the $[-7\,\text{mm}, +7\,\text{mm}]$ interval (2 mm step), for the scaling factor s on the [0.9, 1.1] interval (0.05 step), and for each rotation φ_x, φ_y and φ_z on the $[-9°, +9°]$ interval (0.86° step), with the resulting number of nodes on the interpolation grid equal to 1,985. In the case of spine detection, the initial location of the mean shape model of the lumbar spine, i.e. the origin of interpolation domain Ω_s^3, was set to the average location of the spine, as observed in the set of training images used to generate the mean shape model of the lumbar spine. In the case of vertebra detection and vertebra alignment, the initial location of the mean shape model of each lumbar vertebra, i.e. the origin of interpolation domains Ω_v^3 and Ω_v^7, was set according to the location resulting from the previous optimization step, i.e. according to the results of spine detection and vertebra detection, respectively.

3.2 Results

The performance of the proposed framework was evaluated by the mean symmetric absolute surface distance (MSD), symmetric root-mean-square surface distance (RMSSD), maximal symmetric absolute surface distance (MaxSD) and Dice coefficient (DICE), computed between the resulting 3D meshes and corresponding reference segmentation binary masks. Detailed results for the detection of individual vertebral levels are presented in Table 1. The overall lumbar vertebra detection performance (mean ± standard deviation) was MSD = 1.25 ± 0.41 mm, RMSSD = 1.73±0.55 mm, MaxSD = 8.64±2.46 mm and DICE = 83.67±4.44%. On average, spine detection took around 1 min, vertebra detection around 2 min, and vertebra alignment around 40 s. Figure 2 shows the initialization of the mean shape model of the lumbar spine and the resulting spine detection results, while Fig. 3 shows the results of vertebra detection and vertebra alignment for selected examples of CT lumbar spine images from the evaluation database.

Table 1 Final lumbar vertebra detection results in terms of mean symmetric absolute surface distance (MSD), symmetric root-mean-square surface distance (RMSSD), maximal symmetric absolute surface distance (MaxSD) and Dice coefficient (DICE), reported as mean ± standard deviation

Vertebral level	MSD (mm)	RMSSD (mm)	MaxSD (mm)	DICE (%)
L1	1.02 ± 0.23	1.44 ± 0.33	8.00 ± 2.79	85.95 ± 2.63
L2	1.04 ± 0.18	1.46 ± 0.25	6.93 ± 1.32	85.62 ± 2.06
L3	1.17 ± 0.23	1.61 ± 0.29	8.35 ± 1.18	84.94 ± 2.11
L4	1.45 ± 0.52	2.05 ± 0.73	9.50 ± 2.50	81.56 ± 4.85
L5	1.54 ± 0.50	2.10 ± 0.63	10.42 ± 2.75	80.26 ± 2.75

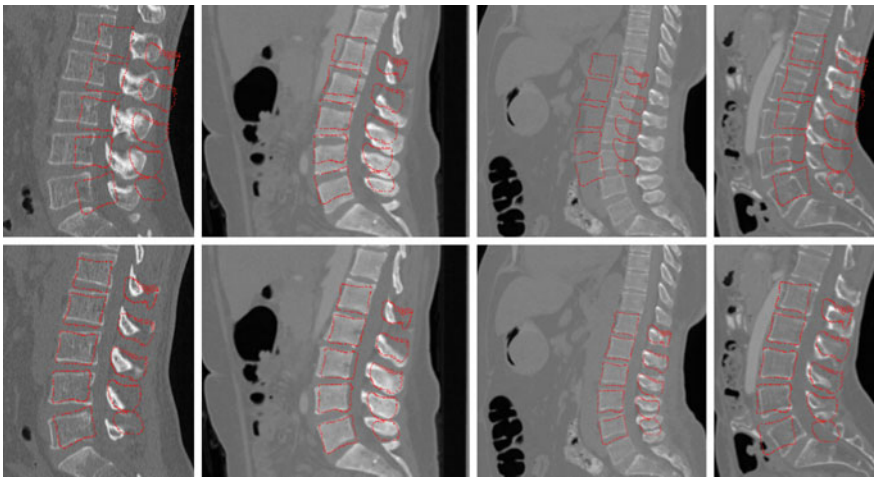

Fig. 2 Selected examples of CT lumbar spine images, shown in mid-sagittal cross-sections, with overlayed initialization of the mean shape model (*top row*) and spine detection results (*bottom row*)

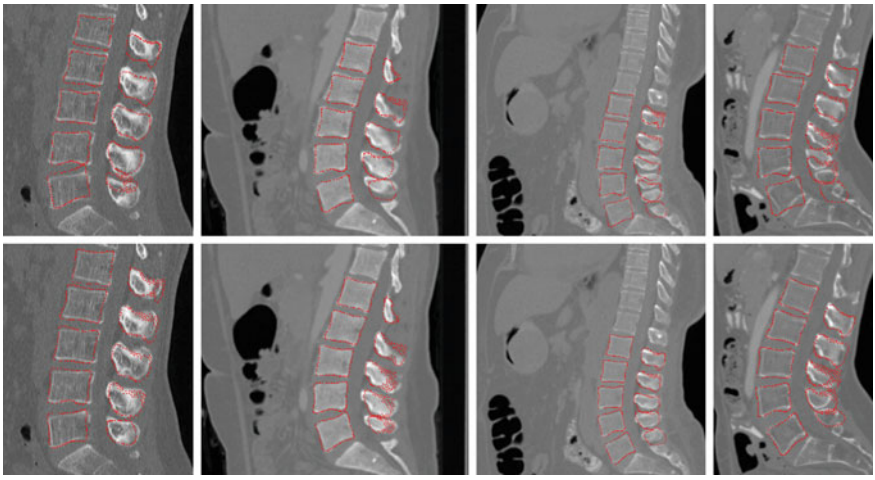

Fig. 3 Selected examples of CT lumbar spine images, shown in mid-sagittal cross-sections, with overlayed vertebra detection results (*top row*) and vertebra alignment result (*bottom row*)

4 Discussion and Conclusion

Computerized spine analysis is above all based on accurate and robust detection of each vertebra, however, this task is hampered by the fact that neighboring vertebrae are similar in shape. As a result, vertebrae cannot be successfully distinguished

without observing the whole spine, or at least a larger spine section, and without observing neighboring structures, such as the sacrum, ribs or specific internal organs. Semi-automatic vertebra detection methods require manual identification of such an anchor structure [12], or manual annotation in image of one modality followed by its propagation to images of different modalities of the same subject [13]. On the other hand, automated vertebra detection approaches split the problem into detection of candidate objects of interest, i.e. vertebrae or intervertebral discs, which is often based on inter-object spatial relationships. Feature-based vertebra detection methods treat vertebra centers or all image voxels belonging to the observed vertebra as points of interest, i.e. landmarks, and model them by single intensity- or Haar wavelet-based features that are combined into landmark detectors by Adaboost or random forests approaches [14–16]. Alternatively, vertebra detection can rely on vertebra shape, which is often described by active shape and appearance models [17–19], generalized Hough transforms [20, 21] or features that are based on vertebral bilateral symmetry and cylindricality of the vertebral body [22]. Although most of the approaches assume that the whole spine or a predefined part of the spine is present in the image, several frameworks also deal with more challenging images of arbitrary field of view [15, 16, 20].

In this paper, we propose a novel lumbar spine and vertebra detection framework. As vertebra detection suffers from convergence to wrong maxima, i.e. the target vertebra can be mis-detected at the location of its neighboring vertebrae, all local maxima of the objective function have to be detected. We therefore cannot use standard optimizers such as the Powell's optimizer [1] or CMA-ES [2], which are usually limited by their initialization and detect a single maximum. We proposed to use a novel optimizer that is based on interpolation theory [3] and takes into account the complete domain, where the object of interest can be located and where the objective function is defined. To avoid brute force analysis over the domain, we computed the objective function on a sparse set of points on that domain and then approximated the objective function on the complete domain. The obtained spine and vertebra detection results can be used to initialize methods for accurate segmentation of individual vertebrae [23]. However, interpolation theory is not limited to vertebra detection or even object detection, but can be used to speed up optimization and smooth the potentially harmful fluctuations of the objective function in various image analysis problems. In our future work, we therefore plan to extend this methodology, validate it on larger databases of spine images with different field of view, and combine it with vertebra segmentation.

Acknowledgments This work was supported by the Slovenian Research Agency (ARRS) under grants P2–0232 and L2–4072.

References

1. Powell, M.J.D.: An efficient method for finding the minimum of a function of several variables without calculating derivatives. Comput. J. **7**(2), 155–162 (1964)
2. Hansen, N.: The CMA evolution strategy: a comparing review. In: Lozano, J.A. et al. (eds.) Towards a New Evolutionary Computation, pp. 75–102. Springer, (2006)
3. Cheney, E.W., Light, W.A.: A course in approximation theory. Am. Math. Soc. (2009)
4. Bungartz, H.-J.: Finite elements of higher order on sparse grids. Shaker Verlag (1998)
5. Gerstner, T., Griebel, M.: Dimension adaptive tensor product quadrature. Computing **71**(1), 65–87 (2003)
6. Lorensen, W.E., Cline, H.E.: Marching cubes: a high resolution 3D surface construction algorithm. In: Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques—SIGGRAPH'87, 163–169 (1987)
7. Botsch, M., Kobbelt, L.: A remeshing approach to multiresolution modeling. In: Proceedings of the 2004 Eurographics/ACM SIGGRAPH Symposium on Geometry Processing—SGP'04, 185–192 (2004)
8. Myronenko, A., Song, X.: Point set registration: coherent point drift. IEEE Trans. Pattern Anal. Mach. Intell. **32**(12), 2262–2275 (2010)
9. Gower, J.C.: Generalized procrustes analysis. Psychometrika **40**(1), 33–51 (1975)
10. Viola, P., Jones, M.: Robust real-time face detection. Int. J. Comput. Vis. **57**(2), 137–154 (2004)
11. Ibragimov, B., Likar, B., Pernuš, F., Vrtovec, T.: Shape representation for efficient landmark-based segmentation in 3D. IEEE Trans. Med. Imaging **33**(4), 861–874 (2014)
12. Weiss, K.L., Storrs, J.M., Banto, R.B.: Automated spine survey iterative scan technique. Radiology **239**(1), 255–262 (2006)
13. Otake, Y., Schafer, S., Stayman, J.W., Zbijewski, W., Kleinszig, G., Graumann, R., Khanna, A.J., Siewerdsen, J.H.: Automatic localization of vertebral levels in X-ray fluoroscopy using 3D–2D registration: a tool to reduce wrong-site surgery. Phys. Med. Biol. **57**(17), 5485–5508 (2012)
14. Huang, S.H., Chu, Y.H., Lai, S.H., Novak, C.L.: Learning-based vertebra detection and iterative normalized-cut segmentation for spinal MRI. IEEE Trans. Med. Imaging **28**(10), 1595–1605 (2009)
15. Glocker, B., Feulner, J., Criminisi, A., Haynor, D.R., Konukoglu, E.: Automatic localization and identification of vertebrae in arbitrary field-of-view CT scans. In: Proceedings of the 15th International Conference on Medical Image Computing and Computer-Assisted Intervention—MICCAI 2012, 590–598 (2012)
16. Glocker, B., Zikic, D., Konukoglu, E., Haynor, D.R., Criminisi, A.: Vertebrae localization in pathological spine CT via dense classification from sparse annotations. In: Proceedings of the 16th International Conference on Medical Image Computing and Computer-Assisted Intervention—MICCAI 2013, 262–270 (2013)
17. Long, L.R., Thoma, G.R.: Use of shape models to search digitized spine X-rays. In: Proceedings of the 13th IEEE Symposium on Computer-Based Medical Systems—CBMS 2000, 255–260 (2000)
18. Howe, B., Gururajan, A., Sari-Sarraf, H., Long, L.: Hierarchical segmentation of cervical and lumbar vertebrae using a customized generalized Hough transform and extensions to active appearance models. In: Proceedings of the 6th IEEE Southwest Symposium on Image Analysis and Interpretation—SSIAI 2004, 182–186 (2004)
19. Koompairojn, S., Hua, K., Bhadrakom, C.: Automatic classification system for lumbar spine X-ray images. In: Proceedings of the 19th IEEE International Symposium on Computer-Based Medical Systems—CBMS 2006, 213–218 (2006)
20. Klinder, T., Ostermann, J., Ehm, M., Franz, A., Kneser, R., Lorenz, C.: Automated model-based vertebra detection, identification, and segmentation in CT images. Med. Image Anal. **13**(3), 471–482 (2009)

21. Larhmam, M.A., Benjelloun, M., Mahmoudi, S.: Vertebra identification using template matching modeling and K-means clustering. Int. J. Comput. Assist. Radiol. Surg. **9**(2), 177–187 (2014)
22. Štern, D., Likar, B., Pernuš, F., Vrtovec, T.: 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 (2010)
23. Korez, R., Ibragimov, B., Likar, B., Pernuš, F., Vrtovec, T.: An improved shape-constrained deformable model for segmentation of vertebrae from CT lumbar spine images. In: Proceedings of the 2nd MICCAI Workshop on Computational Methods and Clinical Applications for Spine Imaging—MICCAI CSI 2014, 76–85 (2014)

An Improved Shape-Constrained Deformable Model for Segmentation of Vertebrae from CT Lumbar Spine Images

Robert Korez, Bulat Ibragimov, Boštjan Likar, Franjo Pernuš and Tomaž Vrtovec

Abstract Accurate and robust segmentation of spinal and vertebral structures from medical images is a challenging task due to a relatively high degree of anatomical complexity and articulation of spinal structures, as well as due to image spatial resolution, inhomogeneity and low signal-to-noise ratio. In this paper, we describe an improved framework for vertebra segmentation that is based on an existing shape-constrained deformable model, which was modified with the aim to improve segmentation accuracy, and combined with a robust initialization that results from vertebra detection by interpolation-based optimization. The performance of the proposed segmentation framework was evaluated on 10 computed tomography (CT) images of the lumbar spine. The overall segmentation performance of 0.43 ± 0.14 mm in terms of mean symmetric absolute surface distance and 93.76 ± 1.61 % in terms of Dice coefficient, computed against corresponding reference vertebra segmentations, indicates that the proposed framework can accurately segment vertebrae from CT images of the lumbar spine.

R. Korez (✉) · B. Ibragimov · B. Likar · F. Pernuš · T. Vrtovec
Faculty of Electrical Engineering, University of Ljubljana, Tržaska 25, 1000 Ljubljana, Slovenia
e-mail: robert.korez@fe.uni-lj.si

B. Ibragimov
e-mail: bulat.ibragimov@fe.uni-lj.si

B. Likar
e-mail: bostjan.likar@fe.uni-lj.si

F. Pernuš
e-mail: franjo.pernus@fe.uni-lj.si

T. Vrtovec
e-mail: tomaz.vrtovec@fe.uni-lj.si

© Springer International Publishing Switzerland 2015

J. Yao et al. (eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, Lecture Notes in Computational Vision and Biomechanics 20, DOI 10.1007/978-3-319-14148-0_8

1 Introduction

Accurate and robust segmentation of spinal and vertebral structures from medical images is a challenging task due to a relatively high degree of anatomical complexity (i.e. vertebrae consisting of the vertebral body, pedicles, laminae and spinous process) and due to the articulation of vertebrae with each other. In addition to the complexity and articulation, the problem also lies in insufficient image spatial resolution, inhomogeneity and low signal-to-noise ratio. Since skeletal structures have high contrast when observed in computed tomography (CT) images, CT is commonly the modality of choice for assessing three-dimensional (3D) skeletal structures, such as the spine and vertebrae.

In recent years, several automated and semi-automated methods focusing on the vertebra segmentation problem have been developed for CT images. Kim and Kim [8] proposed a fully automated method that constructs 3D fences to separate vertebrae from valley-emphasized Gaussian images, and then the region growing algorithm is applied within 3D fences to obtain the final segmentation. Klinder et al. [9] progressively adapted tube-shaped segments to extract the spine curve, performed vertebra detection on curved-planar reformatted images using the generalized Hough transform, identified vertebrae by rigid registration of appearance models to the detected candidates, and obtained the final segmentation by adapting shape-constrained models of the individual vertebrae. Kadoury et al. [6, 7] built an articulated shape manifold from a training database by embedding the data into a low-dimensional sub-space, and applied the Markov random field optimization to infer between the unseen target shape and shape manifold. Lim et al. [10] incorporated local geometrical features using the Willmore flow and prior shape knowledge by kernel density estimation into a level set segmentation framework. Ma and Lu [12] introduced a learning-based bone structure edge detection algorithm and hierarchical coarse-to-fine deformable surface-based segmentation that relied on response maps of a trained edge detector. Rasoulian et al. [14] developed a statistical multi-vertebrae model of shape and pose, and proposed a novel iterative expectation maximization registration technique to align the model to CT images of the spine. Ibragimov et al. [5] presented a segmentation framework, in which a novel landmark-based shape representation of vertebrae was combined with a landmark detection framework based on game theory.

In this paper, we describe an improved framework for vertebra segmentation that is based on the shape-constrained deformable model [9, 15]. Our framework is initialized by the results of a novel vertebra detection and alignment algorithm [4], and the segmentation of each vertebra is then obtained by a mesh deformation technique that moves mesh vertices to their optimal locations while preserving the underlying vertebral shape. The performance of the proposed segmentation framework was evaluated on vertebrae from CT images of the lumbar spine, and the obtained results with the mean error below 0.5 mm indicate that accurate segmentation of vertebrae was achieved.

2 Methodology

2.1 Mean Shape Model of the Lumbar Spine

Let set $\mathscr{T}$ contain 3D images of the lumbar spine, where each image is assigned a series of binary masks representing reference segmentations of each individual lumbar vertebra from level L1 to L5. To extract a shape model of each vertebra from each image in $\mathscr{T}$, the marching cubes algorithm [11] is applied to each corresponding binary mask, resulting in a 3D face-vertex mesh consisting of vertices with triangle connectivity information. The dependency of the number of vertices in each mesh on the size of the image voxel and of the observed vertebra is removed by isotropic remeshing [1]. In order to establish pointwise correspondences among vertices of the same vertebral level, the nonrigid transformation among sets of vertices is recovered using state-of-the-art coherent point drift algorithm [13] that outperforms other methods for point set registration. Finally, the generalized Procrustes alignment [3] is used to remove translation, rotation and scaling from corresponding meshes, yielding the mean shape model of each vertebra, represented by a 3D face-vertex mesh $\mathscr{M} = \{\mathscr{V}, \mathscr{F}\}$ of $|\mathscr{V}|$ vertices and $|\mathscr{F}|$ faces (i.e. triangles). The mean shape model of the whole lumbar spine, i.e. a chain of mean shape models of individual vertebrae, is further used for spine detection, while the mean shape models of individual vertebrae are used for vertebra detection and segmentation in an unknown 3D image I.

2.2 Vertebra Detection

The detection of vertebrae in an unknown 3D image I was performed by a novel optimization scheme that is based on interpolation theory [4]. The optimization scheme consists of three steps: spine detection, vertebra detection and vertebra alignment. To detect the spine in image I, the pose of mesh $\mathscr{M}$ representing the mean shape model of the lumbar spine (i.e. a chain of meshes representing individual vertebrae from L1 to L5) is optimized against three translations (i.e. coordinates x, y and z representing sagittal, coronal and axial anatomical directions, respectively). The global maximum of the resulting interpolation represents the location of the spine in the 3D image, and is further used to initialize the detection of individual vertebrae. To detect individual vertebrae, the pose of mesh $\mathscr{M}$, now representing the mean shape model of the observed lumbar vertebra, is optimized against three translations, however, in this case all local maxima of the resulting interpolation are extracted, corresponding to locations of the observed and neighboring vertebrae. The correct location of each vertebra is determined by the optimal path that passes through a set of locations, where each location belongs to a local maxima at a different vertebral level. Finally, a more accurate alignment of the mean shape model of each observed vertebra is performed by optimizing the pose of each model against three translations, one scaling

(i.e. factor s) and three rotations (i.e. angles φ_x, φ_y and φ_z about coordinate axes x, y and z, respectively). The resulting alignment represents the final vertebra detection result. A detailed description of the interpolation-based optimization scheme can be found in [4].

2.3 Vertebra Segmentation

After the interpolation-based alignment [4] of the mean shape model of each lumbar vertebra to the unknown image I, segmentation of each lumbar vertebra is performed by a mesh deformation technique that moves mesh vertices to their optimal locations while preserving the underlying vertebral shape [9, 15]. In this iterative procedure, the following two steps are executed in each iteration: image object detection for mesh face centroids that are represented by the centers of mass for mesh faces $\mathscr{F} \in \mathscr{M}$, followed by reconfiguration of mesh vertices $\mathscr{V} \in \mathscr{M}$. By combining the robust initialization resulting from vertebra detection (Sect. 2.2) with modifications to the mesh deformation technique, we improve the accuracy of the resulting vertebra segmentation.

2.3.1 Object Detection

By displacing each mesh face centroid $\mathbf{c}_i; i = 1, 2, \ldots, |\mathscr{F}|$ along its corresponding mesh face normal $\mathbf{n}(\mathbf{c}_i)$, a new candidate mesh face centroid $\mathbf{c}_i^*$ is found in each kth iteration:

$$\mathbf{c}_i^* = \mathbf{c}_i + \delta\, j_i^*\, \mathbf{n}(\mathbf{c}_i), \tag{1}$$

where δ is the length of the unit displacement, and j_i^* is an element from set $\mathscr{J}$; $j_i^* \in \mathscr{J}$. Set $\mathscr{J}$ represents the search profile along $\mathbf{n}(\mathbf{c}_i)$, called the sampling parcel (Fig. 1):

$$\mathscr{J} = \Big\{ -j, -j+1, \ldots, j-1, j \Big\}; \quad j = J - k + 1, \tag{2}$$

which is of size $2J+1$ at initial iteration $k = 1$ and $2(J-K+1)+1$ at final iteration $k = K$. The element j_i^* that defines the location of $\mathbf{c}_i^*$ is determined by detecting vertebra boundaries:

$$j_i^* = \underset{j \in \mathscr{J}}{\arg\max} \Big\{ F\big(\mathbf{c}_i, \mathbf{c}_i + \delta\, j\, n(\mathbf{c}_i)\big) - D\, \delta^2\, j^2 \Big\}. \tag{3}$$

where $\mathbf{c}_i' = \mathbf{c}_i + \delta\, j_i\, \mathbf{n}(\mathbf{c}_i)$ is the candidate location for $\mathbf{c}_i^*$ (Eq. 1), and parameter D controls the tradeoff between the response of the boundary detection operator

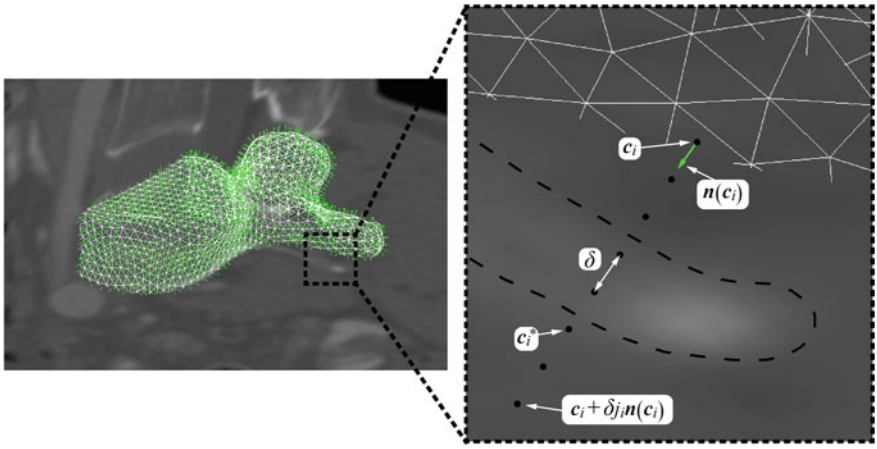

Fig. 1 Each ith face centroid $\mathbf{c}_i$ of the 3D face-vertex mesh of the observed vertebra is displaced for δj_i along the sampling parcel in the direction of its face normal $\mathbf{n}(\mathbf{c}_i)$. In an iterative framework, where the length of the sampling parcel is gradually reduced, each centroid moves to the location $\mathbf{c}_i^*$ that best corresponds to vertebra boundaries (*dashed curve*)

F (Eq. 4) and the distance from $\mathbf{c}_i$ to $\mathbf{c}'_i$. In comparison to the original approach [9, 15], we propose an improved boundary detection operator F that is based on image intensity gradients, weighted by an image appearance operator:

$$F(\mathbf{c}_i, \mathbf{c}'_i) = \frac{g_{\max}\left(g_{\max} + \left\|\mathbf{g}_W(\mathbf{c}'_i)\right\|\right)}{g_{\max}^2 + \left\|\mathbf{g}_W(\mathbf{c}'_i)\right\|^2} \left\langle \mathbf{n}(\mathbf{c}_i), \mathbf{g}_W(\mathbf{c}'_i) \right\rangle, \tag{4}$$

where $\|\cdot\|$ denotes the vector norm, $\langle \cdot, \cdot \rangle$ denotes the dot product, $g_{\max}$ is the estimated mean amplitude of intensity gradients at vertebra boundaries that is used to suppresses the weighted gradients, which may occur if the gradient magnitude at the boundary of the object of interest is considerably smaller than of another object in its neighborhood (e.g. pedicle screws), and $\mathbf{g}_W$ is the image appearance operator at candidate mesh centroid location $\mathbf{c}'_i$:

$$\mathbf{g}_W(\mathbf{c}'_i) = \left(1 + C(\mathbf{c}'_i)\right) \mathbf{g}(\mathbf{c}'_i), \tag{5}$$

where $\mathbf{g}(\mathbf{c}'_i)$ is the intensity gradient at $\mathbf{c}'_i$ and $C(\mathbf{c}'_i) \in [0, 1]$ is the continuous response to the Canny edge operator [2]. By adding additional weights to the image intensity gradients, vertebra boundary points are more likely to be detected. In contrast to the original technique [9, 15], the size of the sampling parcel J (Eq. 2) is reduced in each iteration k and the image intensity gradients $\mathbf{g}$ (Eq. 5) are additionally weighted, both to improve the accuracy of the resulting segmentation.

2.3.2 Mesh Reconfiguration

Once the new candidate mesh face centroids $\mathbf{c}_i^*$ are detected, mesh $\mathscr{M} = \{\mathscr{V}, \mathscr{F}\}$ is reconfigured in each kth iteration by minimizing the weighted sum E of energy terms:

$$\min_{\mathscr{M}}\{E\} = \min_{\mathscr{M}}\{E_{\text{ext}} + \alpha E_{\text{int}}\}, \tag{6}$$

where α is the weighting parameter. The external energy E_{ext} attracts mesh $\mathscr{M}$ to new face centroids $\mathbf{c}_i^*$, $i = 1, 2, \ldots, |\mathscr{F}|$ (Eq. 1), that are located on vertebra boundaries:

$$E_{\text{ext}} = \sum_{i=1}^{|\mathscr{F}|} w_i^* \left\langle \mathbf{c}_i^* - \mathbf{c}_i, \frac{\mathbf{g}_W(\mathbf{c}_i^*)}{\|\mathbf{g}_W(\mathbf{c}_i^*)\|} \right\rangle^2 \tag{7}$$

where $|\mathscr{F}|$ is the number of mesh faces, $\mathbf{g}_W$ is the image appearance operator (Eq. 5), and w_i; $i = 1, 2, \ldots, |\mathscr{F}|$, are weights that are defined according to the obtained j_i^* (Eq. 3) to give a greater influence to more promising centroid locations:

$$w_i^* = \max\left\{0,\, F(\mathbf{c}_i, \mathbf{c}_i^*) - D\,\delta^2\, j_i^{*2}\right\} \tag{8}$$

The internal energy E_{int} restricts the flexibility of mesh $\mathscr{M}$ by penalizing the deviation between deformation vertices $\mathscr{V}$ and mean vertices $\mathscr{V}^m$:

$$E_{\text{int}} = \sum_{i=1}^{|\mathscr{V}|} \sum_{j \in \mathscr{N}_i} \left\| \left(\mathbf{v}_i - \mathbf{v}_j\right) - \left(s R \left(\mathbf{v}_i^m - \mathbf{v}_j^m\right) + \mathbf{t}\right) \right\|^2 \tag{9}$$

where $\mathbf{v}_i$ and $\mathbf{v}_i^m$ are vertices from sets $\mathscr{V}$ and $\mathscr{V}^m$, respectively, $\mathscr{M}^m = \{\mathscr{V}^m, \mathscr{F}^m\}$ represents the mean shape model of the observed lumbar vertebra (Sect. 2.1), and $\mathscr{N}_i$ is the set of vertices neighboring to $\mathbf{v}_i$ (or $\mathbf{v}_i^m$, since the topology is preserved). The scaling factor s, rotation matrix R and translation vector $\mathbf{t}$ that align mesh vertices $\mathbf{v}_i$ to the mean vertices $\mathbf{v}_i^m$ are determined prior to calculation of Eq. (9) by using Procrustes superimposition [3].

3 Experiments and Results

The performance of the described vertebra segmentation framework was tested on a database of 10 CT images of the lumbar spine (i.e. a total of 50 lumbar vertebrae, with axial in-plane pixel size of 0.3–0.8 mm and cross-sectional thickness of 0.7–1.5 mm) by applying a leave-one-out evaluation scheme. A reference segmentation binary mask was available for each vertebra in the database. The framework

was implemented in Matlab, and executed on a personal computer with Intel Core i5 processor at 3.2 GHz and 16 GB of memory without a graphics processing unit.

3.1 Experimental Details

The mean shape model of the lumbar spine (Sect. 2.1) was obtained by applying the marching cubes algorithm [11] to binary masks representing reference segmentations of each vertebra, resulting in 3D face-vertex meshes $\mathscr{M} = \{\mathscr{V}, \mathscr{F}\}$ of genus 1 (i.e. the number of holes is 1, as expected for lumbar vertebrae). The corresponding number of vertices $|\mathscr{V}| = 31.542$–161.790 (the number of faces was $|\mathscr{F}| = 2\,|\mathscr{V}|$) was further reduced to $|\mathscr{V}| = 3.228$–5.642 by isotropic remeshing with mean edge length of 2.25 mm [1]. After establishing correspondences among meshes of the same vertebral level by the coherent point drift algorithm [13] and applying the generalized Procrustes alignment [3], the mean shape model of each lumbar vertebra was obtained. The initialization of vertebra segmentation was obtained from the results of interpolation-based vertebra detection [4] based on spline approximation on an equidistant grid. Segmentation of each vertebra (Sect. 2.3) consisted of 25 iterations (from $k = 1$ to $k = 25$) of object detection and mesh reconfiguration. The corresponding parameters, which were obtained from the original mesh deformation technique [9, 15] and not further tuned to search for a possibly better segmentation performance on the tested database, were set to $J = 25$ (Eq. 2), $D = 0.6\,\text{mm}^{-2}$ (Eqs. 3 and 8), $\delta = 0.3$ mm (Eqs. 1, 3 and 8), $g_{\max} = 100$ HU (Eq. 4) and $\alpha = 33$ (Eq. 6). The minimization of the sum of energy terms (Eq. 6) was performed by the conjugate gradient method.

3.2 Results

The performance of the proposed framework was evaluated by the mean symmetric absolute surface distance (MSD), symmetric root-mean-square surface distance (RMSSD), maximal symmetric absolute surface distance (MaxSD) and Dice coefficient (DICE), computed between the resulting 3D meshes and corresponding reference segmentation binary masks. Detailed results for the segmentation of individual vertebral levels are presented in Table 1 separately for the original mesh deformation framework [9, 15] and for the proposed framework that is based on a robust initialization and additional modifications with the aim to improve the framework performance. The overall vertebral segmentation performance (mean ± standard deviation) was MSD $= 0.43 \pm 0.14$ mm, RMSSD $= 0.83 \pm 0.33$ mm, MaxSD $= 7.32 \pm 3.23$ mm and DICE $= 93.76 \pm 1.61$ % for the proposed framework, compared to MSD $= 0.55 \pm 0.21$ mm, RMSSD $= 1.10 \pm 0.47$ mm, MaxSD $= 9.65 \pm 4.37$ mm and DICE $= 92.19 \pm 2.19$ % of the original framework. The detection of all five lumbar vertebrae (i.e. levels from T1 to T5) took on average

Table 1 Lumbar vertebra segmentation results in terms of mean symmetric absolute surface distance (MSD), symmetric root-mean-square surface distance (RMSSD), maximal symmetric absolute surface distance (MaxSD) and Dice coefficient (DICE), reported as mean ± standard deviation

Vertebral level	MSD (mm)	RMSSD (mm)	MaxSD (mm)	DICE (%)
Original framework [9, 15]				
L1	0.46 ± 0.13	0.92 ± 0.34	8.67 ± 4.11	93.37 ± 1.49
L2	0.42 ± 0.11	0.78 ± 0.23	6.79 ± 2.80	93.63 ± 1.24
L3	0.54 ± 0.14	1.13 ± 0.48	11.60 ± 5.81	92.60 ± 1.21
L4	0.67 ± 0.19	1.34 ± 0.40	11.04 ± 3.07	90.94 ± 1.98
L5	0.69 ± 0.28	1.32 ± 0.62	10.17 ± 4.41	90.43 ± 2.79
Proposed framework				
L1	0.34 ± 0.09	0.69 ± 0.32	6.52 ± 4.04	94.83 ± 0.90
L2	0.36 ± 0.08	0.66 ± 0.16	5.72 ± 1.52	94.47 ± 0.82
L3	0.40 ± 0.11	0.76 ± 0.26	6.28 ± 2.13	94.06 ± 0.97
L4	0.52 ± 0.12	1.05 ± 0.27	9.56 ± 2.26	92.73 ± 1.44
L5	0.51 ± 0.19	0.99 ± 0.44	8.52 ± 4.08	92.71 ± 2.28

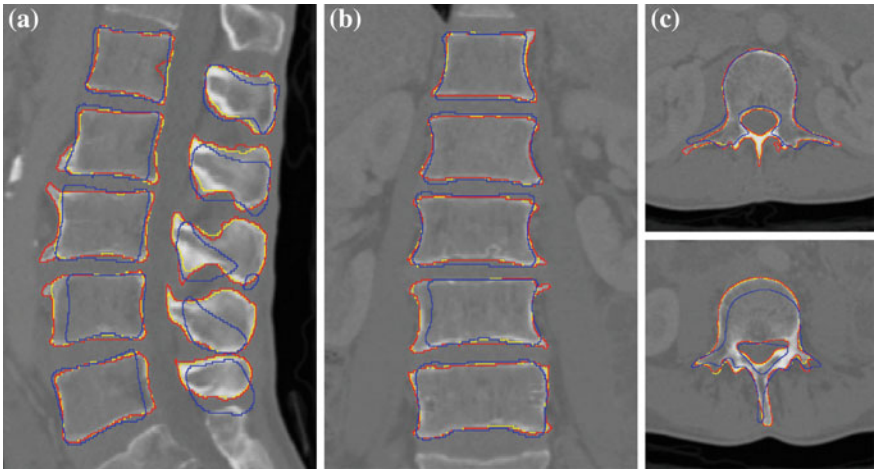

Fig. 2 An example of vertebra segmentation initialization (in *blue*) and vertebra segmentation results (in *yellow*) in comparison to reference segmentation (in *red*) for a selected CT lumbar spine image, shown in selected **a** mid-sagittal, **b** mid-coronal and **c** mid-axial cross-sections (Color figure online)

around 220 s, while the segmentation of each individual vertebra took on average around 1 min. An example of the resulting segmentation is for a selected CT lumbar spine image shown in Fig. 2.

4 Discussion and Conclusion

In this paper, we combined robust detection of the object of interest [4] with an improved shape-constrained deformable model to segment vertebrae from CT images of the lumbar spine. Vertebra segmentation from 3D spine images has been already addressed in several studies [6–10, 12, 14]. The best performance in terms of accuracy was reported by Klinder et al. [9], who reported a mean point-to-surface error (i.e. one surface is represented by a set of surface points and the other by a surface mesh model) of 0.76 mm by applying the original shape-constrained deformable model technique, but also progressively adapted tube-shaped segments to extract the spine curve, performed vertebra detection on curved-planar reformatted images using the generalized Hough transform, and identified vertebrae by rigid registration of appearance models to the detected candidates. On the other hand, Kadoury et al. [6, 7] reported the highest Dice coefficient, i.e. of 92.5 %, which was obtained by building an articulated shape manifold from a training database and embedding the data into a low-dimensional sub-space, followed by the Markov random field optimization to infer between the unseen target shape and shape manifold. Although the overall results of the proposed method of 0.43 ± 0.14 mm in terms of MSD and 93.76 ± 1.61 % in terms of the Dice coefficient can not be directly compared to the results reported by the existing studies because of different evaluation methodologies and data collection techniques, as well as because of different databases, we can conclude that the proposed automated spine and vertebra detection segmentation framework produces accurate results. Moreover, to objectively compare the effects of the performed modifications of the original shape-constrained deformable models, we report vertebra segmentation results also for the original technique obtained on the same database of CT lumbar spine images. From Table 1 it can be observed that an improvement of around 20 % in terms of MSD in favor of the proposed framework was achieved. It can be therefore concluded that the performed modifications improved the accuracy of vertebra segmentation, and that when combined with a robust initialization, the proposed framework can accurately segment vertebrae from CT images of the lumbar spine.

Acknowledgments This work was supported by the Slovenian Research Agency (ARRS) under grants P2-0232 and L2-4072.

References

1. Botsch, M., Kobbelt, L.: A remeshing approach to multiresolution modeling. In: Proceedings of the 2004 Eurographics/ACM SIGGRAPH symposium on geometry processing, SGP '04, pp. 185–192 (2004)
2. Canny, J.F.: A computational approach to edge detection. IEEE Trans. Pattern Anal. Mach. Intell. **8**(6), 679–698 (1986)
3. Dryden, I.L., Mardia, K.V.: Statistical shape analysis. Wiley, Chichster (1998)

4. Ibragimov, B., Korez, R., Likar, B., Pernuš, F., Vrtovec, T.: Interpolation-based detection of lumbar vertebrae in CT spine images. In: Proceedings of the 2nd MICCAI Workshop on Computational Methods and Clinical Applications for Spine Imaging—MICCAI CSI 2014, pp. 65–75 (2014)
5. Ibragimov, B., Likar, B., Pernuš, F., Vrtovec, T.: Shape representation for efficient landmark-based segmentation in 3D. IEEE Trans. Med. Imaging **33**(4), 861–874 (2014)
6. Kadoury, S., Labelle, H., Paragios, N.: Automatic inference of articulated spine models in CT images using high-order Markov random fields. Med. Image Anal. **15**(4), 426–437 (2011)
7. Kadoury, S., Labelle, H., Paragios, N.: Spine segmentation in medical images using manifold embeddings and higher-order MRFs. Med. Image Anal. **32**(7), 1227–1238 (2013)
8. Kim, Y., Kim, D.: A fully automatic vertebra segmentation method using 3D deformable fences. Comput. Med. Imaging Graph. **33**(5), 343–352 (2009)
9. Klinder, T., Ostermann, J., Ehm, M., Franz, A., Kneser, R., Lorenz, C.: Automated model-based vertebra detection, identification, and segmentation in CT images. Med. Image Anal. **13**(3), 471–482 (2009)
10. Lim, P.H., Bagci, U., Bai, L.: Introducing Willmore flow into level set segmentation of spinal vertebrae. IEEE Trans. Biomed. Eng. **60**(1), 115–122 (2013)
11. Lorensen, W.E., Cline, H.E.: Marching cubes: a high resolution 3D surface construction algorithm. Comput. Graph. **21**(4), 163–169 (1987)
12. Ma, J., Lu, L.: 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 (2013)
13. Myronenko, A., Song, X.: Point set registration: coherent point drift. IEEE Trans. Pattern Anal. Mach. Intell. **32**(12), 2262–2275 (2010)
14. Rasoulian, A., Rohling, R., Abolmaesumi, P.: Lumbar spine segmentation using a statistical multi-vertebrae anatomical shape+pose model. IEEE Trans. Med. Imaging **32**(10), 1890–1900 (2013)
15. Weese, J., Kaus, M., Lorenz, C., Lobregt, S., Truyen, R., Pekar, V.: Shape constrained deformable models for 3D medical image segmentation. Inf. Process. Med. Imaging Lect. Notes Comput. Sci. **2082**, 380–387 (2001)

Detailed Vertebral Segmentation Using Part-Based Decomposition and Conditional Shape Models

Marco Pereañez, Karim Lekadir, Corné Hoogendoorn, Isaac Castro-Mateos and Alejandro Frangi

Abstract With the advances in minimal invasive surgical procedures, accurate and detailed extraction of the vertebral boundaries is required. In practice, this is a difficult challenge due to the highly complex geometry of the vertebrae, in particular at the processes. This paper presents a statistical modeling approach for detailed vertebral segmentation based on part decomposition and conditional models. To this end, a Vononoi decomposition approach is employed to ensure that each of the main subparts the vertebrae is identified in the subdivision. The obtained shape constraints are effectively relaxed, allowing for an improved encoding of the fine details and shape variability at all the regions of the vertebrae. Subsequently, in order to maintain the statistical coherence of the ensemble, conditional models are used to model the statistical inter-relationships between the different subparts. For shape reconstruction and segmentation, a robust model fitting procedure is introduced to exclude outlying inter-part relationships in the estimation of the shape parameters. The experimental results based on a database of 30 CT scans show significant improvement in accuracy with respect to the state-of-the-art and the potential of the proposed technique for detailed vertebral modeling.

M. Pereañez (✉)
Center for Computational Imaging and Simulation Technologies in Biomedicine (CISTIB), Universitat Pompeu Fabra, 08018 Barcelona, Spain
e-mail: marco.pereanez@upf.edu

K. Lekadir · C. Hoogendoorn
CISTIB, Universitat Pompeu Fabra, 08018 Barcelona, Spain
e-mail: karim.lekadir@upf.edu

C. Hoogendoorn
e-mail: corne.hoogendoorn@upf.edu

I. Castro-Mateos · A. Frangi
CISTIB, University of Sheffield, S1 3JD, Sheffield, UK
e-mail: isaac.casm@sheffield.ac.uk

A. Frangi
e-mail: a.frangi@sheffield.ac.uk

© Springer International Publishing Switzerland 2015

J. Yao et al. (eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, Lecture Notes in Computational Vision and Biomechanics 20, DOI 10.1007/978-3-319-14148-0_9

1 Introduction

Automatic segmentation of the vertebrae is an important pre-requisite for a number of clinical applications, ranging from the assessment of spinal disorders to image-guided interventions. The latter one in particular, with the recent advances in minimal invasive surgical procedures, requires accurate and detailed extraction of the vertebral boundaries. However, this is challenging in practice due to the highly complex geometry of the vertebrae, in particular at the region of the processes. Figure 1 shows some examples of typical areas of high geometrical complexity and curvature in the lumbar vertebra L5.

Amongst existing techniques for vertebral image segmentation, statistical models of 3D shape [4] have been extensively used [2, 3, 5, 7, 8] due to their ability to build a shape prior from a representative training population. However, these methods consider at best a whole vertebra as the smallest unit for the construction of the point distribution models (PDMs). Due to the large variability of the vertebrae in particular at the processes and the generally small number of samples available for training, the obtained models are too constraining and not flexible enough to localize the fine details at areas of high curvatures.

In this paper, we present a new method for detailed modeling and segmentation of the vertebrae. The fundamental idea behind the proposed technique is to decompose each vertebra into a set of subparts based on their geometrical properties. By using a Vononoi [9] decomposition approach, we ensure that each of the main subparts of the vertebrae is well identified. With this approach, the shape constraints are effectively relaxed, allowing for an improved encoding of the fine details and shape variability at all the regions of the structures. Subsequently, in order to maintain the statistical coherence of the ensemble, conditional models are used to model the statistical inter-relationships between the different subparts. For shape reconstruction and segmentation, a robust model fitting procedure is introduced to exclude outlying

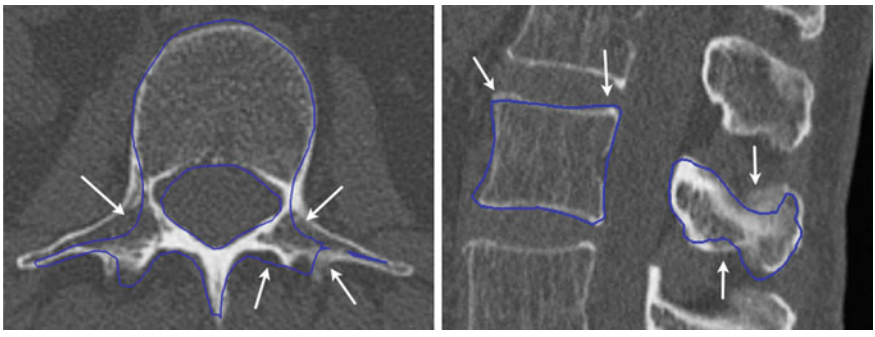

Fig. 1 Examples of segmentations (in *blue*) obtained with local PDMs, showing suboptimal fitting in areas of complex geometry and high curvature on the spine (Color figure online)

inter-part relationships in the estimation of the shape parameters. The proposed technique is validated with a total of 30 spinal CT scans.

2 Method

The proposed framework consists of three main stages. Firstly, in Sect. 2.1, a subdivision of each vertebra into a number of subparts is proposed based on an approximate Voronoi region decomposition. Subsequently, the conditional models describing the statistical inter-relationships between the subparts are presented in Sect. 2.2. Finally, a model fitting approach based on all pair wise conditional models is introduced in Sect. 2.3, with the aim to estimate the shape parameters for each subpart robustly.

2.1 Vertebral Decomposition

Let us denote $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)^T$ the landmark-based shape representation of each vertebra, where n is the number of landmarks used to discretize the 3D shape. The aim of this section is to obtain a subdivision of $\mathbf{x}$ into K sub-components.

We do this in this paper by using a polygon clustering algorithm described in [9], which provides a compact subdivision of the shape based on the concept of a Voronoi diagram [1]. Furthermore, let us denote $V = (C_1, \ldots, C_m)^T$ the triangulation of the shape $\mathbf{x}$, where $C_i, i = 1, \ldots, m$, represents each face on the mesh, and E_j the set of edges between all adjacent triangles. Given the centroids c_i corresponding to each of the triangular triangles C_i, the algorithm computes an approximation of a centroidal Voronoi diagram (CVD). The energy term to minimize is:

$$F = \sum_{k=1}^{K} \left(\sum_{i \in R_k} w_i ||c_i - c_{R_k}||^2 \right), \tag{1}$$

where R_k is a subset of V (i.e. subpart of the shape), c_{R_k} is the center of the region, and w_i is the area of triangle C_i.

To minimize Eq. 1, we use an iterative approach over the subset of edges E_j between adjacent regions. The regions R_k are initialized as a single triangle that is randomly chosen amongst V. The remaining triangles are assigned to the *null* region R_0. We then iterate only over the edges that are between two regions R_k and R_l, or between any R_k and R_0. Then we assign one of the two triangles adjacent to the current edge to the region that minimizes Eq. 1. At some point, the region R_0 will become empty. The iterative algorithm will continue until no modification of the region assignments for the triangles lead to an improvement of the subdivision according to Eq. 1.

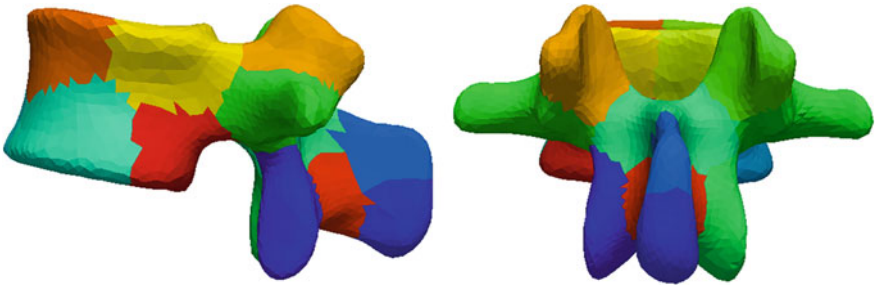

Fig. 2 The obtained Voronoi decomposition with 18 subparts

Given that the regions of highest complexity and curvature on the vertebra are the different vertebral processes, we chose the lowest number of clusters K such that points selected roughly at the distal region of every process (transverse, superior/inferior articular and spinous) are assigned to a different region after the subdivision.

Once the algorithm converges, all $C_i \in V$ will belong to a unique region R_i. However, the points that lie on the boundary edges between subparts will now belong to two adjacent regions. To resolve this ambiguity, we perform a last step where we go through the different regions sequentially and assign boundary points to the current region unless previously assigned.

Figure 2 shows the obtained subdivision with $K = 18$, where it can be seen that the main regions of high curvature now belong to a unique subpart.

2.2 Conditional Model Parametrization

In the previous section we obtained K subcomponents $\mathbf{x}_k$, $k = 1, \ldots, K$. The aim of this section is to describe the statistical modeling of the inter-part probability distributions, i.e. $P(\mathbf{x}_k|\mathbf{x}_l)$, where $k, l = 1, \ldots, K$ and $k \neq l$. More specifically, we would like to obtain new constraints for each part $\mathbf{x}_k$ based on its conditional relationship with $\mathbf{x}_l$, that is, a new mean and covariance in the space of the shape parameters $\mathbf{b}_k$. Let us denote μ_{kl} and Σ_{kl} the values that form the new conditional constraints. In this paper, we choose to model $P(\mathbf{x}_k, \mathbf{x}_l)$ using a normal probability distribution. Thus, the mean and the covariance estimates are calculated as:

$$\mu_{kl} = \Sigma_{kl} \Sigma_{ll}^{-1} \mathbf{b}_l \tag{2}$$

$$\Sigma_{kl} = \Sigma_{kk} - \Sigma_{kl} \Sigma_{ll}^{-1} \Sigma_{lk} \tag{3}$$

$\mathbf{b}_l$ in Eq. 2 is the parametric representation after eigendecomposition of the a given conditioning shape $\mathbf{S}_l$, and the covariance matrices in Eqs. 2 and 3 are obtained from the partitioned covariance matrix in Eq. 4:

$$\Sigma = \begin{bmatrix} \Sigma_{kk} & \Sigma_{kl} \\ \Sigma_{lk} & \Sigma_{ll} \end{bmatrix} \tag{4}$$

In Eqs. 2 and 3 $\Sigma_{kl}\Sigma_{ll}^{-1}$ are called the matrix regression coefficients of $\mathbf{b}_k$ on $\mathbf{b}_l$.

In order to compute the conditional mean μ and covariance matrix Σ, we need to compute the inverse of the covariance matrix of the predictor shape, however, as the dimensionality of the shapes is much larger than the number of training samples available, the covariance matrix becomes singular, and cannot be inverted. Also, the computational burden of computing inverse of matrices representing several thousands of points can become cumbersome. We address this issues by reducing the dimensionality of the problem using PCA before computation of the mean and covariance matrix as follows [6]:

given subshapes $\mathbf{x}_k$ and $\mathbf{x}_l$,

$$\mathbf{x}_k = \bar{\mathbf{x}}_k + \Phi_k \mathbf{b}_k \tag{5}$$

$$\mathbf{x}_l = \bar{\mathbf{x}}_l + \Phi_l \mathbf{b}_l, \tag{6}$$

their parametric representation is

$$\mathbf{b}_k = \Phi_k^T \mathbf{x}_k - \bar{\mathbf{x}}_k \tag{7}$$

$$\mathbf{b}_l = \Phi_l^T \mathbf{x}_l - \bar{\mathbf{x}}_l. \tag{8}$$

Then the cross-covariance matrix $\Sigma_{kl} = \mathbf{B}_k \mathbf{B}_l^T$ is the product of parametric shapes matrices $\mathbf{B}_k$ and $\mathbf{B}_l$. The self-covariance matrices Σ_{kk} and Σ_{ll} are the diagonal eigenvalue matrices Λ_k and Λ_l obtained by eigendecomposition of the individual subparts.

The proposed parametrization of the conditional model has two important benefits. Firstly, it decreases the over-constraining of the global model caused by the dimensionality disparity between the available samples and the natural variabilty of the shapes. Additionally, and as detailed in next section, the inter-part models can be used as a mechanism to find the optimal domain of valid subregions and exclude incorrect localized segmentations due to image inhomogeneities.

2.3 Robust Model Fitting

To maintain the coherence of the ensemble in spite of the decomposition, the estimation of the shape parameters must be carried out by considering all pairwise

conditional probabilities $P(\mathbf{x}_k|\mathbf{x}_l)$. This is challenging at the segmentation stage because all subparts are being optimized simultaneously, and therefore, there is a degree of uncertainty surrounding the values of the different $\mathbf{x}_l$ in $P(\mathbf{x}_k|\mathbf{x}_l)$. This could lead to inaccurate constraining and parameter estimation of $\mathbf{x}_k$ if some of the $\mathbf{x}_l, l = 1, \ldots, K, k \neq l$ are erroneous during the segmentation procedure. To address this problem, we use a median-based estimation approach to exclude potentially incorrect conditional relationships.

Firstly, we calculate the initial shape parameter $\mathbf{b}_k^0$ by projecting the boundary feature points (as obtained using normal search profile) onto the standard PDM of $\mathbf{x}_k$. Subsequently, we calculate $K-1$ shape parameters $\mathbf{b}_{kl}$ by considering the $K-1$ shape constraints formed by the conditional mean parameter μ_{kl} and its corresponding bounds λ_{kl} (derived from the eigenvalues of Σ_{kl}), i.e.,

$$\mathbf{b}_{kl} = \begin{cases} \mathbf{b}_k^0 & \text{if } |\mathbf{b}_k^0 - \mu_{kl}| \leq 3\sqrt{\lambda_{kl}} \\ \mu_{kl} + 3\sqrt{\lambda_{kl}} & \text{if } \mathbf{b}_k > \mu_{kl} + 3\sqrt{\lambda_{kl}} \\ \mu_{kl} - 3\sqrt{\lambda_{kl}} & \text{if } \mathbf{b}_k < \mu_{kl} - 3\sqrt{\lambda_{kl}} \end{cases} \tag{9}$$

Due to the fact that some subparts are inevitably erroneous during the image search due to imaging inhomogeneities, some of the $\mathbf{b}_{kl}$ values will be incorrect. To exclude these values and obtain a consensus robust estimation of the shape parameters, we use a median-based final estimation of $\mathbf{b}_{kl}$, i.e.,

$$\mathbf{b}_k^{final} = median(\mathbf{b}_{kl}). \tag{10}$$

3 Validation

We validate our method using 30 image volumes of the lumbar spine (L1-L5) from CT scans. The image datasets were collected at the National Center for Spinal Disorders (Budapest, Hungary). The images have an in-plane resolution of 0.6×0.6 mm and slice thickness of 0.62 mm. All images were manually segmented using open source software.

All segmentations were performed by preserving 98 % of the total variance, and allowing ± 3 standard deviations from the mean. All segmentations are performed following a leave-one-out scheme. Accuracy is measured as the RMS point to surface distance between the manual segmentations and the reconstructions. Furthermore, we compare the proposed approach against the results obtained with a standard ASM using single-vertebra PDMs.

Figure 3 shows the segmentation errors for all the 30 scans using both ASM methods. It is evident that the proposed technique outperforms the single model ASMs for nearly all cases (with the exception of case 24, with minor differences). The average improvement is of 16 % and in some cases the improvement is over

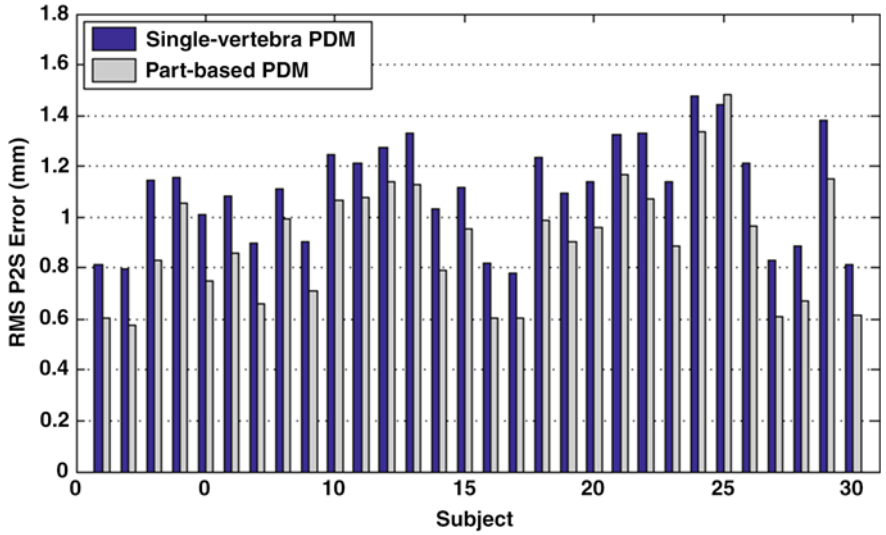

Fig. 3 Point to surface segmentation error comparison between the proposed method and a standard ASM single-vertebra PDM

Table 1 Image segmentation errors (mm) comparing the performance of our parts-based models (pb), and a single-vertebra PDM (sv). Errors are shown individually for each lumbar vertebra

Structure	L1		L2		L3		L4		L5	
	pb	sv	**pb**	sv	**pb**	sv	**pb**	sv	**pb**	sv
Mean	**0.81**	1.05	**0.82**	1.08	**0.84**	1.05	**0.88**	1.12	**1.06**	1.23
± Std	**0.13**	0.15	**0.15**	0.23	**0.17**	0.19	**0.20**	0.25	**0.24**	0.18

20 % due to the ability of the proposed technique to better encode the fine details of the vertebrae.

Table 1 summarizes the segmentation results for the proposed part-based technique (pb) and the single vertebra ASM (sv) for the different lumbar vertebrae (L1 to L5). It can be seen that the performance of the proposed technique is consistently better for the entire lumbar spine.

Finally, we show in Fig. 4 two illustrations of the error distribution for both the standard ASM and the proposed technique. It can be seen that the errors introduced locally by the use of a single vertebral model are corrected by the proposed parts-based approach. For both examples, the errors are consistently low in all regions of the vertebra.

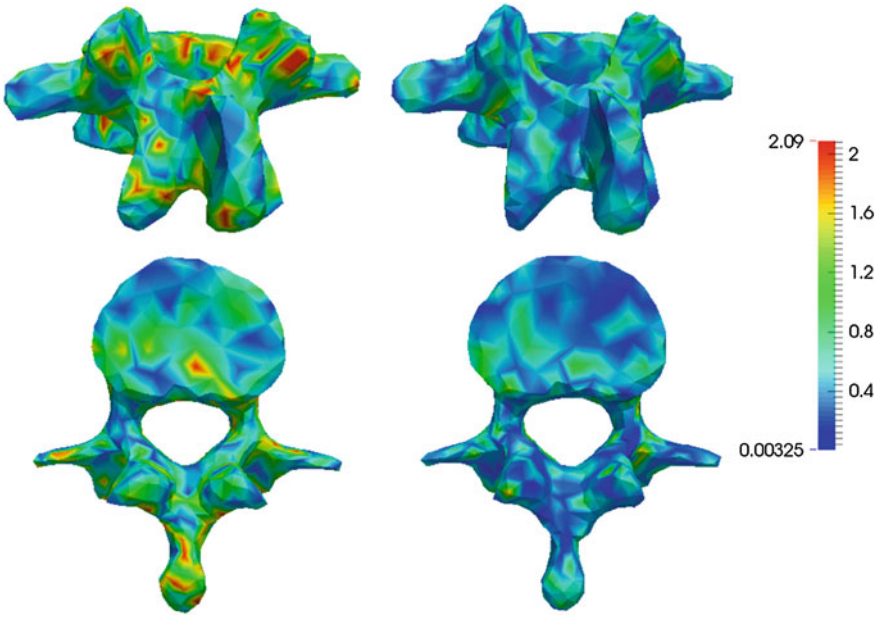

Fig. 4 Point to surface segmentation error comparison between the proposed method (*right top/bottom*), and a standard ASM single-vertebra PDM (*left top/bottom*)

4 Conclusions

In this paper, we presented a new part-based ASM approach for detailed segmentation of the lumbar vertebrae. The proposed technique addresses the difficulty to model the variability in the area of high complexity and curvature by decomposing the vertebrae into a set of subparts, which are subsequently linked using conditional shape models. A robust median-based estimation of the shape parameters of each subpart is used to minimize potential errors due to the presence of image inhomogeneities. The results indicate potential for more detailed localization of the fine details of the vertebrae. Future work include the study of the effect of the number of subparts on the models and segmentation properties.

References

1. Aurenhammer, F.: Voronoi diagrams: a survey of a fundamental geometric data structure. ACM Comput. Surv. (CSUR) **23**(3), 345–405 (1991)
2. Benameur, S., Mignotte, M., Parent, S., Labelle, H., Skalli, W., de Guise, J.: 3d/2d registration and segmentation of scoliotic vertebrae using statistical models. Comput. Med. Imaging Graph. **27**(5), 321–337 (2003)

3. de Bruijne, M., Nielsen, M.: Image segmentation by shape particle filtering. ICPR **3**, 722–725 (2004)
4. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models—their training and application. Comput. Vis. Image Underst. **61**(1), 38–59 (1995). doi:10.1006/cviu.1995.1004
5. Kadoury, S., Labelle, H., Paragios, N.: Spine segmentation in medical images using manifold embeddings and higher-order mrfs. Med. Imaging, IEEE Trans. **32**(7), 1227–1238 (2013)
6. Metz, C., Baka, N., Kirisli, H., Schaap, M., van Walsum, T., Klein, S., Neefjes, L., Mollet, N., Lelieveldt, B., de Bruijne, M., et al.: Conditional shape models for cardiac motion estimation. In: Medical Image Computing and Computer-Assisted Intervention-MICCAI 2010, pp. 452–459. Springer (2010)
7. Rasoulian, A., Rohling, R., Abolmaesumi, P.: Lumbar spine segmentation using a statistical multi-vertebrae anatomical shape+pose model. Med. Imaging IEEE Trans. **32**(10), 1890–1900 (2013). doi:10.1109/TMI.2013.2268424
8. Roberts, M.G., Cootes, T.F., Adams, J.E.: Linking sequences of active appearance sub-models via constraints: An application in automated vertebral morphometry. In: BMVC, pp. 1–10 (2003)
9. Valette, S., Chassery, J.M.: Approximated centroidal Voronoi diagrams for uniform polygonal mesh coarsening **2**4(3), 381–389 (2004). doi:10.1111/j.1467-8659.2004.00769.x

Part III
MR Image Processing

Automatic Segmentation of the Spinal Cord Using Continuous Max Flow with Cross-sectional Similarity Prior and Tubularity Features

Simon Pezold, Ketut Fundana, Michael Amann, Michaela Andelova, Armanda Pfister, Till Sprenger and Philippe C. Cattin

Abstract Segmenting tubular structures from medical image data is a common problem; be it vessels, airways, or nervous tissue like the spinal cord. Many application-specific segmentation techniques have been proposed in the literature, but only few of them are fully automatic and even fewer approaches maintain a convex formulation. In this paper, we show how to integrate a cross-sectional similarity prior into the convex continuous max-flow framework that helps to guide segmentations in image regions suffering from noise or artefacts. Furthermore, we propose a scheme to explicitly include tubularity features in the segmentation process for increased robustness and measurement repeatability. We demonstrate the performance of our approach by automatically segmenting the cervical spinal cord in magnetic resonance images, by reconstructing its surface, and acquiring volume measurements.

S. Pezold (✉) · K. Fundana · P.C. Cattin
Department of Biomedical Engineering, University of Basel, Basel, Switzerland
e-mail: simon.pezold@unibas.ch

K. Fundana
e-mail: ketut.fundana@unibas.ch

P.C. Cattin
e-mail: philippe.cattin@unibas.ch

M. Amann · M. Andelova · A. Pfister · T. Sprenger
University Hospital Basel, Basel, Switzerland
e-mail: michael.amann@usb.ch

M. Andelova
e-mail: michaela.andelova@usb.ch

A. Pfister
e-mail: armanda.pfister@usb.ch

T. Sprenger
e-mail: till.sprenger@usb.ch

© Springer International Publishing Switzerland 2015

J. Yao et al. (eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, Lecture Notes in Computational Vision and Biomechanics 20, DOI 10.1007/978-3-319-14148-0_10

1 Introduction

The segmentation of oriented tubular structures in the body is a common task in medical applications. Examples include measuring functional vessel volumes in patients of cardiovascular diseases, or quantifying spinal cord atrophy (i.e., the loss of nervous tissue) in a variety of neurodegenerative diseases. Multiple sclerosis (MS) is a prominent example among the latter diseases. Clinical MS studies have shown relationships between the degree of cord atrophy and both the strength of disease [1] and disease duration [2]. Therefore, in recent years, assessing spinal cord atrophy has become a highly active topic of research, resulting in a number of methods that were specifically tailored towards the segmentation of the spinal cord (see e.g. the recently published segmentation approaches of Asman et al. [3], De Leener et al. [4] and the methods referenced therein, or the earlier review of Miller et al. [5]). Only few of these methods, however, make extensive use of the fact that the spinal cord is an inherently tubular structure.

In this paper, we present an automated method that aims at the more general goal of segmenting tubular structures in image volumes. Manual intervention on the target data is reduced to placing a landmark if the segmentation result is ambiguous. As a proof of concept, we successfully demonstrate the practicability of our method by segmenting the spinal cord in magnetic resonance (MR) images and acquiring volume measurements from surface reconstructions of the segmentation results.

We adjust Yuan et al.'s continuous max-flow framework [6] to include a cross-sectional similarity prior. This prior exploits the fact that an oriented elongated structure shows only little change in shape along its orientation. Thus, the prior may guide the segmentation in regions where image information is missing or ambiguous. A related approach of including a similarity prior is pursued by Qiu et al. [7]. Due to their different problem setting (they aim for axial symmetry), they formulate parts of the problem in a discrete setting, while our formulation is continuous. We also propose a way to include tubularity features in the segmentation process. Specifically for the segmentation of the spinal cord, we furthermore introduce the new *csfness* feature, which is designed to improve discrimination between the spinal cord and the cerebrospinal fluid (CSF) that immediately surrounds it.

2 Method

In the following subsections, we introduce our adaptation of the max-flow approach and define the flow capacity functions together with the features that we use in experiments. We present an algorithm to solve the adapted problem, and we conclude the section by proposing a scheme to reconstruct the surface from the segmentation result, which we use for quantitative measurements.

Notation. Let $I : \Omega \to \mathcal{I}$ denote the intensity non uniformity corrected image [8] with intensities in the normalized intensity space $\mathcal{I} = [0, 1]$, where

$x = (x_1, x_2, x_3)^\mathsf{T} \in \Omega$ are the coordinates in the continuous image domain $\Omega \subset \mathbb{R}^3$. Throughout the whole section, we furthermore assume that the tubular structure of interest is roughly oriented parallel to the x_3 axis. Figuring out the orientation should be straightforward for most clinical applications, as the subject's orientation with respect to the image can be determined from the image's meta data for most clinical imaging modalities.

Original max-flow formulation. A general formulation for the continuous max-flow problem with spatial flow $p(x)$, source flow $p_s(x)$, sink flow $p_t(x)$, and corresponding flow capacities $C(x)$, $C_s(x)$, $C_t(x)$ is stated by Yuan et al. [6] as

$$\max_{p_s, p_t, p} \int_\Omega p_s \mathrm{d}x, \tag{1}$$

subject to the flow capacity constraints

$$p_s(x) \leq C_s(x), \quad p_t(x) \leq C_t(x), \quad \|p(x)\| \leq C(x) \tag{2}$$

and the flow conservation constraint

$$\operatorname{div} p(x) - p_s(x) + p_t(x) = 0. \tag{3}$$

2.1 Cross-Sectional Similarity Prior

Following our goal to impose a cross-sectional similarity prior on the segmentation, we split the spatial flow $p(x)$ into an in-slice component $q : \Omega \to \mathbb{R}^2$ and a through-slice component $r : \Omega \to \mathbb{R}$ with respect to slices that lie perpendicular to the x_3 axis (see Fig. 1a). The resulting continuous max-flow problem can then be written as follows:

$$\max_{p_s, p_t, q, r} \int_\Omega p_s \mathrm{d}x, \tag{4}$$

subject to the new flow capacity constraints

$$p_s(x) \leq C_s(x), \quad p_t(x) \leq C_t(x), \quad \|q(x)\| \leq \alpha(x), \quad |r(x)| \leq \beta(x) \tag{5}$$

and the new flow conservation constraint

$$\operatorname{div}_{12} q(x) + r'(x) - p_s(x) + p_t(x) = 0, \tag{6}$$

where $\operatorname{div}_{12} q$ denotes the divergence of q perpendicular to the x_3 axis and r' denotes the derivative of r along the x_3 axis.

The flow formulation now possesses the desired property of having the spatial flow capacity $C(x)$ of [6] represented by two separate terms, namely the in-slice flow capacity $\alpha(x)$ and the through-slice flow capacity $\beta(x)$. The latter capacity,

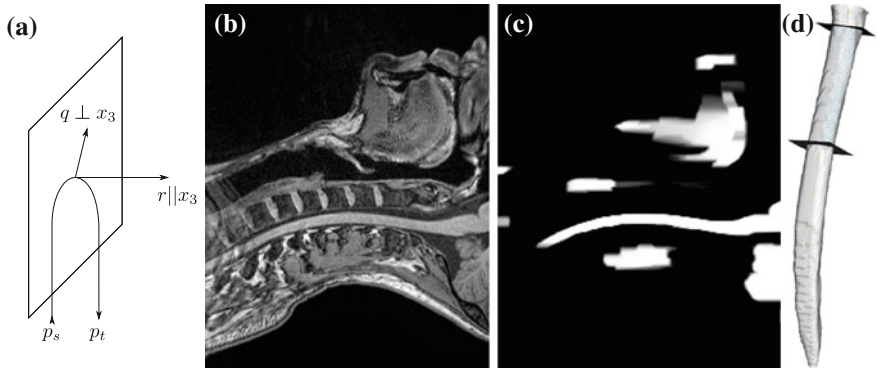

Fig. 1 Method overview. **a** Proposed flow configuration: the spatial flow is split into an in-slice component q, perpendicular to the axis along which the tubular structure is oriented, and a through-slice component r, parallel to the axis. **b** Sample sagittal slice of one of the images used for evaluation. **c** Segmentation result. **d** Surface reconstruction with cutting planes for volume measurement

$\beta(x)$, represents the cross-sectional similarity prior that allows for precise control over the through-slice flow behavior: For example, we may choose an edge-based cost function for $\alpha(x)$ that drives the segmentation towards edges in I, while setting $\beta(x) = \beta_0$ to enforce constant similarity throughout all slices. Or we may calculate $\beta(x) = \beta(x_3)$ as a slice wise cost-function that, for each slice, adjusts the similarity prior to the in-slice noise level (reinforcing the similarity prior if the noise level is high and relaxing it if the noise level is low). Other combinations are possible, of course: note that both α and β may be formulated pointwise.

Dual formulation. Introducing the Lagrange multiplier $u = u(x)$ and following the steps in [6], the max-flow problem can be reformulated as the equivalent primal-dual model

$$\max_{p_s, p_t, q, r} \min_{u} \int_{\Omega} p_s \mathrm{d}x + \int_{\Omega} u \cdot (\mathrm{div}_{12}\, q + r' - p_s + p_t)\mathrm{d}x \tag{7}$$

subject to the capacity constraints (5). The equivalent dual model representing a relaxed min-cut problem then becomes

$$\min_{u \in [0,1]} E(u) := \int_{\Omega} \left\{(1-u)C_s + uC_t + \alpha|\nabla_{12}\, u| + \beta|u'|\right\} \mathrm{d}x. \tag{8}$$

Here, $\nabla_{12}\, u$ denotes the in-slice gradient and u' denotes the through-slice derivative of u with respect to the x_3 axis, similar to the definitions of $\mathrm{div}_{12}\, q$ and r' above. It can be shown that each level set function $u^\ell(x),\ \ell \in (0, 1]$ given by

$$u^\ell(x) := \begin{cases} 1, & u^*(x) > \ell \\ 0, & u^*(x) \leq \ell \end{cases} \quad \text{with} \quad u^* := \operatorname*{argmin}_u E(u) \tag{9}$$

is a global binary solution of the adapted problem stated in Eq. (4).

2.2 Tubularity Features

As our goal is to segment tubular structures in the image, it appears natural to include tubularity features in the flow capacity calculations. A well-known tubularity feature is Frangi's measure of *vesselness* [9] (see Fig. 2b), $v^*(x) = \max_{\xi \in S_v} v(x;\xi)$, where, for each scale ξ in the predefined set of scales S_v, the vesselness $v(x;\xi)$ of bright tubular structures on dark background is

$$v(x;\xi) = \begin{cases} 0, & \lambda_2 \geq 0 \vee \lambda_3 \geq 0 \\ \left(1 - \exp(-2\frac{\lambda_2^2}{\lambda_3^2})\right) \exp(-2\frac{\lambda_1^2}{\lambda_2 \lambda_3}) \left(1 - \exp(-\frac{\sum_{i=1}^3 \lambda_i^2}{2h^2})\right) & \text{else,} \end{cases} \tag{10}$$

with $\lambda_i = \lambda_i(x)$ denoting the ordered eigenvalues ($|\lambda_1| \leq |\lambda_2| \leq |\lambda_3|$) of the pointwise Hessian matrices that result from convolving the input image I with Gaussian derivatives of standard deviation ξ. We define h as half of the maximum Hessian norm at the current scale as suggested by Frangi [9].

In our experiments on segmenting the spinal cord, we decided to include another feature that specifically describes the background that immediately surrounds the target structure. The spinal cord is embedded in cerebrospinal fluid (CSF), which appears dark in the used MR sequences. As the CSF also appears largely elongated, but exhibits both tube-like and plate-like properties, we adapt Frangi's vesselness feature to a *csfness* feature $w^*(x)$ (see Fig. 2c) that discriminates between blob-like structures and non-blobs. We do so by replacing the eigenvalue ratio terms of v^* with an equivalent term composed of λ_1 and λ_3, as it is the latter ratio that discriminates both vessels and plates from blobs in Hessian eigenvalue analysis [9]. Consequently, we define $w^*(x) = \max_{\xi \in S_w} w(x;\xi)$ for dark non-blobs on bright background in the scales S_w with

$$w(x;\xi) = \begin{cases} 0, & \lambda_3 \leq 0 \\ \exp(-2\frac{\lambda_1^2}{\lambda_3^2}) \left(1 - \exp(-\frac{\sum_{i=1}^3 \lambda_i^2}{2h^2})\right) & \text{else.} \end{cases} \tag{11}$$

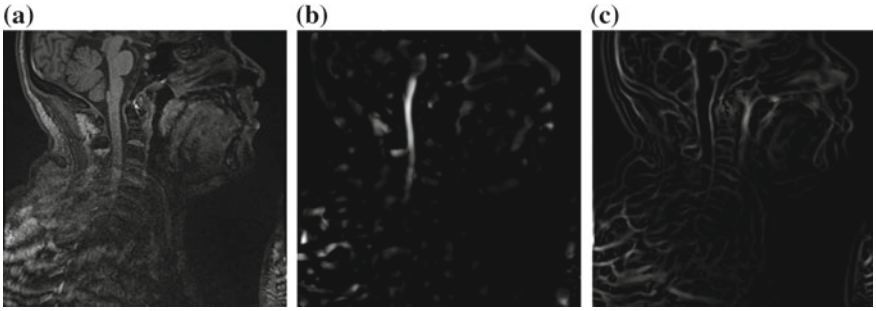

Fig. 2 Features used in segmentation. **a** Image intensities. **b** Vesselness response. **c** Csfness response

Combining the features. Let $\mathcal{V} = [0, 1] \ni v^*$, $\mathcal{W} = [0, 1] \ni w^*$ be the vesselness and csfness feature spaces, let $\mathcal{Y} = \mathcal{I} \times \mathcal{V} \subset \mathbb{R}^2$ and $\mathcal{Z} = \mathcal{I} \times \mathcal{V} \times \mathcal{W} \subset \mathbb{R}^3$ be two combined feature spaces, let $I_2 : \Omega \to \mathcal{Y}$, $I_3 : \Omega \to \mathcal{Z}$ be two new image functions that map to the combined feature spaces, and let $y \in \mathcal{Y}$, $z \in \mathcal{Z}$ be the coordinates in the combined feature spaces.

Furthermore, let $Y_f = \{y_f^i\}_{i=1}^M$, $Y_b = \{y_b^j\}_{j=1}^N$ be two sets holding samples of $\mathcal{Y}$ with known foreground and background membership, respectively. Based on these training sets, we propose to calculate the capacities for the terminal flow constraints (5) using kernel density estimates:

$$C_s(y) = C_s(I_2(x)) = \frac{\frac{1}{M}\sum_{i=1}^{M} K_{\Sigma_f}(y - y_f^i)}{\frac{1}{M}\sum_{i=1}^{M} K_{\Sigma_f}(y - y_f^i) + \frac{1}{N}\sum_{j=1}^{N} K_{\Sigma_b}(y - y_b^j)}, \tag{12}$$

$$C_t(y) = C_t(I_2(x)) = 1 - C_s(y), \tag{13}$$

where $K_{\Sigma_\cdot}$ is a Gaussian kernel with zero mean and diagonal covariance matrix $\Sigma_\cdot$, holding variances σ_d^2 for the feature dimensions d as diagonal elements. Terminal capacities for the feature space $\mathcal{Z}$ may be calculated in a similar way. For the sake of simplicity, we choose the non-terminal capacities as constants in our experiments: $\alpha(x) = \alpha_0$, $\beta(x) = \beta_0$.

2.3 Algorithm

In accordance with the original max-flow approach, we propose to find a global solution to our adapted formulation by setting up the respective augmented Lagrangian equation as

$$\begin{aligned} L_c(p_s, p_t, q, r, u) := &\int_\Omega p_s \mathrm{d}x + \int_\Omega u \cdot (\mathrm{div}_{12}\, q + r' - p_s + p_t)\mathrm{d}x \\ &- \frac{c}{2}\left\|\mathrm{div}_{12}\, q + r' - p_s + p_t\right\|^2, \end{aligned} \tag{14}$$

and iteratively optimizing it using Algorithm 1, based on the algorithm in [6].

2.4 Surface Reconstruction

As can be concluded from Eq. (9), reconstructing the surface of the segmented structure amounts to finding the isosurface of level $\ell \in (0, 1]$ in the segmentation result u^* (see Fig. 1c, d). We propose to extract the isoline as a polygon of m vertices for each slice along the x_3 axis and successively connect the resulting dots in space.

Algorithm 1 Augmented Lagrangian based max-flow algorithm.

Arbitrarily initialize p_s^1, p_t^1, q^1, r^1, u^1; initialize C_s, C_t, α, β; specify a tolerance $\hat{\varepsilon}$ and a step size c; set the iteration count $k = 1$. For each iteration, perform the following:

- Optimize q, fixing the other variables:

$$q^{k+1} \leftarrow \underset{\|q\| \le \alpha}{\operatorname{argmax}} L_c = \underset{\|q\| \le \alpha}{\operatorname{argmax}} -\frac{c}{2} \left\| \operatorname{div}_{12} q + r'^k - p_s^k + p_t^k - \frac{u^k}{c} \right\|^2,$$

  using a projected gradient ascent step of step size γ_q, as suggested in [10]:

  - update q: $q^{k+1} \leftarrow q^k + \gamma_q \cdot \nabla_{12}(\operatorname{div}_{12} q^k + r'^k - p_s^k + p_t^k - \frac{u^k}{c})$,
  - project q: $q^{k+1} \leftarrow \begin{cases} \frac{q^{k+1}}{\|q^{k+1}\|} \cdot \min\{\|q^{k+1}\|, \alpha\}, & q^{k+1} \neq 0 \\ 0, & q^{k+1} = 0 \end{cases}$.

- Optimize r, fixing the other variables:

$$r^{k+1} \leftarrow \underset{|r| \le \beta}{\operatorname{argmax}} L_c = \underset{|r| \le \beta}{\operatorname{argmax}} -\frac{c}{2} \left\| \operatorname{div}_{12} q^{k+1} + r' - p_s^k + p_t^k - \frac{u^k}{c} \right\|^2,$$

  using a projected gradient ascent step of step size γ_r:

  - update r: $r^{k+1} \leftarrow r^k + \gamma_r \cdot \frac{\partial}{\partial x_3}(\operatorname{div}_{12} q^{k+1} + r'^k - p_s^k + p_t^k - \frac{u^k}{c})$,
  - project r: $r^{k+1} \leftarrow \operatorname{sgn}(r^{k+1}) \cdot \min\{|r^{k+1}|, \beta\}$.

- Optimize p_s and p_t pointwise:

  - i) $p_s^{k+1} \leftarrow \frac{1}{c} - \frac{u^k}{c} + \operatorname{div}_{12} q^{k+1} + r'^{k+1} + p_t^k$, ii) $p_s^{k+1} \leftarrow \min\{p_s^{k+1}, C_s\}$,
  - iii) $p_t^{k+1} \leftarrow \frac{u^k}{c} - \operatorname{div}_{12} q^{k+1} - r'^{k+1} + p_s^{k+1}$, iv) $p_t^{k+1} \leftarrow \min\{p_t^{k+1}, C_t\}$.

- Calculate the pointwise error ε: $\varepsilon^{k+1} \leftarrow c \cdot (\operatorname{div}_{12} q^{k+1} + r'^{k+1} - p_s^{k+1} + p_t^{k+1})$.
- Update u: $u^{k+1} \leftarrow u^k - \varepsilon^{k+1}$.
- Terminate if $\frac{1}{|\Omega|} \int_\Omega |\varepsilon^{k+1}(x)| \, dx < \hat{\varepsilon}$, otherwise update $k \leftarrow k + 1$ and continue.

This provides us with the slicewise contours of the segmentation at no additional cost, which then facilitates estimating the centerline, namely as a curve fit through the centroids of the contours. A centerline estimate, in turn, may be useful to acquire quantitative measurements from the reconstruction (see Sects. 3, 4).

If there are multiple foreground regions in u^*, a point of reference may be used to choose the region closest to it. Likewise, heuristic criteria like sudden jumps of the centroid or a threshold on the contour line's convexity may be used to determine a cutoff for the tubular structure of interest. In the spinal cord segmentation experiments below, we define a point of reference by an anatomical landmark, and we define two cutoff criteria as finding either a distance $> d$ between the centroids of two consecutive slices or finding a contour line with convexity $< t$. As a measure of convexity, we employ the ratio of the contour line's area and the area of its convex hull.

3 Materials

Applicability of our approach is shown by segmenting the spinal cord in MR images of healthy volunteers (Figs. 1b, 2a) and MS patients.

To assess accuracy and reproducibility, 11 healthy volunteers (3 female, 8 male, mean age 32.7 year, range 26–44 year) were scanned on a 3 T whole-body MR scanner (Verio, Siemens Medical, Germany) with a T1-weighted MPRAGE sequence (TR/TI/TE/α = 2.0 s/1.0 s/3.4 ms/8°); 192 slices in sagittal orientation parallel to the interhemispheric fissure were acquired with an isotropic resolution of 1 mm^3. Image volumes were corrected for gradient nonlinearity distortions using the scanner manufacturer's correction routine.

To show applicability to clinical data, we used follow-up data of 32 MS patients (21 female, 11 male, mean age 47.1 year, range 22–60 year; 22 patients with relapsing-remitting MS, 10 patients with primary progressive MS, mean disease duration 13.8 year, range 3–31 year, median EDSS 3.0, range 1.5–6.0). The patients were scanned on a 1.5 T whole-body MR scanner (Avanto, Siemens Medical, Germany) with a T1-weighted MPRAGE sequence (TR/TI/TE/α = 2.08 s/1.1 s/3.93 ms/15°); 160 slices in sagittal orientation parallel to the interhemispheric fissure were acquired with an in-slice resolution of 0.98 mm × 0.98 mm and a slice thickness of 1 mm. Scans were acquired at two points in time approximately 5 years apart (mean 5.04 year, range 4.55–5.41 year); demographic data above is given with respect to the earlier scan. Distortion correction was applied to the surface reconstructions using the method of Janke et al. [11].

For the calculation of the terminal flow constraints $C_\cdot(y)$ and $C_\cdot(z)$, sample sets were acquired on 150 separate scans of MS patients. The training patients were scanned with the same MPRAGE sequence as the 32 MS patients above. Foreground/background membership of the training samples was determined using a graph cuts-based [12] semi automated method described as *presegmentation* in [13]. To speed up calculations, features were discretized to 50 bins in the $[0, 1]$ interval in each feature dimension. Silverman's rule of thumb with $\sigma_d = 4^{\frac{1}{D+4}} (n(D+2))^{\frac{-1}{D+4}} \hat{\sigma}_d$ provided a σ_d^2 estimate, where n is the number of samples, $\hat{\sigma}_d$ is the sample standard deviation in d, and D is 2 for $C_\cdot(y)$ and 3 for $C_\cdot(z)$. To avoid zero bins, a small additive constant of $0.0001k$ was added to the resulting bin values, where k is the maximum value of all bins.

For all experiments, the following parameters were applied: $\alpha_0 = 0.5$, $\beta_0 = 2.5$, $S_v = [2\,\text{mm}, 4\,\text{mm}]$ (16 values), $S_w = [1\,\text{mm}, 2\,\text{mm}]$ (8 values) for the flow capacities; $\ell = 0.5$, $m = 60$, $d = 10\,\text{mm}$, $t = 0.95$ for the surface reconstruction. Likewise for all reported volume measurements, the volume of a spinal cord surface segment of 50 mm centerline length, which was clipped by planes perpendicular to the centerline and which was located approximately 25 mm inferior of a manually marked landmark, was evaluated as described in [13].

4 Results

Scan–rescan evaluation. In two experiments on scan–rescan data, we evaluated the accuracy and reproducibility of our method. To show the benefits of splitting the spatial flow into an in-slice component and a through-slice component, we repeated all experiments using Yuan et al.'s original formulation [6] for segmentation, setting the spatial flow capacity $C(x)$ (see Eq. (2)) to a value of $C(x) = 1.0$, as this parameter choice provided the highest number of successful surface reconstructions in the second experiment below. For the experiments, we scanned the 11 healthy volunteers (see Sect. 3) three times in a row (scans S1, S2, S3), without repositioning between S1 and S2 and with repositioning between S2 and S3, resulting in 33 scans altogether.

Accuracy: As we work with human *in-vivo* data, it was not possible to acquire quantitative ground truth measurements, for example, via histologic specimen. We therefore used manual segmentations of the image data as a gold standard for comparison in the first experiment. To make such manual measurements feasible, a semiautomated approach that allows for human feedback in the segmentation process seemed appropriate. We thus segmented all scan–rescan datasets with the method described as *presegmentation* in [13], placing foreground/background seeds manually and adjusting them in an iterative manner until we acquired a satisfying binary segmentation. We then compared the overlap of this gold standard segmentation with the binarized results of the automated segmentation for a 50 mm cord segment, located 25 mm inferior of the manually marked landmark. As a measure of overlap agreement, we calculated Dice coefficients for the region overlaps.

With our approach, we gained a mean Dice coefficient of 0.88 using the proposed feature combination $\mathcal{Z}$ (i.e., intensity + vesselness + csfness) and 0.82 using feature combination $\mathcal{Y}$ (i.e., intensity + vesselness). With Yuan et al.'s approach, we gained a mean Dice coefficient of 0.86 using $\mathcal{Z}$ and 0.79 using $\mathcal{Y}$. Therefore, our approach proves superior in the given problem setting. Furthermore, it can be seen that including the csfness feature into the segmentation process improves the segmentation accuracy.

Reproducibility: In the second experiment, we assessed the reproducibility of our method. The cervical spinal cord was segmented using feature combinations $\mathcal{Y}$ and $\mathcal{Z}$, its surface was reconstructed, and the volume of the 50 mm cord surface segment (see Sect. 3) was compared between scans and rescans. As a measure of reproducibility, we calculated the coefficients of variation (CV; i.e., the sample standard deviation over the mean) of the measured volumes for all possible S1–S2 comparisons (i.e., without repositioning) and S1–S3 comparisons (i.e., with repositioning). An overview of the mean CVs is given in Table 1.

For our proposed segmentation approach, the subsequent reconstruction of the complete surface segment succeeded for 30 out of 33 scans using $\mathcal{Y}$ and 32 out of 33 scans using $\mathcal{Z}$. All failures happened for the same subject, whose scans showed an extremely low signal-to-noise ratio upon visual inspection. For Yuan et al.'s segmentation approach, the surface reconstruction succeeded for 25 out of 33 scans using $\mathcal{Y}$ and 28 out of 33 scans using $\mathcal{Z}$.

Table 1 Coefficients of variation (%) for spinal cord segment volume measurements, using feature combinations $\mathcal{Y}$ and $\mathcal{Z}$ with Yuan et al.'s and our segmentation method

	$\mathcal{Y}$		$\mathcal{Z}$	
	S1 versus S2	S1 versus S3	S1 versus S2	S1 versus S3
Ours	**2.02**	**5.81**	**2.13**	4.90
Yuan et al.	2.54	6.84	2.85	**4.42**

As one could expect, CVs are lower for the S1–S2 comparison, due to the fact that the subjects were not repositioned. Furthermore, including the csfness feature makes the segmentation more robust (more successful surface reconstructions) while at the same time having beneficial effects on the reproducibility (substantially lower CVs for the more realistic S1–S3 comparisons). Similar statements on improved robustness and reproducibility can be made when comparing our adapted max-flow formulation with the original formulation: in both aspects, our method proves largely superior. And while for feature combination $\mathcal{Z}$ the S1–S3 CV of the original approach is better than ours, one has to keep in mind that ours is calculated on a higher number of successful reconstructions, including the more challenging ones on which the original approach failed.

An exemplary case where the surface reconstruction failed for the original max-flow formulation while succeeding for our adapted formulation is shown in Fig. 3. As can be seen, the segmentation stops early for the original formulation while it extends further down into the noisy image regions for ours. Relaxing $C(x)$ in this case would possibly enable the original formulation to also extend further down; however, this would come at the price of an overall higher susceptibility to noise. By contrast, controlling $\alpha(x)$ and $\beta(x)$ separately in our approach enables us to largely circumvent this tradeoff.

On the whole, the CVs we obtained by our method are higher than those of established methods that are actually used in MS research (most notably, Losseff et al. [1] and the methods compared in [13]). On the one hand, however, one should keep

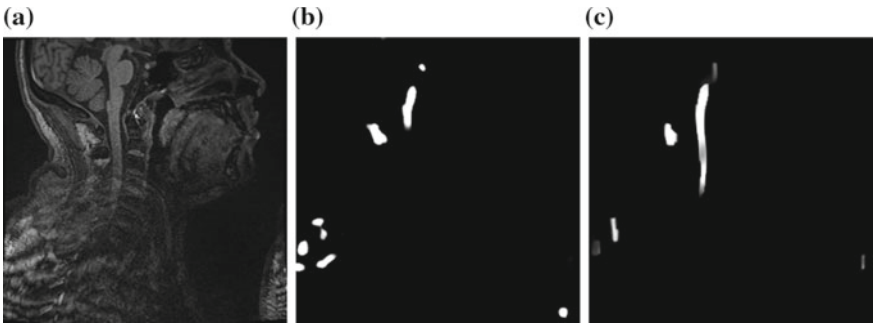

Fig. 3 Comparison of the segmentation approaches on noisy, low-contrast case. **a** Input image. **b** Segmentation result using the original max-flow formulation [6]. **c** Segmentation result using our max-flow formulation

in mind the substantially higher amount of manual intervention in these approaches. On the other hand, we see our presented framework in its current state more as a proof of concept than as a tool that is ready for clinical use.

Evaluation on patient data. A comparison of the five-year follow-up MS patient data (see Sect. 3), using the proposed max-flow formulation for segmentation, showed a mean yearly atrophy of 25.4 mm^3 in the 50 mm cord surface segment (maximum loss: 194.3 mm^3, maximum gain: 53.4 mm^3). The mean yearly percentage loss was 0.9 % (maximum loss: 7.0 %, maximum gain: 2.0 %). These measurements agree well with the observation of cord atrophy during MS progression reported in the literature [2]. Nevertheless, due to the high variability, our measurements should again be interpreted as a proof of concept for our segmentation method rather than as hard clinical data.

Computational performance. As we implemented the max-flow segmentation on the GPU based on code provided by the authors of [6, 10], results can be acquired extremely fast, namely in the order of seconds. Other parts of the implementation also show a high parallelization potential in that they are mainly pointwise (such as the feature calculation and the surface extraction). We therefore assume that the complete chain of steps from feature calculation to quantitative measurements could be optimized to run in less than a minute per subject.

5 Conclusion

We presented a new segmentation algorithm based on continuous max flow that was specifically tailored towards the segmentation of elongated structures: a cross-sectional similarity prior was introduced, which guides the segmentation in regions of missing or contradictory image information. We showed how tubularity features may be used in the flow capacity constraints to increase segmentation robustness and measurement repeatability. Finally, we successfully demonstrated the clinical applicability of our method by segmenting the spinal cord in both healthy volunteers and multiple sclerosis patients.

Acknowledgments We would like to thank Ernst-Wilhelm Radue and the MIAC AG, University Hospital Basel, Basel, Switzerland, for their support.

References

1. Losseff, N.A., Webb, S.L., O'Riordan, J.I., Page, R., Wang, L., Barker, G.J., Tofts, P.S., McDonald, W.I., Miller, D.H., Thompson, A.J.: Spinal cord atrophy and disability in multiple sclerosis. Brain **119**(3), 701–708 (1996)
2. Rashid, W., Davies, G.R., Chard, D.T., Griffin, C.M., Altmann, D.R., Gordon, R., Thompson, A.J., Miller, D.H.: Increasing cord atrophy in early relapsing-remitting multiple sclerosis: a 3 year study. J. Neurol., Neurosurg. Psychiatry **77**(1), 51–55 (2006)

3. Asman, A., Smith, S., Reich, D., Landman, B.: Robust GM/WM segmentation of the spinal cord with iterative non-local statistical fusion. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) Medical Image Computing and Computer-Assisted Intervention—MICCAI 2013. Lecture Notes in Computer Science, vol. 8149, pp. 759–767. Springer, Heidelberg (2013)
4. De Leener, B., Kadoury, S., Cohen-Adad, J.: Robust, accurate and fast automatic segmentation of the spinal cord. NeuroImage **98**, 528–536 (2014)
5. Miller, D.H., Barkhof, F., Frank, J.A., Parker, G.J.M., Thompson, A.J.: Measurement of atrophy in multiple sclerosis: pathological basis, methodological aspects and clinical relevance. Brain **125**(8), 1676–1695 (2002)
6. Yuan, J., Bae, E., Tai, X.C.: A study on continuous max-flow and min-cut approaches. In: Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2217–2224 (2010)
7. Qiu, W., Yuan, J., Ukwatta, E., Sun, Y., Rajchl, M., Fenster, A.: Fast globally optimal segmentation of 3d prostate mri with axial symmetry prior. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) Medical Image Computing and Computer-Assisted Intervention—MICCAI 2013. Lecture Notes in Computer Science, vol. 8150, pp. 198–205. Springer, Heidelberg (2013)
8. Tustison, N., Avants, B., Cook, P., Zheng, Y., Egan, A., Yushkevich, P., Gee, J.: N4ITK: improved n3 bias correction. IEEE TMI **29**(6), 1310–1320 (2010)
9. Frangi, A., Niessen, W., Vincken, K., Viergever, M.: Multiscale vessel enhancement filtering. In: Wells, W., Colchester, A., Delp, S. (eds.) Medical Image Computing and Computer-Assisted Interventation—MICCAI'98. Lecture Notes in Computer Science, vol. 1496, pp. 130–137. Springer, Heidelberg (1998)
10. Yuan, J., Bae, E., Tai, X.C., Boykov, Y.: A study on continuous max-flow and min-cut approaches. Tech. Rep. CAM 10–61, UCLA, CAM, UCLA (2010)
11. Janke, A., Zhao, H., Cowin, G.J., Galloway, G.J., Doddrell, D.M.: Use of spherical harmonic deconvolution methods to compensate for nonlinear gradient effects on MRI images. Magn. Reson. Med. **52**(1), 115–122 (2004)
12. Boykov, Y.Y., Jolly, M.P.: Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. In: Eighth IEEE International Conference on Computer Vision, 2001. ICCV 2001. Proceedings, vol. 1, pp. 105–112 (2001)
13. Pezold, S., Amann, M., Weier, K., Fundana, K., Radue, E., Sprenger, T., Cattin, P.: A semi-automatic method for the quantification of spinal cord atrophy. In: Yao, J., Klinder, T., Li, S. (eds.) Computational Methods and Clinical Applications for Spine Imaging, Lecture Notes in Computational Vision and Biomechanics, vol. 17, pp. 143–155. Springer International Publishing (2014)

Automated Radiological Grading of Spinal MRI

Meelis Lootus, Timor Kadir and Andrew Zisserman

Abstract This paper describes a fully automatic system for obtaining the standard Pfirrmann degeneration grading of individual intervertebral spinal discs in T2 MRI scans. It involves detecting and labeling all the vertebrae in the scan and then learning a regression from the disc region to the grading. In developing the regression function we investigate a spectrum of support regions which involve differing degrees of segmentation of the scan: our intention is to ascertain to what extent segmentation is necessary or detrimental in obtaining robust and accurate measurements. The methods are assessed on a heterogeneous clinical dataset containing 1,710 Pfirrmann-graded discs, from 285 symptomatic back pain patients. We are able to predict the grade to ± 1 precision at 85.8 % accuracy. Our novel method proposes new image features that outperform previous features and utilizes techniques to improve robustness to MR imaging variations.

Keywords Spine · Radiological measurement · MRI · Grading · Discs · Regression

1 Introduction

Our primary goal in this paper is to automate radiological measurements in multi-slice clinical Magnetic Resonance Imaging (MRI) spinal scans, and to this end we describe a system to extract the standard clinical Pfirrmann disc grading that is used in the diagnosis and management of back pain patients, exploring accuracy and robustness in the process. The task is defined in Fig. 1.

M. Lootus (✉) · A. Zisserman
Oxford University, Oxford, UK
e-mail: meelis@robots.ox.ac.uk; meelis.lootus@gmail.com

A. Zisserman
e-mail: az@robots.ox.ac.uk

T. Kadir
Mirada Medical, Oxford, UK
e-mail: timork@mirada-medical.com

© Springer International Publishing Switzerland 2015

J. Yao et al. (eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, Lecture Notes in Computational Vision and Biomechanics 20, DOI 10.1007/978-3-319-14148-0_11

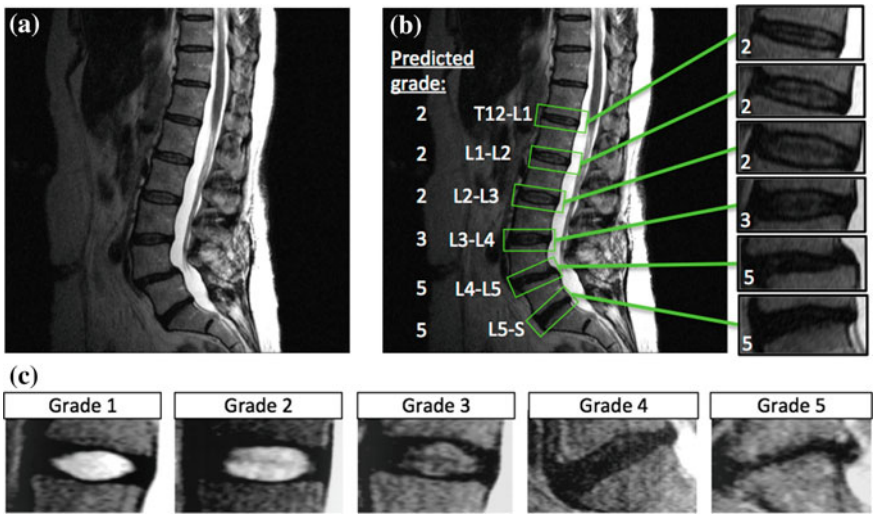

Fig. 1 The task. Given a clinical MRI volume of the lumbar spine (**a**) as input, fully automatically localize, label, and radiologically measure (**b**) the six lumbar discs in that volume, according to a standard radiological grading system (**c**) [15]. The radiological measurement is the 'degeneration grade' describing drying out of the disc (darkening in T2 MRI), and disc space collapse, in terms of four radiological features as defined in the main text. Note that in evaluation, the grade is considered correct if predicted to ± 1 accuracy, due to the ground truth intra-observer variability. **a** Input. **b** Output. **c** Radiological grading atlas

Our secondary goal is to investigate the need for a complete segmentation of the disc in order to accomplish this task. On the one hand, voxel wise segmentation can help better define the grading task. On the other hand, anatomical units (discs and vertebrae in this case) may be inseparable due to pathological changes, rendering the task ill-defined. Also, in practice, segmentation is often prone to failure so avoiding it can possibly improve the overall results. Thus, we ask the question: to segment or not to segment?

We answer the question by formulating the task as one of regressing between an image support region and the Pfirrmann disc grading, and then investigate a spectrum of ways of obtaining the support region which cover: no segmentation, segmentation of only the vertebrae, and finally segmentation of the disc.

The method is evaluated over a large heterogeneous clinical dataset, and this adds to the challenge of the task since T2 images of the same anatomy and pathology look different in different MRI machines and under different protocols (different "tissue contrasts"). We introduce a normalization scheme to address this problem. The task is also challenging because MR imaging artifacts can be confused with pathology.

Background. The normal intervertebral disc is composed of a soft liquid central part, the nucleus pulposus (NP), and a hard ligamentous surrounding, the annulus fibrosus. It is interfaced to the vertebral bodies above and below by cartilaginous endplates.

The disc acts as both a cushion and a pivot point in the spine. Disc problems are a common cause of back pain.

One common disc problem is *Degeneration*, the drying out and collapse of the disc space, and this abnormality is clinically measured using the standard radiological **Pfirrmann grade** [15], illustrated in Fig. 1. Pfirrmann defines the categorical five-score grade in terms of sequential changes to four MRI features: brightness of the NP, uniformity of the NP, distinction between the NP and the annulus fibrosus, and the disc height.

Since our ground-truth labelling is not perfect—the intra-observer grade agreement in our database is only 71 % based on grading 121 patients twice, while agreement to ± 1 is 98 %—we assess our scores to within ± 1 grade accuracy. Note, Pfirrmann [15] measured 88–92 % intra-observer agreement over measurements in a single day. Our database was annotated by one radiologist over several years—achieving similar variability to *inter*-observer variability witnessed by Pfirrmann.

While the Pffirrmann grade is widely used in clinical practice to assess the overall disc quality, conflicting accounts have been presented in research studies regarding the correlation between the grade, back pain, and surgical outcomes [5, 9, 12].

Given the quite high intra- and inter-observer variability of radiological measurements, one advantage of automating measurement is that it should lead to an improvement in consistency. In turn, this consistency may lead to improvements in both clinical research studies on the correlation of back pain with radiological measurements, and communication between radiologists.

1.1 Related Work

Recently, multiple medical imaging papers have been published attempting to automatically diagnose a number of spinal conditions [1, 7, 14, 16, 19].

The existing publications on Disc Degeneration deal with the binary classification task—e.g. the presence or absence of desiccation/degeneration [1, 7, 14, 19]—rather than measuring the standard radiological Pfirrmann grade or a similar radiological quantity. In addition, they are generally restricted to homogeneous data collected from the same scanner, using the same [1, 19], or a relatively narrow range [14] of protocols.

Image features for Pfirrmann grading have been proposed before [13], however their computation is not fully automatic and they have not been used to automatically measure the grade. Alomari et al. [1] and Neubert et al. [14] automatically predicted a binary grading to high accuracy in MRI scans over a homogeneous dataset; and we compare to their features here. Often, the methods require a segmentation step [14] to accurately delineate the discs, or to find an exact square in the disc [1]. The robustness of the methods to segmentation has not been explicitly studied, however it is an important point.

A number of vertebra and disc segmentation algorithms have been proposed [6, 8, 10, 18]. Vertebra segmentation methods have been more successful than disc

segmentation ones. This is largely because vertebrae in MRI have well-defined edges and consistent appearances across patients. In contrast, discs have variable appearance, lack clear boundaries, and vary considerably due to degradation (the very process we are assessing).

1.2 Contributions and Paper Layout

Contributions. We fully automatically measure the Pfirrmann grade to ± 1 accuracy, and are the first to present results generalizing across clinical data from a number of sites, scanner types and imaging sequences. We train and assess our results based on ground truth annotations on clinical data, by an expert radiologist with 25 years of experience. We experimentally assess the effect of varying the amount of segmentation to define the feature support region, and compare our features to those proposed in previous work on binary auto-grading [1, 14].
Paper layout. The method is presented in full in Sect. 2, explaining in detail the three steps of the pipeline: feature support region definition, image feature extraction, and regression. The dataset is described in Sect. 3.1, the evaluation protocol in Sect. 3.2, and the experimental results and discussion in Sect. 3.3.

2 Regression of Radiological Measures

In this section, we describe a method for predicting the Pfirrmann grading. It is formulated as a regression task, and we use standard machine learning methods to learn the regressor from expert-annotated training scans, and then apply the regressor to previously unseen clinical scans.

Although Pfirrmann grading is categorical, the underlying fundamental disc degeneration process is continuous over time, and that is why we choose to formulate it as regression rather than classification.

The pipeline from a raw multi-slice MRI scan to radiological measurement of disc degeneration has three steps, described in more detail in the following sections: first, finding the support region; second, extracting image features; third, predicting the Pfirrmann grade. In Sect. 2.1, we explore three alternative methods of obtaining the feature support region, named **V-det** (vertebra detection), **V-seg** (vertebra segmentation), and **D-seg** (disc segmentation), illustrated in Fig. 2. This explores three points on the 'no segmentation' to 'full disc segmentation' spectrum.

2.1 Step 1: Three Alternative Support Regions

Taking a clinical MRI scan as input, this step outputs the feature support region for the six lumbar discs (annotated in Fig. 1) in three different ways, as contrasted in

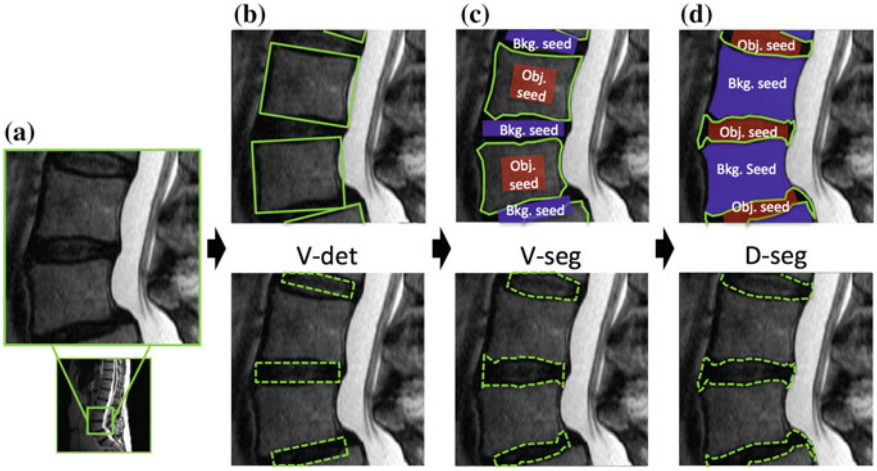

Fig. 2 Computation of the feature support regions. The *top row* illustrates the detection and segmentation steps; the *bottom row* shows the three corresponding feature support regions for the **b** V-det, **c** V-seg, and **d** D-seg methods. The full segmentation pipeline consists of vertebrae detection, vertebrae segmentation, and intervertebral disc segmentation. In the *top row*, the *green lines* show the detection and segmentation outputs, and the Obj. (object) and Bkg. (background) seeds show the initializations used for the segmentations as the *red* and *blue* areas. The resulting support regions are shown as *green dashed lines* in the *bottom row*. The V-det pipeline is the shortest, involving no segmentation, and the D-seg pipeline is the longest, involving both vertebrae and disc segmentation. See Sect. 2.1 for more detail. **a** Input image (zoom). **b** VB Detections. **c** VB Segmentations. **d** IVD Segmentations (Color figure online)

Fig. 2. In **V-det**, the region is defined as a rectangle between vertebrae bounding boxes. In **V-seg**, the region is defined as a rectangle based on vertebrae segmentations and excluding any vertebrae voxels. In **D-seg**, the region is defined as the disc segmentation result, the disc mask. The full algorithm from image to vertebrae and disc segmentation is sequential: (1) vertebrae detection, (2) vertebrae segmentation, (3) disc segmentation, with each step automatically initializing the next. Both the vertebrae and the discs are segmented using the standard graph cuts algorithm of Boykov-Jolly [3] with region and boundary terms. As might be expected, each step in the process has some degree of failure rate. So the more steps we employ, the greater the potential for failure.

Implementation details. The detections are performed in all slices using the method of Lootus et al. [11], picking—for each vertebra v_i—the tight labelled bounding box $B(v_i)$ from the slice with the most confident detector output. This way, the system is robust to scoliotic spines where not all lumbar vertebrae may be clearly visible in the same sagittal slice. Note that the labelling starts from the sacrum and is performed in the same manner and is consistent with the ground truth labelling process. Therefore any labelling errors due to the natural variability in vertebra number will not affect the results presented here, but will need to be considered in the future work. The segmentations and gradings are performed in the mid-sagittal slice (as

clinical standard), initializing the support region selection based on the automatically labelled bounding boxes. The segmentation of the vertebrae T12 (v_1) to S1 (v_7) is initialized automatically for each v_i independently, by placing object (Obj.) and background (Bkg.) seeds according to the bounding boxes $B(v_{i-1})$, $B(v_i)$, and $B(v_{i+1})$. The Obj. seed for v_i is placed in an area obtained by eroding $B(v_i)$ to half its size. The two Bkg. seeds are placed as rectangles between $B(v_{i-1})$ and $B(v_i)$, and $B(v_i)$ and $B(v_{i+1})$ respectively, at the arithmetic mean position of the corners of the adjacent bounding box edges. Each seed is a quarter the height of $B(v_i)$. The disc segmentation is performed for discs T12/L1 to L5/S1. For each disc, it is automatically initialized by seeding the foreground as the space between, and background as the area of the neighboring vertebra segmentations. The region terms are modeled as three-component Gaussian Mixture Models (GMMs) according to the image intensities in the Obj. and Bkg. seeds for both the vertebrae and disc segmentations. Three components were picked as best performing at earlier experiments. To arrive at the final box regions for V-det, the box is placed at the angle of its lower neighboring vertebra, with its height equal to a quarter of the mean heights, and its width equal to the mean width of the neighboring vertebrae bounding boxes (optimal size found by gird search). In V-seg, the vertical 'walls' of the region are placed at the same positions as for the V-det case.

2.2 Step 2: Image Features

To characterize the disc, a number of intensity and shape features are extracted from the feature support region, as described below. As mentioned in the introduction, one of the challenges is the variation of MRI contrast across different protocols. Therefore, before extraction, the image is normalized to 1.25 times the median vertebral intensity as found from the bounding boxes (for V-det) or the segmented vertebrae (for V-seg/D-seg). In early experiments, we tested a number of intensity normalization techniques including normalizing to CSF and found this approach performed best and provided sufficient robustness to variations in protocol and MRI scanners. The conventional CSF-based normalization [2] also suffers from inconsistent CSF signal due to various deformations in the dural sac. After normalization the new median vertebral intensity becomes 0.8; intensities above one are truncated to one. Thus, there is still dynamic range kept above the vertebrae intensity (e.g. for grade 1 or 2 discs), unlike in the case if the vertebra intensity was normalized to one.

Intensity features. Two groups of intensity features are extracted: first, a histogram normalized so that the highest entry is one; second, four global statistical features: standard deviation, kurtosis, skewness, and entropy.

Shape features. The mid-height to width ratio h/w of the feature support region, approximating that of the disc, is used as the shape measure.

Implementation details. The histogram feature is modelled with 21 bins, making up a 26-dimensional feature vector for each disc. The mid-height to width ratio h/w is measured in the middle of the feature support region—the rectangle for V-det, the disc space for V-seg, and the segmented disc for D-seg.

Baseline Features. In addition to the above features, as a baseline we assess the performance of our system using two previously proposed intensity features: Disc Mean Intensity [1] and a GMM fitted to the support region [14].

2.3 Step 3: Regression

A linear epsilon Support Vector Regressor [17] is learned to map the 26-dimensional feature vector **x** described above to the grading as:

$$f(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} + b \tag{1}$$

where f is the predicted (continuous) disc grade, and the vector **w** and bias b are learnt on the training set. The fitting process is influenced by two unitless parameters: $C > 0$, and $\varepsilon > 0$. The parameter C determines the trade-off between the flatness of f and the amount up to which deviations larger than ε are tolerated.
Implementation details. The values of ε and C for the Support Vector Regression cost function are learnt by a grid search on a hold-out validation set, with a range of 0.01 to 1.0, and 0.1 to 10,000 respectively, using the LIBSVM package [4].

3 Evaluation and Comparison

3.1 Dataset

We evaluate our algorithm using T2 sagittal MRI scans from a large, clinical, heterogeneous dataset of 1,710 radiologically annotated lumbar discs in 285 symptomatic back pain patients, exhibiting all the challenges highlighted in the introduction including degeneration, herniation and scoliosis. Of the patients, 116 were male, 151 female (30 unknown), with ages 10–88 years. The patients were split into a 114-scan training (684 discs), 57-scan validation (342 discs), and 114-scan (684 discs) testing set.

In contrast to previous work where images were acquired using the same scanner and protocol, our database includes scans from 25 different sites acquired using a wide gamut of T2 sequences, fields-of-view, and 14 different scanner models. The scan parameters ranged as follows: magnetic field 0.6–3T, TE 69–139 ms, TR 1,180–1,210 ms, voxel size 0.34–1.68 mm, slice spacing 3.85–15 mm.

Each patient had each of their six lumbar discs (T12-L1, L1-L2, L2-L3, L3-L4, L4-L5, and L5-S) Pfirrmann-graded by an expert spine radiologist with 25 years of clinical experience. There were 550 grade 1, 194 grade 2, 379 grade 3, 379 grade 4, and 208 grade 5 discs in the dataset.

3.2 Evaluation Protocol

We assess vertebrae detection accuracy by distance between the detected bounding box center from ground truth bounding box centre in millimeters, and segmentation accuracy according to an overlap measure—the ratio of intersection to union between manually segmented vertebrae/discs, and the automatic segmentation results.

We assess grading accuracy by measuring the proportion of discs that are graded correctly, as proposed by [20] for eye cataract grading. However, as discussed in the introduction, our ground-truth labelling is not perfect due to intra-observer variability, and thus we evaluate our regression performance as the fraction of scores which are predicted to ± 1 accuracy.

3.3 Results and Discussion

A summary of the overall results and the D-seg confusion matrix is presented in Fig. 3, with example grading result shown for two patients in Fig. 4, and for two more in Fig. 5. The median detection error was 2.0 mm; the mean vertebrae segmentation overlap measure was 0.808 ± 0.132 on fifty randomly selected patients. Overall, the best performing method is D-seg (85.8 %) using the Hist+ features. For all three support region methods, the Hist+ features outperforms the baseline GMM and the Mean Int. features. Also, the segmentation methods (D-seg and V-seg) outperform detection only (V-det). For both Mean Int. and GMM, V-seg outperforms D-seg by a small margin but performs similarly with Hist+ features. The D-seg with Hist+ features confusion matrix shows that the greatest errors are in predicting a grade of '3'.

Based on qualitative analysis on all discs from levels 1–5 (1,425 in total), there were 50 discs with detection/segmentation failure (3.5 %), 21 discs (1.5 %) with low imaging quality, 19 discs (1.3 %) failed slice selections. This sums to 6.3 %, to approximately half of all failure cases in the test set, and covers the principal causes of error.

The Hist+ features provide a clear improvement. This may be because the Mean Int. is insufficiently discriminative, while the GMM parameters might vary significantly between discs (since GMM fitting minimizes error to the underlying distribution, but does not constrain the component centers). The Hist+ features provide improved discriminative power, are repeatable, and also include the global

Fig. 3 Numerical results. Percentage of discs with Pfirrmann grade predicted to ± 1 accuracy using each of the features (*columns*) on each of the support regions (*rows*)

	Mean Int.	**GMM**	**Hist+**
V-det	73.4%	77.5%	79.7%
V-seg	81.1%	84.6%	85.2%
D-seg	79.6%	83.3%	**85.8%**

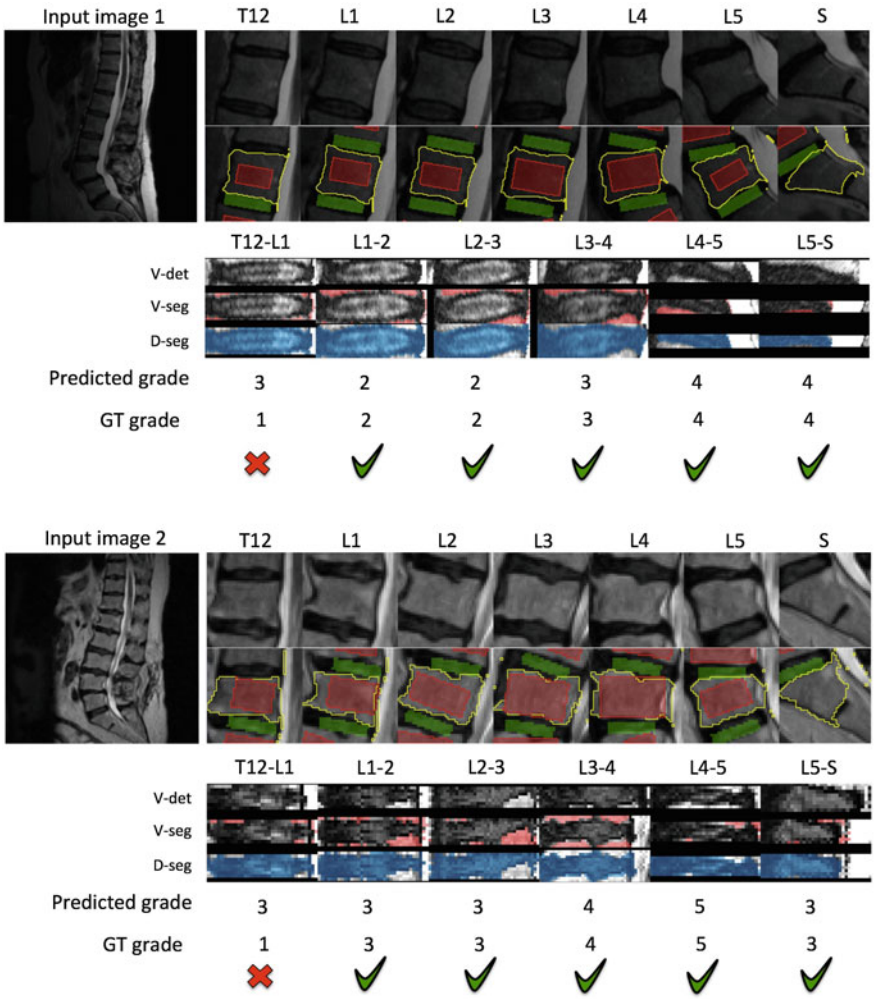

Fig. 4 Example results 1–2. The full segmentation and grading results are given for two example input images. For each patient, in the *second row* from the *top*, the vertebrae segmentations are shown as *yellow lines*. In the *bottom* three rows, the extracted regions for feature extraction, along with the segmented areas for the vertebrae and the discs are given in the three *bottom rows* for V-det, V-seg, and D-seg methods. At the *bottom*, the grading result is given for the D-seg method. For both patients 1 and 2, five disc gradings succeed, and one fails. Note that the fact that both the discs T12-L1 are predicted two grades too high is not a systematic error, but randomly present in those two cases (Color figure online)

descriptors of the distribution shape (that the GMM can capture). In early experiments, we found the global descriptors to improve performance.

We investigated the role of segmentation in our processing pipeline by testing three different approaches to defining the feature support region—vertebra detection

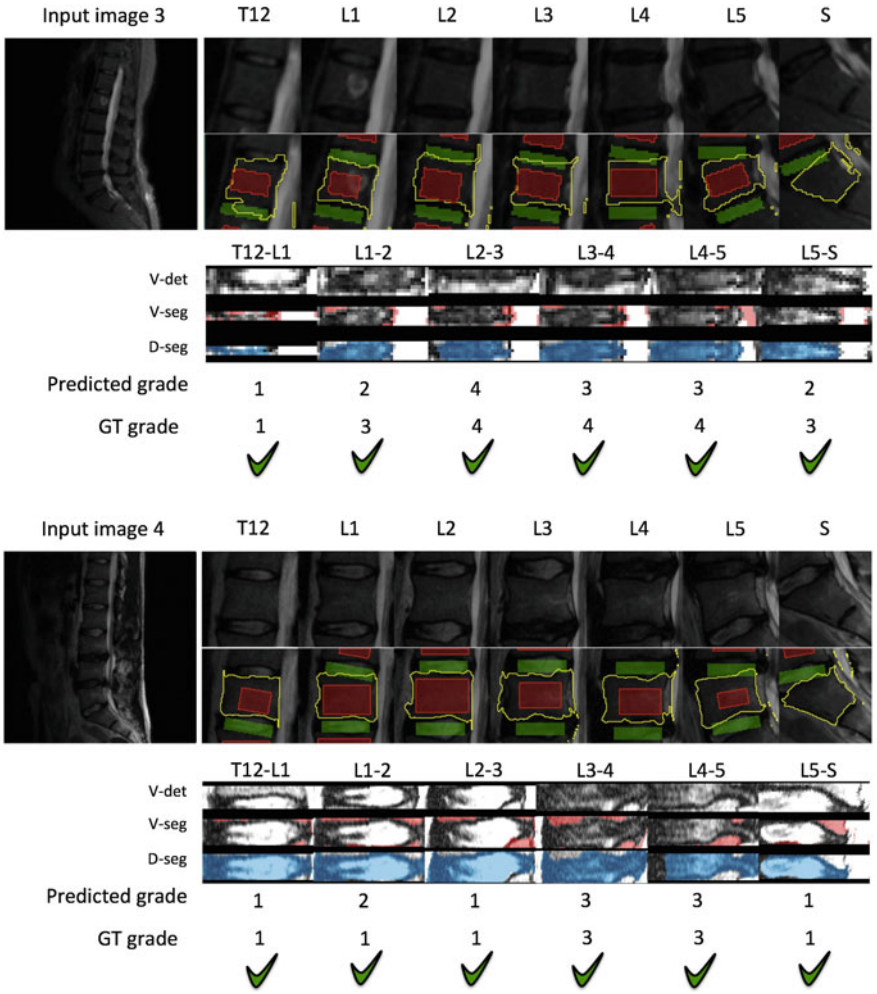

Fig. 5 Example results 3–4. The full segmentation and grading results are given for two example input images. For each patient, in the *second row* from the *top*, the vertebrae segmentations are shown as *yellow lines*. In the *bottom* three rows, the extracted regions for feature extraction, along with the segmented areas for the vertebrae and the discs are given in the three *bottom rows* for V-det, V-seg, and D-seg methods. At the *bottom*, the grading result is given for the D-seg method. For both patients 3 and 4, all the six discs succeed. Note that in patient 3, discs L1-2, L3-L4, L4-5, and L5-S, and in patient 4, disc L1-L2, are off by one grade, but still considered correct predictions according to our ± 1 criterion (Color figure online)

only, vertebra segmentation and direct disc segmentation. This spectrum represents a trade-off between specificity of the feature support region and the likelihood of failure in that step. In principle, though subject to failure, a support region that encompasses only the object of primary interest, i.e. the disc, will outperform more generic support regions.

There was a 5–7 % difference in performance between the methods employing segmentation and detection only, while there was little difference between the two segmentation approaches. That vertebrae are easier to segment than discs due to their more consistent and distinct appearance, as noted in previous work, may explain the small margin between V-seg and D-seg.

An interesting question is whether the lower performance of V-det is due its box-shaped support region or sub-optimal setting of its parameters. To answer this, we replaced the simple vertebra height based adaptation with a box height set from the height of the V-seg region. In other words, the support region is still a box but its height is based on the V-seg support region. The regression performance for this variation was 80.2, 82.6 and 84.3 % for the Mean Intensity, GMM and Hist+ features, which is very similar to the V-seg and D-seg methods. So indeed it seems that it is the size of the support region more than its exact shape that is responsible for the lower performance.

A final question is whether the sub-optimal size setting is affecting all the discs or just a subset. One might expect that the support region for the L5-S disc to be problematic due to variability in the curvature of the spine at that location. Indeed, by excluding all L5-S discs the average V-det performance improves to 80.4, 82.2 and 83.3 % for the three features (computed on the box of original size).

One limitation of our study is that the grades could only be evaluated to within ± 1 Pfirrmann grade, in effect reducing the five grades to three in the evaluation step. This was due to the variability in the ground-truth mark-up. Nevertheless, the system could still output the full range and prior work has only reported results on a binary classification of healthy versus degraded.

4 Conclusion

We have presented the first fully automated system to predict all five Pfirrmann grades. In a large dataset of 1,710 discs from 285 patients using standard clinical MRI scans acquired on different scanners, protocols and sites, our system correctly predicted to within ± 1 Pfirrmann grade in 85.8 % of discs. Our novel method proposes new features that outperform previous ones, improves robustness to MR imaging variations, and shows that disc segmentation is not essential for for this level of performance.

References

1. Alomari, R.S., Corso, J.J., Chaudhary, V.: Desiccation diagnosis in lumbar discs from clinical MRI with a probabilistic model. In: ISBI (2009)
2. Battie, M.C., Viderman, T., Gibbons, L.E., Fisher, L.D., Manninen, H., Gill, K.: 1995 volvo award in clinical sciences—determinants of lumbar disc degeneration. Spine **20**(24), 2601–2612 (1995)

3. Boykov, Y., Jolly, M.P.: Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. In: Proceedings ICCV (2001)
4. Chang, C.-C., Lin, C.-J., LIBSVM: A library for support vector machines. ACM TIST, 2:27: 1–27:27 (2011)
5. Cheng, F., You, J., Rampersaud, Y.R.: Relationship between spinal magnetic resonance imaging findings and candidacy for spinal surgery. Can. Fam. Physician, **56**, 323–330 (2010)
6. Gamio, J.C., Belongie, S.J., Majumdar, S.: Normalized cuts in 3D for spinal MRI segmentation. TMI **23**, 36–44 (2004)
7. Hao, S., Jiang, J., Guo, Y., Li, H.: Active learning based intervertebral disk classification combining shape and texture similarities. Neurocomputing **101**, 252–257 (2013)
8. Huang, S.-H., Chu, Y.-H., Lai, S.-H., Novak, C.L.: Learning-based vertebra detection and iterative normalized-cut segmentation for spinal MRI. TMI **28**(10), 1595–1605 (2009)
9. Lam, K., Anbar, A., O'Brien, A.: The correlation of the severity of lumbar disc degeneration with discogenic low back pain—a study utilizing a validated classification with awake provocative discography. J Bone Joint Surg. **92-B** (2010)
10. Law, M.W.K., Tay, K., Leung, A., Garvin, G., Shuo, L.I.: Intervertebrald disc segmentation in MR images using anisotropic oriented flux, In: Proceedings of MIA (2012)
11. Lootus, M., Kadir, T., Zisserman, A.: Vertebrae detection and labelling in lumbar MR images. In: MICCAI CSI Workshop (2013)
12. Lurie, J.D., Moses, R.A., Tosteson, A.N.A., Tosteson, T.D., Carragee, E.J., Carrino, J.A., Kaiser, J.A., Herzog, R.J.: Magnetic resonance imaging predictors of surgical outcome in patients with lumbar intervertebral disc herniation. Spine **38**, 1216–1225 (2013)
13. Michopoulou, S., Costaridou, L., Vlychou, M., Speller, R., Todd-Pokropek, A.: Texture-based quantification of lumbar intervertebral disc degeneration from conventional T2-weighted MRI. Acta Radiol. **52**, 91–98 (2011)
14. Neubert, A., Fripp, J., Engstrom, C., Walker, D., Weber, M.-A., Schwarz, R., Crozier, S.: Three-dimensional morphological and signal intensity features for detection of intervertebral disc degeneration from magnetic resonance images. J. Am. Med. Inform. Assoc. **4**, 1082–1090 (2013)
15. Pfirrmann, C.W.A., Metzdorf, A., Zanetti, M., Hodler, J., Boos, N.: Magnetic resonance classification of lumbar intervertebral disc degeneration. Spine **26**, 1873–1878 (2001)
16. Shen, J., Parent, S., Kadoury, S.: Classification of spinal deformities using a parametric torsion estimator. In: MICCAI CSI Workshop (2013)
17. Smola, A.J., Schölkopf, B.: A tutorial on support vector regression. Stat. Comput. **14**, 199–222 (2004)
18. Štern, D., Likar, B., Pernus, F., Vrtovec, T.: Parametric modelling and segmentation of vertebral bodies in 3D CT and MR images. Phys. Med. Biol. **56**, 7505 (2011)
19. Unal, Y., Kocer, H.E., Akkurt. H.E.: A comparison of feature extraction techniques for diagnosis of lumbar intervertebral degenerative disc disease. INISTA (2011)
20. Xu, Y., Gao, X., Lin, S., Wong, D.W.K., Liu, J., Xu, D. Cheng, C.-Y. Cheung, C.Y., Wong, T.Y.: Automatic grading of nuclear cataracts from slit-lamp lens images using group sparsity regression. In: MICCAI (2013)

Automated 3D Lumbar Intervertebral Disc Segmentation from MRI Data Sets

Xiao Dong and Guoyan Zheng

Abstract This paper proposed an automated 3D lumbar intervertebral disc (IVD) segmentation strategy from MRI data. Starting from two user supplied landmarks, the geometrical parameters of all lumbar vertebral bodies and intervertebral discs are automatically extracted from a mid-sagittal slice using a graphical model based approach. After that, a three-dimensional (3D) variable-radius soft tube model of the lumbar spine column is built to guide the 3D disc segmentation. The disc segmentation is achieved as a multi-kernel diffeomorphic registration between a 3D template of the disc and the observed MRI data. Experiments on 15 patient data sets showed the robustness and the accuracy of the proposed algorithm.

1 Introduction

Intervertebral disc (IVD) degeneration is a major cause for chronic back pain and function incapacity [1]. Magnetic Resonance Imaging (MRI) has become one of the key investigative tools in clinical practice to image the spine with IVD degeneration not only because MRI is non-invasive and does not use ionizing radiation, but more importantly because it offers good soft tissue contrast which allows visualization of the disc's internal structure [2].

MRI quantification has great potential as a tool for the diagnosis of disc pathology but before quantifying disc information, the IVDs need to be extracted from the MRI data. IVD extraction from MRI data comprises two key steps. Firstly, all IVDs have to be detected from the images and secondly, the regions belonging to IVDs have to be segmented. Manual extraction methods [3, 4] as well as automated extraction

X. Dong
School of Computer Science and Engineering, Southeast University, Nanjing, China
e-mail: xiao.dong@seu.edu.cn

G. Zheng (✉)
Institute for Surgical Technology and Biomechanics, University of Bern, Bern, Switzerland
e-mail: guoyan.zheng@istb.unibe.ch

© Springer International Publishing Switzerland 2015

J. Yao et al. (eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, Lecture Notes in Computational Vision and Biomechanics 20, DOI 10.1007/978-3-319-14148-0_12

methods [5–11] have been presented before. Since manual extraction is a tedious and time-consuming process which lacks repeatability, automated methods are preferred.

There are different approaches for automatizing the extraction of IVDs from medical images such as graphical model [5], probabilistic model [6], Random Forest regression and classification [12, 13], watershed algorithm [7], atlas registration [8], statistic shape model [10], graph cuts [9], and anisotropic oriented flux [11]. But stable and accurate IVD segmentation remains a challenge.

In this paper we propose an automated 3D lumbar IVD extraction method with minimal user interaction from MRI data sets. The main contribution of our method is a combination of graphical model-based spine column localization with a multi-kernel diffeomorphic registration based segmentation. The 3D IVDs are extracted with a two-step procedure where we first identify the spine column structure and then carry out the IVD segmentation. The motivation behind this two-step procedure can be explained as follows. The IVD geometries are highly constrained by the geometry of the spine column. If the geometrical parameters of the spine column and each individual vertebral body can be estimated accurately from the observed images, then they can provide both geometrical and appearance information about the intervertebral discs, which helps to improve the accuracy and robustness of the IVD segmentation.

The work flow of the proposed algorithm consists of following three steps

- **Initialization**. Two user supplied landmarks are required to indicate the centers of L1 and L5 vertebral bodies.
- **Lumbar spine column identification and modeling**. Starting from the user initialization, the 3D geometry of the lumbar spine column is automatically extracted from the 3D data sets. The outputs are the 3D geometric information of each individual vertebral body of L1-L5 and a *soft-tube* model that fits the outer surface of the lumbar spine column.
- **Lumbar disc segmentation**. Based on the prior information of the extracted lumbar spine column, the disc segmentation is achieved as a multi-kernel diffeomorphic registration between a disc template and the observed data.

2 Methods

2.1 Data Sets

All datasets used in this paper were generated from a 1.5 Tesla MRI scanner (Siemens, Erlangen, Germany). Dixon protocol was used to reconstruct four aligned high-resolution 3D volumes during one data acquisition: in-phase, opposed-phase, water and fat images, as shown in Fig. 1. Each volume has a resolution of $2 \times 1.25 \times 1.25\,\text{mm}^3$ and the data set size is $40 \times 512 \times 512$. The advantage of working with such datasets is that different channels provide complementary information for our disc segmentation task. In our proposed segmentation strategy, we always first extract

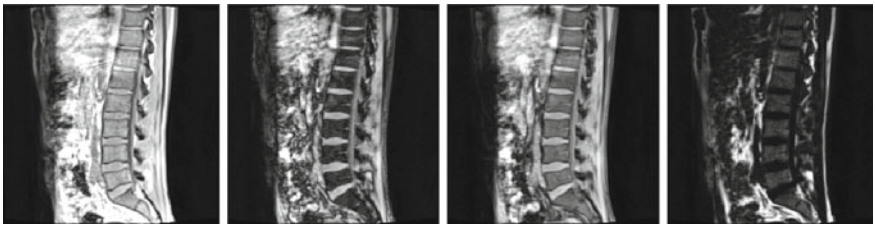

Fig. 1 The four aligned channels of a patient data (for visualization purpose, we only show the middle sagittal (mid-sagittal) slice of each channel.)

either intensity or feature information about different tissues on each channel and then combine the 4 channel data into a single dataset.

2.2 Lumbar Spine Column Identification

On the mid-sagittal slice, two landmarks are picked to indicate the centers of L1 and L5 vertebral bodies as shown in Fig. 2a. Starting from the initialization, we first carry out a 2D vertebral body and disc identification to localize vertebrae L1-L5 and the 5 target discs from the mid-sagittal slice. The geometrical information of the 2D identification is then used to guide a further 3D lumbar spine column modeling.

2.2.1 2D Vertebral Body and Disc Identification

Solutions for spine location and disc labeling include feature-based bottom-up methods, statistical model-based methods and graphical model-based solutions. For a detailed review of the existing methods, we refer to [14]. In this paper, the 2D vertebral body and disc identification is achieved using a graphical model based strategy introduced in [14]. Compared with the graphical models in [5, 6], the advantage of the graphical model in [14] is that both the low level image observation model and the high level vertebra context potentials need not to be learned from training data. Instead they are capable of self-learning from the image data during the inference procedure. For completeness, here we describe the key components of the method that we previously introduced [14].

1. **The graphical model**: The graphical model is given in Fig. 2d. Each node V_i represents a connected disc-vertebrae-disc chain of the spine, whose geometrical parameters are given by X_i. We define
   - **The component observation model**. $p(I|X_i)$ of a single component V_i representing the probability that the configuration X_i of the node V_i match the observed images I.

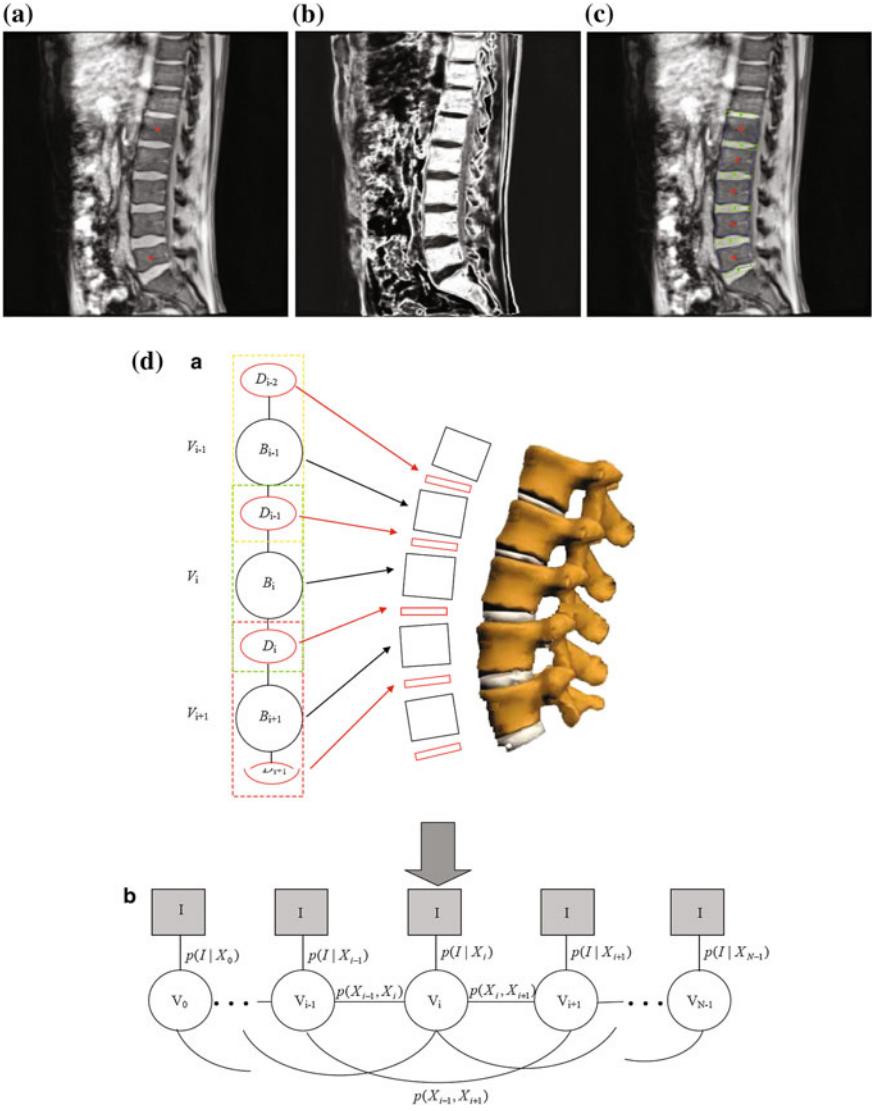

Fig. 2 Initialization and 2D lumbar spine column detection. **a** User initialization by picking two landmarks indicating the centers of L1 and L5 in the middle sagittal slice. **b** Probability assignment (displayed as grey values) of the bone tissue in the mid-sagittal slice for 2D lumbar spine column detection. **c** 2D lumbar spine column detection result using the graphical model based detection algorithm, *blue* and *green* rectangles representing the vertebral bodies and IVDs respectively. **d** Graphical model of 2D lumbar column detection (Color figure online)

- **The potentials**. $p(X_i, X_j)$ *between neighboring components* V_i and V_j encoding the geometrical constraints between components which are defined by the anatomical structure of the spine column.

The identification of the spine column from the mid-sagittal slice can then be formalized as to find the optimal configurations of $\{V_i\}$, $X = \{X_i\}$ that maximize

$$P(X|I) \propto \Pi_i p(I|X_i)\Pi_{i,j} p(X_i, X_j) \tag{1}$$

with

$$p(I|X_i) = p_I(I|X_i)p_G(I|X_i) \tag{2}$$

and

$$p(X_i, X_j) = p_S(X_i, X_j)p_O(X_i, X_j)p_D(X_i, X_j) \tag{3}$$

$p_I(I|X_i)$ and $p_G(I|X_i)$ stand for the probabilities that the observed image intensity and image gradient distributions match the geometrical parameters X_i respectively. $p_S(X_i, X_j)$, $p_O(X_i, X_j)$ and $p_D(X_i, X_j)$ are the geometrical constraints on the sizes, orientations and distances between neighboring components. All the observation models and constraints can be designed according to the observed data and prior anatomical knowledge of the spine structure. For detailed formulation of these terms, we refer to [14].

2. **Optimization**: The optimization is achieved as an inference on the graphical model. The method introduced in [14], which is a particle based nonparametric belief propagation on the graphical model, is used here to carry out the inference.

Figure 2b shows an example of the bone tissue probability assignment on the mid-sagittal slice, which is computed from the user supplied 2 landmarks (Fig. 2a) and 4 channel volume data (see Fig. 1 for an example) by a Gaussian distribution modeling and an equally weighted combination of the intensity distributions of the bony tissue in the 4 channels. This image is used for the computation of the intensity observation model $p_I(I|X_i)$ during the 2D lumbar column detection. Figure 2c gives the 2D lumbar column detection result. It can be observed that the centers, sizes and orientations of the vertebral bodies and IVDs are correctly identified.

2.2.2 3D Lumbar Spine Column Modeling

We model each lumbar vertebral body as an elliptical cylinder and the lumbar spine column as a variable-radius soft tube. Details of the modeling procedure are described as follows:

- **3D modeling of each vertebral body**: From the 2D vertebral body identification results, the position, hight, radius and orientation of each vertebral body and the image intensity distribution of the bone region can be estimated by modeling the vertebral body as a cylinder. Accordingly for each voxel in the neighbourhood of

the estimated cylinder, we can assign the probability that it belongs to the bony tissue. To refine the 3D modeling of the vertebral bodies, we then further model the vertebral body as an elliptical cylinder, a least-squares geometric fitting to the voxels assigned with a high probability (>0.8) of belonging to the bony tissue can extract the 3D geometry of each vertebral body, including the center, height, orientation and the major radius and minor radius of the elliptical cylinder model.

- **3D modeling of the spine column**: The lumbar column can be modeled as a variable-radius soft tube that contains all the extracted vertebral bodies. Given the 3D models of L1-L5 vertebral bodies, the central axis and the variable-radius of the soft tube can be obtained by a linear interpolation on the centers and radii of the extracted 3D models of vertebral bodies. This results in a 3D variable-radius soft-tube spine column model as shown in Fig. 3a.

Given the 3D soft-tube lumbar spine column model, the spine column region can be extracted from the observed data sets (Fig. 3b). By further eliminating the bony tissue region using the 3D models of vertebral bodies, the candidate region for each target disc can be localized as shown in Fig. 3c–f. The following 3D IVD segmentation is then carried out on the extracted candidate IVD regions.

2.3 3D Disc Segmentation

We solve the 3D disc segmentation as a template based registration between a geometrical disc template and the observed data.

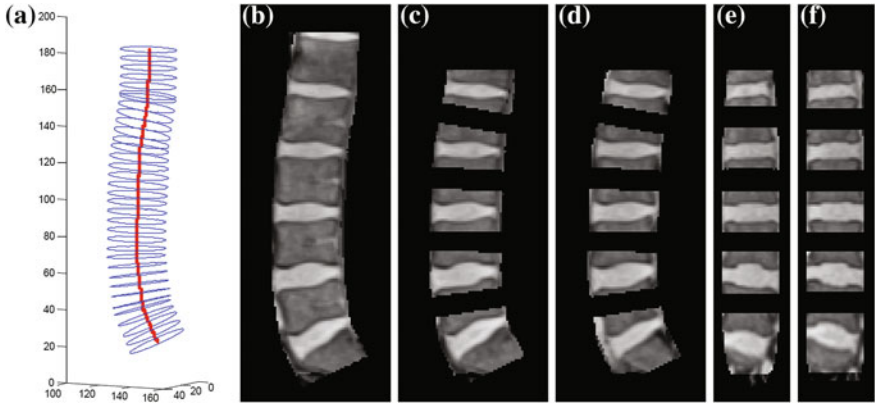

Fig. 3 3D lumbar spine column detection and modeling. **a** the 3D soft-tube model of the lumbar spine column; **b** segmented lumbar spine column image; **c**–**d** segmented disc candidate regions in sagittal slices; **e**–**f** segmented disc candidate regions in coronal slices. Although all tasks are conducted in 3D, here we show the results in 2D slices for visualization purpose

- The IVD template is set as a thin elliptical cylinder. Considering the anatomical structure of the spine column, i.e., each IVD must fall between its neighbouring vertebral bodies, the initial geometries (center, radii, orientation, hight) of the IVD cylinder template can be estimated using the 3D spine column model and the geometries of its neighboring vertebral bodies, which are all available from the previous 3D lumbar spine column modeling procedure.
- For the segmentation of a specific IVD, the correspondent observed data to be matched is just the extracted candidate IVD region as shown in Fig. 3c–f
- The registration algorithm we choose is a multi-kernel diffeomorphic image matching in the Large Deformation Diffeomorphic Metric Mapping (LDDMM) framework as described in [15] and related literatures [16–18].

2.3.1 Multi-kernel LDDMM Registration

LDDMM framework [17] is one of the two main computational frameworks in computational anatomy [16]. Existing works show that LDDMM is a general solution for nonrigid image registration with a high flexibility and accuracy. In [15] multi-kernel LDDMM registration algorithms were investigated. Compared with the LDDMM registration with a single kernel, multi-kernel LDDMM has the capability to optimize the deformation in multiple spatial scales [15].

Following the general idea of LDDMM framework, we formalize the multi-kernel image registration between two images I_0 and I_1 as an optimization problem to find the optimal time dependent velocity field $v(t)$ that minimizes the sum of a similarity and a deformation energy formalized as

$$\mathscr{E}(v^{\alpha}(t)) = \frac{1}{2}\sum_i w^i \int_0^1 \|v^i(t)\|^2_{V^i} dt + \|I_0 \circ \phi_v^{-1}(1) - I_1\|^2_{L^2} \tag{4a}$$

$$\frac{\partial}{\partial t}\phi_v(t) = v(t) \circ \phi_v(t) \tag{4b}$$

$$v(t) = \sum_i v^i(t) \tag{4c}$$

$$\phi_v(0) = Id \tag{4d}$$

where $\|v^i(t)\|_{V^i} =< v^i(t), v^i(t) >^{\frac{1}{2}}_{V^i}$ is the norm induced by the inner product $< u, v >_{V^i} =< L_{V^i}u, L_{V^i}v >_{L^2}$, and $\{K_{V^i} = (L^+_{V^i}L_{V^i})^{-1}\}$ are the kernels. The $\phi_v(t)$ is the time-dependent deformation computed as the integration of the velocity field $v(t)$ and $I_0 \circ \phi_v^{-1}(t)$ is the transformed image of I_0 by the deformation $\phi_v(t)$.

Using the optimal control based approach introduced in [19, 20], we get the Euler-Poincare equation (EPDiff) for the multi-kernel LDDMM registration algorithm as

$$\dot{I}(t) = -\nabla I(t) \cdot \sum_i v^i(t) \tag{5a}$$

$$\dot{P}(t) = -\nabla (P(t) \cdot \sum_i v^i(t)) \tag{5b}$$

$$v^i(t) = -(w^i)^{-1} K_{V^i} \star (P(t) \nabla I(t)) \tag{5c}$$

$$P(1) = -(I(1) - I_1) \tag{5d}$$

$$I(0) = I_0 \tag{5e}$$

The registration can then be carried out by updating the deformation velocity fields $v^i(t)$ iteratively from an initial value of $v^i(t)$ using Eqs. (5a)–(5e). For more details on the computation routine, we refer to [17, 19, 20].

2.3.2 Disc Segmentation by Diffeomorphic Registration

The IVD segmentation is achieved as a template based registration between the thin cylinder IVD template and the correspondent candidate disc region as shown in Fig. 3.

In order to explore both intensity and feature information to enhance the accuracy and robustness of the segmentation, we consider a simultaneous registration of two pairs of images, I_0^I / I_1^I and I_0^E / I_1^E, which stand for the image intensity and edge information template/observation pairs respectively. Accordingly in the cost function of the LDDMM registration (4a), the image similarity term includes two components $\| I_0^I \circ \phi_v^{-1}(1) - I_1^I \|_{L^2}^2 + \beta \| I_0^E \circ \phi_v^{-1}(1) - I_1^E \|_{L^2}^2$ (Fig. 4).

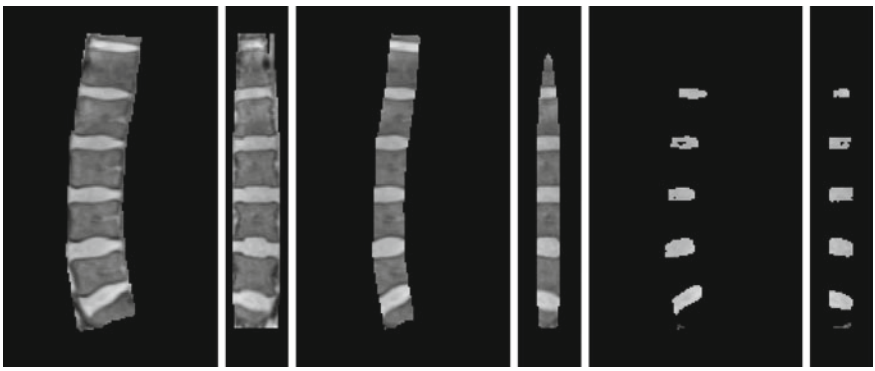

Fig. 4 Determination of the high confidence disc region using the spine column model. *Left* to *right*: The spine column region extracted using the spine column model shown in a sagittal and a coronal slice; The central region of the spine column obtained by shrinking the radius of the spine column model by a factor 0.5 shown in the same two slices; The detected high confidence disc regions by further cutting out the bone tissue using the spine column model

I_0^I / I_1^I and I_0^E / I_1^E can be computed from the initialized IVD template and the extracted IVD candidate regions. Details about how to compute them are omitted due to limited space. An example of the template images and correspondent target images and the time dependent registration procedure is shown in Fig. 5.

The final segmented IVD can then be obtained as the deformed template achieved by the multi-kernel LDDMM registration.

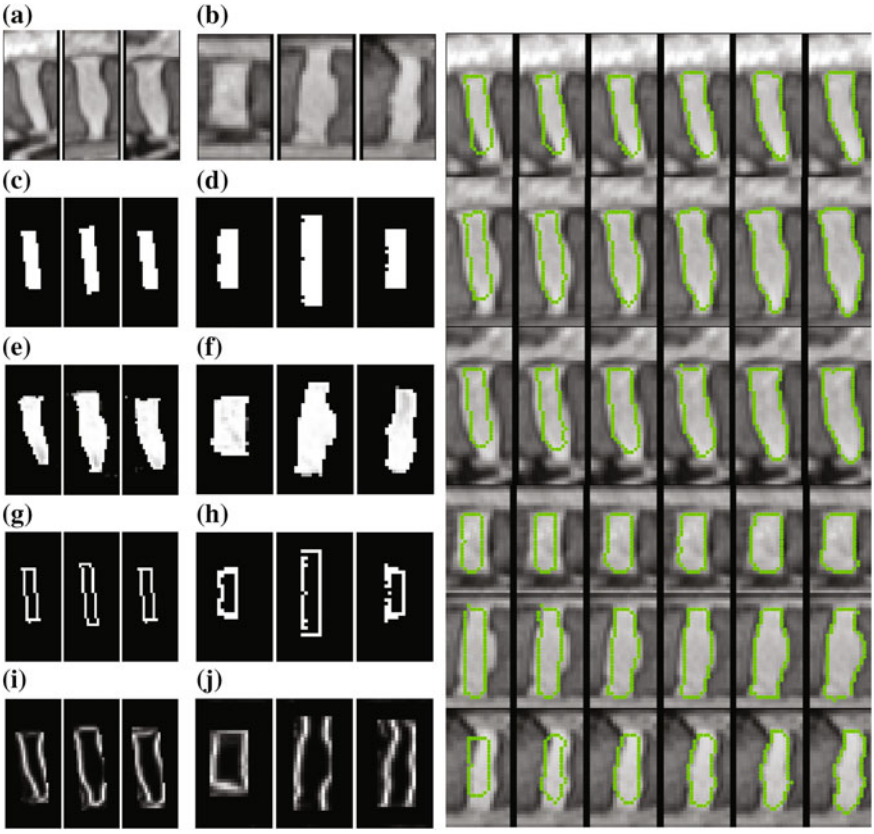

Fig. 5 3D IVD segmentation by multi-kernel LDDMM registration *Left side*: The data used in diffeomorphic registration based 3D lumbar disc segmentation. In the target images, the bone tissue regions are extracted using the spine column model. **a**–**b**: 3 sagittal/coronal slices of the candidate disc region (disc L4-L5 in Fig. 2); **c**–**d**: the intensity disc template in 3 sagittal/coronal slices; **e**–**f**: intensity information extracted from MRI data sets in 3 sagittal/coronal slices; **g**–**h**: the edge disc template in 3 sagital/coronal slices; **i**–**j**:edge information computed from MRI data sets in 3 sagittal/coronal slices; *Right side*: The time-dependent deformation of the disc template during the multi-kernel diffeomorphic registration for a L4-L5 disc segmentation. *left* to *right*: the deformations of the template at 6 time slots $t = 0, 0.2, 0.4, 0.6, 0.8, 1$. $t = 0$ means the initial template and $t = 1$ gives the final registration results; from *top row* to *bottom row*: the evolution of the template visualized in 6 different slices

3 Experiments

The proposed algorithms are verified on MRI datasets of 15 patients obtained with the Dixon protocol. In all the data sets, based on the two landmarks obtained from the initialization step, both the 2D lumbar spine column and the 3D spine column models are correctly extracted. Examples of the disc segmentation results on 4 patient data sets are shown in Fig. 6.

We also carried out quantitative evaluation of our algorithm. To do this, we manually segmented all datasets (we only need to segment one channel for each patient as all four channel volumes are aligned according to Dixon imaging protocol) and took the binary volumes from the manual segmentation as the ground truth to verify the accuracy of the present algorithm. We computed the Dice coefficient D which is usually used to measure the overlap between two binary images:

$$D = \frac{2 \times |A \cap B|}{|A| + |B|} \times 100 \tag{6}$$

Table 1 shows the average dice coefficients of the 5 discs on all 15 patients when the automated segmentation was compared to the manual segmentation. The highest average dice coefficient was found for patient #8 (87.9 %) and the lowest average dice coefficient was found for patient #9 (80.5 %). We also computed the average

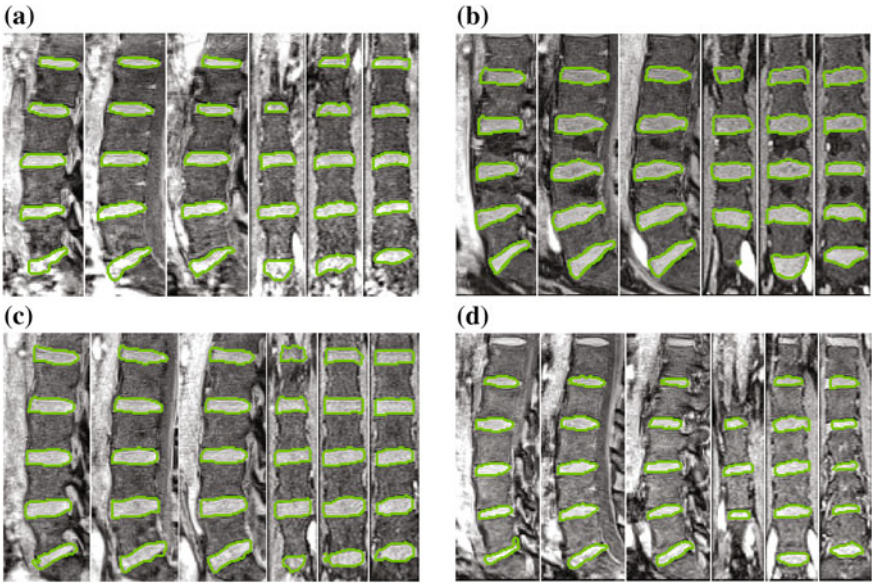

Fig. 6 3D intervertebral disc segmentation results on 4 patients. For visualization purpose, we display the results on 2D slices. For each image, the left three columns are sagittal slices and the right three are coronal slices

Table 1 Average Dice coefficients (%) of the 5 discs between the manual segmentation and the proposed algorithm on different patients

Patient	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15
Dice	86.1	81.9	82.6	86.3	86.8	83.6	87.6	87.9	80.5	84.1	86.3	85.4	86.9	87.7	83.1

Table 2 Average Dice coefficients (%) between the manual segmentation and the proposed algorithm on different discs on all 15 data sets

Disc	L1-L2	L2-L3	L3-L4	L4-L5	L5-S1
Dice	81.2	87.1	88.2	86.5	82.7

dice coefficients for all discs and the results are presented in Table 2. We note that Neubert et al. [10] reported a mean Dice of 76–80 % in their 3D IVD segmentation paper.

4 Conclusions

In this paper we proposed an automated lumbar intervertebral disc segmentation strategy, whose key components include a graphical model based spine column identification algorithm and a multi-kernel LDDMM registration algorithm to achieve the disc segmentation. By identifying the lumbar spine column structure before carrying out the segmentation, we acquire geometrical and appearance information about the spine column. These information can be used to accurately locate the candidate disc region and provide constraints to enhance the performance of the disc segmentation. By converting the segmentation problem as a template based diffeomorphic registration, we can explore both the intensity and edge information of the observed data while keeping a smooth deformation of the template so that the final segmented discs will possess smooth surfaces. The experiments on 15 patient data sets verified the robustness and accuracy of our method.

We also noticed that for abnormal cases, such as with missing/additional vertebrae or the scoliosis case, the automated lumbar column identification may not be reliable although the graphical model can handle the unknown vertebra number as shown in [14]. A possible solution for these extreme cases is to ask the user to indicate the center of each vertebra body during the initialization step. Once the centers are known, the particle filtering-based inference can then achieve a reliable 2D lumbar column identification and the following up 3D lumbar column modeling and disc segmentation.

References

1. Modic, M., Ross, J.: Lumbar degenerative disk disease. Radiology **245**(1), 43–61 (2007)
2. Parizel, P., Goethem, J.V., den Hauwe, L.V., Voormolen, M.: Degenerative disc disease. In: Van Goethem, J. (ed.) Spinal Imaging—Diagnostic imaging of the spine and spinal cord, pp. 122–133. Springer, Berlin (2007)
3. Tsai, M., Jou, J., Hsieh, M.: A new method for lumbar herniated intervertebral disc diagnosis based on image analysis of transverse sections. Comput. Med. Imaging Graph. **26**(6), 369–380 (2002)
4. Niemelainen, R., Videman, T., Dhillon, S., Battie, M.: Quantitative measurement of intervertebral disc signal using mri. Clin. Radiol. **63**(3), 252–255 (2008)
5. Schmidt, S., Kappes, J.H., Bergtholdt, M., Pekar, V., Dries, S., Bystrov, D., Schnorr, C.: Spine detection and labeling using a parts-based graphical model. In: N. Karssemeijer, B.L. (ed.) IPMI, pp. 122–133. Springer, Berlin (2007)
6. Corso, J., Alomari, R., Chaudhary, V.: Lumbar disc localization and labeling with a probabilistic model on both pixel and object features. In: Metaxas, e.a., D. (ed.) MICCAI, pp. 202–210. Springer, Berlin (2008)
7. Chevrefils, C., Cheriet, F., Aubin, C., Grimard, G.: Texture analysis for automatic segmentation of intervertebral disks of scoliotic spines from mr images. IEEE Trans. Inf. Technol. Biomed. **13**(4), 608–620 (2009)
8. Michopoulou, S., Costaridou, L., Panagiotopoulos, E., Speller, R., Panayiotakis, G., Todd-Pokropek, A.: Atlas-based segmentation of degenerated lumbar intervertebral discs from mr images of the spine. IEEE Trans. Biomed. Eng. **56**(9), 2225–2231 (2009)
9. Ayed, I.B., Punithakumar, K., Garvin, G., Romano, W., Li, S.: Graph cuts with invariant object-interaction priors: application to intervertebral disc segmentation. In: G. Szekely, H.H. (ed.) IPMI, pp. 221–232. Springer, Berlin (2011)
10. Neubert, A., Fripp, J., Schwarz, R., Lauer, L., Salvado, O., Crozier, S.: Automated detection, 3d segmentation and analysis of high resolution spine mr images using statistical shape models. Phys. Med. Biol. **57**, 8357–8376 (2012)
11. Law, M., Tay, K., Leung, A., Garvin, G., Li, S.: Intervertebral disc segmentation in mr images using anisotropic oriented flux. Med. Image Anal. **17**(1), 43–61 (2013)
12. Glocker, B., Feulner, J., Criminisi, A., Haynor, D., Konukoglu, E.: Automatic localization and identification of vertebrae in arbitrary field-of-view ct scans. MICCAI, pp. 590–598. Springer-Verlag, Berlin (2012)
13. Glocker, B., Zikic, D., Konukoglu, E., Haynor, D., Criminisi, A.: Vertebrae localization in pathological spine ct via dense classification from sparse annotations. MICCAI, pp. 262–270. Springer, Berlin (2013)
14. Dong, X., Lu, H., Sakurai, Y., Yamagata, H., Zheng, G., Reyes, M.: Automated intervertebral disc detection from low resolution, sparse mri images for the planning of scan geometries. In: Wang, F., Yan, P., Suzuki, K., Shen, D. (eds.) Machine Learning Medical Imaging. Lecture Notes in Computer Science, pp. 10–17. Springer, Berlin Heidelberg (2010)
15. Risser, L., Vialard, F.X., Wolz, R., Murgasova, M., Holm, D.D., Rueckert, D.: Simultaneous multiscale registration using large deformation diffeomorphic metric mapping. IEEE Trans. Med. Imaging **30**(10), 1746–1759 (2011)
16. Grenander, U., Miller, M.I.: Computational anatomy: An emerging discipline. Quarterly of Applied Mathematics LVI(4), pp. 617–694. (1998)
17. Beg, M.F., Miller, M.I., Trouv, A., Younes, L.: Computing large deformation metric mappings via geodesic flow of diffeomorphisms. Int. J. Comput. Vis. **61**, 139–157 (2005)
18. Miller, M.I., Trouv, A., Younes, L.: Geodesic shooting for computational anatomy. J. Math. Imaging. Vis. **24**(2), 209–228 (2006)
19. Hart, G.L., Zach, C., Niethammer, M.: An optimal control approach for deformable registration. In: Computer Vision and Pattern Recognition. (2009)
20. Vialard, F.X., Risser, L., Rueckert, D., Cotter, C.J.: Diffeomorphic 3d image registration via geodesic shooting using an efficient adjoint calculation. Int. J. Comput. Vis. **97**, 229–241 (2012)

Minimally Supervised Segmentation and Meshing of 3D Intervertebral Discs of the Lumbar Spine for Discectomy Simulation

Rabia Haq, Rifat Aras, David A. Besachio, Roderick C. Borgie and Michel A. Audette

Abstract A framework for 3D segmentation of healthy and herniated intervertebral discs from T2-weighted MRI was developed that exploits weak shape priors encoded in simplex mesh active surface models. An ellipsoidal simplex template mesh was initialized within the disc image boundary through affine landmark-based registration, and was allowed to deform according to image gradient forces. Coarse-to-fine multi-resolution approach was adopted in conjunction with decreasing shape memory forces to accurately capture the disc boundary. User intervention is allowed to turn off the shape feature and guide model deformation when internal shape memory influence hinders detection of pathology. For testing, 16 healthy discs were automatically segmented, and 5 pathological discs were segmented with minimal supervision. A resulting surface mesh was utilized for disc compression simulation under gravitational and weight loads and Meshless-Mechanics (MM)-based cutting using Simulation Open Framework Architecture (SOFA). The surface-mesh based segmentation method is part of a processing pipeline for anatomical modeling to support interactive surgery simulation. Segmentation results were validated against expert guided segmentation and demonstrate mean absolute shape distance error of less than 1 mm.

R. Haq (✉) · R. Aras · M.A. Audette
Old Dominion University, Norfolk, VA, USA
e-mail: rhaqx001@odu.edu

R. Aras
e-mail: raras001@odu.edu

M.A. Audette
e-mail: maudette@odu.edu

D.A. Besachio · R.C. Borgie
Naval Medical Center Portsmouth, Portsmouth, VA, USA
e-mail: David.Besachio@med.navy.mil

R.C. Borgie
e-mail: Roderick.Borgie@med.navy.mil

© Springer International Publishing Switzerland 2015

J. Yao et al. (eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, Lecture Notes in Computational Vision and Biomechanics 20, DOI 10.1007/978-3-319-14148-0_13

1 Introduction

Imaging studies indicate that 40 % of patients suffering from chronic back pain showed symptoms of inter-vertebral disc degeneration (IDD) [1, 2]. Primary treatment for lower back pain consists of non-surgical treatment methods. If non-surgical treatments are ineffective, a surgical procedure may be required to treat IDD, a procedure known as spinal discectomy. Approximately 300,000 discectomy procedures, over 90 % of all spinal surgical procedures [3], are performed each year, totaling up to $11.25 billion in cost per year. Other spinal surgeries include treatment for metastatic spinal tumors and spinal cord injury.

A patient-specific, high-fidelity spine anatomical model that faithfully represents any existing spine pathology can be utilized to facilitate the fusion of several spine medical images into a probabilistic intensity atlas of the spine that mirrors a brain atlas [4] and that in turn could provide priors for identifying pathologies [5].

In this paper we propose an automatic method for segmentation of intervertebral discs, in conjunction with manual guidance in presence of pathology, from high-resolution T2-weighted Magnetic Resonance Imaging (MRI). A simplex active surface mesh is initialized in the sagittal plane of a patient MRI volume, and allowed to deform using weak shape priors to capture the disc boundary. In the event that the simplex model is unable to automatically capture the boundary accurately, the user can manually guide the model deformation through constraint points placed on the image volume. We present the application of a simplex model featuring weak shape priors through automatic segmentation and controlled-resolution meshing of healthy and pathological intervertebral discs, on the basis of landmark-based registration. In addition, we utilize a surface-mesh result to initiate a tetrahedral finite-element based *Behavior Model* in SOFA to simulate healthy disc compression and herniated disc Meshless Mechanics (MM)-based cutting. Our results are validated on a clinical dataset of 16 healthy disc cases and 5 herniated disc cases, and achieve mean absolute shape distance segmentation error of less than 1 mm.

Section 2 surveys research related to intervertebral disc segmentation, Sect. 3 explains the framework of our segmentation method, compression and cutting simulation, and Sect. 4 provides research results. We discuss future work and conclusions in Sect. 5.

2 Related Research

3D segmentation of intervertebral discs is a prerequisite for the development of a computer-based surgery planning tool. Various inter-vertebral disc segmentation methods have been introduced, but are either limited to 2D segmentation of disc pathology, or 3D segmentation of healthy intervertebral discs.

Michopoulou et al. [6] proposed an atlas-based method for segmentation of degenerated lumbar intervertebral discs limited to 2D MRI scans coupled with

intensity-based classifiers, with segmentation accuracy of 91.7 % for normal and 88.6 % of degenerated discs. Alomari et al. [7] proposed a herniated disc diagnostic method that classifies pathology limited to 2D MRI scans using a Bayesian classifier with a coupled active shape model and a gradient vector flow snake for segmentation. Klinder et al. [8] and Kelm et al. [9] proposed automatic, learning-based 3D detection and segmentation frameworks of the spine, and suggested that existing disc pathologies can be reliably segmented without specifying segmentation accuracy of the pathological structure. Lalonde et al. [10] proposed kriging-based deformation of a tetrahedral template mesh of the spine, which is based on high-resolution meshes that are essentially inapplicable to interactive surgery simulation as a result of a high element count. Neubert et al. [11] results indicate potential of using statistical shape-aware models for segmentation of disc pathology without explicitly addressing herniated disc segmentation accuracy. Disc pathology generally cannot be accurately represented using strong shape priors, given that there is no average disc pathology shape.

Our method supports the ability to successfully segment disc pathology, based on semi-supervised, spatially variable weighting of weak prior shape information. We also exploit controlled-resolution meshing conducive to a multi-resolution approach to segmentation, as well as producing anatomical models with low element count for interactive simulation.

3 Method

Our segmentation approach is based on the discrete Simplex surface model. A Simplex deformable model is a physically-based system, where point vertices are treated as point masses and edges model physical properties, such as a spring-like behavior, or boundary smoothness, and is further discussed in Sect. 3.4.

Weak shape priors in Simplex mesh deformable models are exploited to deform an ellipsoidal template mesh for segmentation of an intervertebral disc. In the event that the simplex fails to accurately capture a herniated disc boundary, the user is allowed to manually facilitate the segmentation process by placing constraint points in the image volume. Similarly, ground truth for healthy intervertebral discs has also been generated by implementing this semi-supervised technique, where the user is allowed to manually assist the deformation to correct existing segmentation errors.

The remainder of this section discusses the image dataset and image preprocessing steps, followed by the automatic Simplex mesh deformation and optional user-guidance through constraint points. The data validation technique is presented to quantify the proposed framework's performance.

3.1 Image Dataset

Our test and validation dataset consists of MR images of the lumbar spine pertaining to 5 patients with various pathologies, such as herniated discs. Four intervertebral discs have been identified as degenerated, due to vertebral fractures, disc compression, and have been discarded from the dataset. Herniated discs are mostly located in the L4–L5 and L5–S1 lumbar region and have been identified in the dataset under expert supervision. T2-weighted MRI scans, acquired on a 1.5 T device using spin-echo scanning sequence with repetition time (TR) = 1,500 ms, echo time (TE) = 147 ms, flip angle = 150 and number of averages = 2, having a resolution of $0.5 \times 0.5 \times 0.9\,\text{mm}^3$ have been utilized for testing and validating the segmentation approach.

Five herniated discs has been manually segmented under expert [12] supervision to be used for quantitative evaluation of the proposed semi-supervised automatic segmentation framework, with results discussed in the Sect. 4.

3.2 Image Preprocessing

An anisotropic diffusion [13] filter (conductance = 0.8, timestep = 0.5, iterations = 50) has been applied to the volumetric images to reduce image noise within the structures while preserving image boundaries. The filter mitigates image intensity inhomogeneity located around the disc due to overlapping image intensities of the herniated disc boundary and the surrounding posterior ligament [14]. The Insight Segmentation and Registration Toolkit (ITK) [15] has been utilized for applying image preprocessing filters.

3.3 Model Initialization Through Landmark-Based Affine Registration

An ellipsoidal template mesh is initialized within the herniated disc image volume for simplex mesh deformation. Arbitrary translation, rotation and scaling effects need to be captured between the template mesh and MRI image. Six landmarks are manually placed on the ellipsoidal template mesh corresponding to landmarks within the herniated disc image boundary to initialize the template within the herniated disc image through affine registration. These landmarks are placed at the center, as well as the superior, inferior, anterior and posterior points of the disc, as well as one at the center of the superior disc surface to characterize rotation. The initialized template mesh is then allowed to automatically deform using a multi-resolution surface model, as described in Sect. 3.4.

3.4 Automatic Multi-resolution Simplex Deformation

This research exploits simplex mesh discrete deformable models for segmentation of intervertebral discs. Introduced by Delingette [16] for 3D shape reconstruction and segmentation, a k-simplex mesh is a k-manifold discrete mesh with exactly $k+1$ distinct neighbors [27]. A simplex mesh has the property of constant vertex connectivity. Simplex meshes can represent various objects depending on the connectivity k, where 1-simplex represents a curve, a 2-simplex represents a surface, and a 3-simplex represents a volume. Our research is focused on surface representation for image segmentation using 2-simplex meshes with constant 3-vertex connectivity.

The constant connectivity of the 2-simplex mesh leads to three simplex parameters corresponding to a vertex with a mass and its three neighboring vertices that are invariant under similarity transformations [16]. These independent simplex parameters can be utilized to represent the geometric constraints enforced upon a vertex with respect to its three neighbor vertices. The dynamics of each vertex P are governed by a Newtonian law of motion represented by the equation

$$m\frac{d^2P_i}{dt^2} = -\gamma\frac{dP_i}{dt} + \alpha\overrightarrow{F}_{int} + \beta\overrightarrow{F}_{ext} \tag{1}$$

where m is the vertex mass, γ is the damping force and α and β are the weight factors of the internal and external forces respectively. α has an average value of 0.4 and β has an average value of 0.5 in our approach. Appropriate values of α and β are selected prior to segmentation and are consistent across a dataset. $\overrightarrow{F}_{ext}$ is the sum of external forces governed by image edge information and gradient intensity values that minimize the distance between a vertex P and maximum gradient intensity in the neighborhood of P along the normal direction. $\overrightarrow{F}_{int}$ is the sum of internal forces represented by an elastic force that enforces smoothness and weak shape-based constraints. This physically-based deformable model is governed by forces to maintain internal stabilization through $\overrightarrow{F}_{int}$. Weak shape memory is enforced by constraining the internal forces along the normal direction of vertex P. This is implemented by constraining the mean curvature at vertex P governed by the simplex angle ϕ by setting $\phi = \phi_c$, where ϕ_c is a constant [16].

3.5 Multi-resolution Simplex Model Creation

Delingette [16] proposed four topological operators for transformation of a k-simplex mesh, which were refined by Gilles et al. [17] to optimize the simplex mesh quality by applying quality constraints. These operators yield regular mesh faces with desired edge length, leading to high quality multi-resolution meshes.

The global mesh resolution is adapted to the complexity of the anatomical shape being segmented in a coarse-to-fine segmentation approach. Thus, various simplex

Fig. 1 Multi-resolution segmentation refinement herniated disc model

mesh resolutions of a disc and toroidal shape have been generated through a multi-resolution scheme without loss of vertex connectivity for segmentation refinement, as demonstrated in Fig. 1.

Template mesh deformation is guided by the presence of MR image gradient forces, resulting in 3–6 min of segmentation time per disc or vertebra. The resulting simplex mesh is converted to a dual triangulated surface mesh with resolution control, which in turn is directly input to a tetrahedralization of similar resolution for simulation, as discussed in Sect. 3.7.

3.6 User-Guided Pathology Segmentation

In the event that the internal simplex shape memory influence hinders detection of pathology, as detected via visual inspection, user input is allowed to locally turn off the shape feature and assist model deformation. This assistance is implemented by placing *internal* and *external constraint points* on the volumetric image that gracefully constrain the deformation to correct under-and over-segmentation. Constraint point forces are enforced as an addition to the total external force. The number of constraint points applied to the images typically range between 37 and 60, while requiring 5–7 min, depending on the shape of the pathology that the automatic Simplex model deformation may fail to capture.

This semi-supervised segmentation method has also been utilized for manual correction of healthy intervertebral disc segmentation, which serves as ground truth for validation of our healthy disc segmentation results. Similar segmentation evaluation techniques have been employed in literature using manually corrected active shape models for segmentation of anatomical structures [8, 18]. Manual segmentation is a labor intensive process; each MRI volume consists of about 85–90 image slices in the sagittal plane, with an average segmentation time of approximately 7 h of manual segmentation time per disc. An anatomist performed manual and semi-supervised segmentation for generation of ground truth verified by a neuroradiologist.

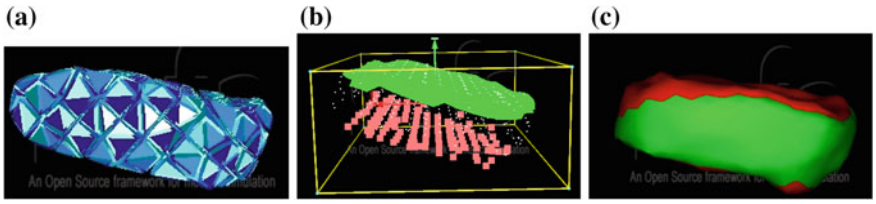

Fig. 2 3D simulation of a healthy intervertebral disc under pressure. **a** Tetrahedral FEM. **b** *Behavioral Model:* The *bottom* nodes (*red*) are constrained to be fixed and the Neumann boundary condition is applied to the *top* surface (*green*) of the disc model. **c** *Visual Model:* Comparison of the disc model at rest (*red*) and deformed (*green*) configurations (Color figure online)

3.7 Intervertebral Disc Simulation Model

Simulation Open Framework Architecture (SOFA) [19] is an open-source object-oriented software toolkit that is targeted towards real-time interactive medical simulations. Several components of a model can be combined in hierarchies through an easy-to-use scene file format to represent various model parameters such as material properties, deformable behavior, constraints and boundary conditions, which makes SOFA a very powerful and efficient prototyping tool. The following section describes two SOFA applications using the disc surface mesh: a deformation based on an FEM model and cutting based on an enriched Meshless Mechanics (MM)-based model.

3.7.1 Healthy Disc Compression Simulation

A healthy lumbar intervertebral disc has been modeled using SOFA to simulate the biomechanical and physiological changes of the disc under compression [19]. The tetrahedral mesh of the healthy L2–L3 disc has been generated from the segmented surface mesh using the isosurface stuffing method [20]. This volumetric mesh has been used to define the tetrahedral corotational finite element model of the disc, depicted in Fig. 2a, which corresponds to the *Behavior Model*[1] of the deformable object. The boundary conditions and external compression forces have been defined through the segmented surface mesh, which is linked to the underlying *Behavior Model* of the deformable object (Fig. 2b). Following the actual anatomy of the simulated intervertebral disc, the bottom nodes that are in direct contact with the below rigid vertebral body have been constrained to be fixed to their initial locations, and a prescribed vertical pressure of $100\,\mathrm{N/cm^2}$ has been applied to the top surface of the disc using SOFA's TrianglePressureForceField[2] component.

In the simulation phase, we have assumed a uniform isotropic material model for representing the intervertebral disc. Our interbertebral disc biomechanical properties

[1] http://www.sofa-framework.org/sofa/doc/sofadocumentation.pdf.

[2] http://www.sofa-framework.org/classes?show=Triangle-PressureForceField.

are consistent with values published by Malandrino et al. [21] and Spilker [22]. Using these studies, we have chosen Poisson's ratio to be 0.4 and Young's modulus to be 1,5800 Pa, representing the ratio of disc model expansion versus compression and the stiffness of the elastic model respectively. The effect of the compression force on the disc has been captured in terms of the relative displacement of the surfaces of the original and deformed configurations (Fig. 2c), where the uncompressed disc is depicted in red and the compressed configuration in green. The simulation results in a slightly bulging disc.

3.7.2 Herniated Disc Meshless Mechanics-Based Cutting

We present a proof of concept for point-based simulation and cutting of herniated disc soft-tissue to be used in the context of interactive surgery simulation setting. Figure 3a. demonstrates the application of a Meshless Mechanics (MM)-based cutting engine that runs on SOFA, in conjunction with a controlled-resolution herniated intervertebral disc model. The surgical tool that is depicted next to the disc is an 8 mm curette. A coarse-resolution triangulation leads to a tetrahedral meshing of comparable resolution, in conjunction with the it Behavior Model, while the fine-resolution surface meshing is used in conjunction with the *Visual Model*.

The continuum equation of the deformable body is discretized using the meshless moving least square (MLS) based approximation scheme. The cutting operation is realized with a novel easy-to-use intrinsic meshless distance-based enrichment technique that handles discontinuities, such as cutting, based on the work of Barbieri et al. [23]. The utilized enrichment function [24, 25] improves on [23] multiplicative enrichment in a manner that leads to full, seamless compatibility with multiple cuts. Figure 3b. depicts the meshless cutting mechanics of a simulation object in the *Visual Model*, which is controlled by the *Behavior Model*. An enrichment grid represented by point primitives is generated for the cut, which is used to update the underlying models of the deformable object. The integration of the meshless cutting model depicted in Fig. 3c, described by [25], is currently being refined for the curette tool geometry as shown in 3a.

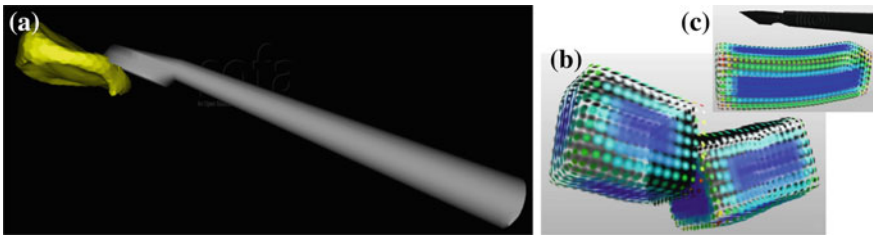

Fig. 3 Meshless Mechanics-based cutting simulation in SOFA. **a** A herniated L5–S1 disc *Visual Model* using an 8 mm curette. **b** *Visual Model* of an object represented by point primitives. **c** Meshless cutting model [25]

While there are similarities between our work and related research, our work features innovations essential to the development of an interactive spine surgery simulator. First, the anatomical modeling under development enables a trade-off between weak shape priors and limited user supervision near the pathology of interest to the simulation. Second, our approach specifically emphasizes resolution control with the final simplex surface mesh, which leads to a controlled-resolution triangulated mesh by duality; moreover the latter controlled-resolution triangulated mesh in turn leads to a like-resolution tetrahedral mesh bounded by it. Both aspects of the meshing are essential to the low element count needed for an interactive virtual tissue response.

This implementation is intended as a proof of concept to demonstrate use of segmentation results to initiate a patient-specific simulation in SOFA, such that an interactive response is feasible. Meanwhile, competing spine modeling methods emphasize dense tetrahedral decomposition and onerous finite element computations that preclude an interactive response. In particular, our controlled-resolution modeling technique can produce a coarse triangular surface for constraining a coarse tetrahedralization for a *Behavior Model*, a medium-resolution surface mesh for a *Collision Model*, and a fine-resolution surface mesh for a *Visual Model*, all running on SOFA and mapped to each other.

4 Results

MeshValmet [26] has been utilized for calculation of quantitative validation metrics. The mean absolute shape distance, MASD, (in mm) and absolute standard deviation of all errors (in mm), absolute mean square distance MSD (in mm), the Hausdorff distance (in mm) and DICE similarity coefficient comparison metrics have been calculated to compare the quality of our segmentation approach with ground truth. The Hausdorff distance is the maximum surface distance between two surface meshes and quantitatively represents a measure of the worst segmentation error. DICE similarity coefficient compares the similarity between the resulting segmentation and ground truth, and has been calculated as $s = (2|X \cap Y|)/(|X| + |Y|)$.

Statistical comparison of 16 automatic segmentations of healthy lumbar intervertebral discs with minimally supervised segmentation results, considered ground truth, is represented in Table 1. The average absolute mean error of healthy disc

Table 1 Average validation metrics comparing automatic segmentation results with corresponding semi-supervised segmentation of 16 healthy lumbar intervertebral discs

Validation metric (mm)	Healthy disc
MASD	0.321
Absolute std. dev.	0.455
MSD	0.342
Average hausdorff distance	3.261
DICE coefficient	0.954

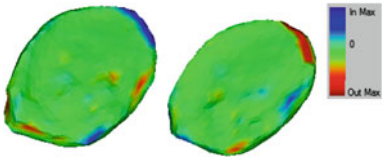

Fig. 4 Comparison of an automatic L5–S1 healthy disc segmentation result against its corresponding semi-supervised segmentation (ground truth), with −2.452 mm max. in, 2.081 mm max. out error

segmentation approach is 0.323 mm ± 0.455 mm, with an average Hausdorff distance of 3.26 mm and average DICE score of 0.954. The maximum surface error was generally located at the lateral margins of the intervertebral disc, where the automatic segmentation approach failed to faithfully capture the image boundary due to image intensity ambiguity caused by surrounding spine tissues and ligaments. Figure 4 compares automatic segmentation of a healthy L5–S1 disc with the semi-supervised segmentation result, considered as ground truth. Maximum In error corresponds with maximum under-segmentation error and maximum Out error represents the over-segmentation error. Our automatic segmentation approach under-segmented the lateral margins with a maximum In error of −2.45 mm, and a mean absolute segmentation error of 0.19 mm ± 0.29 mm.

Average results of 5 herniated discs comparing semi-supervised segmentation results against manual segmentation have been calculated. Evaluation results have been obtained by calculating the surface to mesh difference between the manual segmentation, considered ground truth, and the simplex model result from our approach of the corresponding intervertebral disc. Our approach demonstrates mean absolute shape distance of 0.61 mm ± 0.52 mm of segmentation of 5 herniated intervertebral discs (Table 2). Our results are favorable in comparison with competing 2D segmentation methods of herniated discs, and 3D segmentation methods of healthy discs respectively. Michopoulou et al. [6] reported a 2D mean absolute distance of 0.61 mm, whereas Neubert et al. [11] achieved a 3D segmented Hausdorff distance of 3.55 mm for healthy discs in high-resolution $0.34 \times 0.34 \times 1 - 1.2\,\text{mm}^3$ MR images; in our case, the average Hausdorff distance is 3.261 mm.

Table 2 Average validation metrics comparing semi-supervised segmentation results with corresponding manual segmentation of 5 herniated lumbar intervertebral discs

Validation metric (mm)	Herniated disc
MASD	0.608
Absolute std. dev.	0.518
MSD	0.638
Average hausdorff distance	3.485
DICE coefficient	0.917

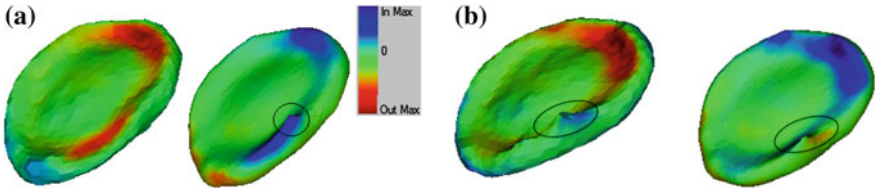

Fig. 5 Spatial segmentation error of an L5–S1 herniated disc, with disc herniation *circled* **a** Comparison of automatic segmentation using weak shape priors against manual segmentation, considered ground truth (5.022 mm max. in, 2.603 mm max. out). **b** Comparison of semi-supervised segmentation against its corresponding manual segmentation (3.369 mm max. in, 3.487 mm max. out)

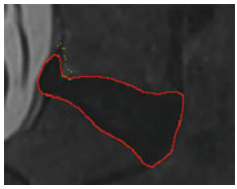

Fig. 6 Sagittal MRI slice of a herniated disc with corresponding segmentation and constraint points

Figure 5 shows the spatial distribution of error between initial automatic segmentation using weak shape priors, semi-supervised segmentation result after constraining model deformation, and the corresponding manual segmentation of the herniated disc results. Weak shape priors are successfully able to segment the disc with a maximum in error of −5.022 mm near the disc pathology. This error is reduced to −3.369 mm through semi-supervised segmentation of pathology. Figure 6 displays disc pathology with its corresponding segmentation using constraint points in a sagittal MRI slice. It can be observed that maximum error in our semi-supervised segmentation result is located at the lateral portion of the intervertebral disc. This is likely due to ambiguity in determining the intervertebral disc boundary at the lateral margins of the anatomy during manual segmentation.

Robustness to variability in user supervision and landmark-placed template mesh initialization is demonstrated in a series of experiments where the same anatomist's results are compared over several initializations, and where two anatomist results are also compared.

Table 3 displays the intra-rater and inter-rater user variability during semi-supervised segmentation of an L5–S1 herniated disc. As demonstrated in Fig. 7, intra-rater variability is present at the disc pathology where constraint points were required to correctly segment the herniated part of the anatomy. More variability exists between different anatomists, with a larger mean segmentation error of 0.254 mm, present at the lateral margins of the disc as well as the disc pathology, where manual interaction was required.

Table 3 Validation metrics comparing two sets of semi-supervised segmentations of a herniated intervertebral disc performed by the same anatomist and two different anatomists, demonstrating intra-rater and inter-rater variability respectively

Validation metric (mm)	Intra-rater	Inter-rater
MASD	0.050	0.254
Absolute std. dev.	0.062	0.323
Maximum out error	−1.214	−2.593
Maximum in error	0.905	2.290
Hausdorff distance	1.214	2.593

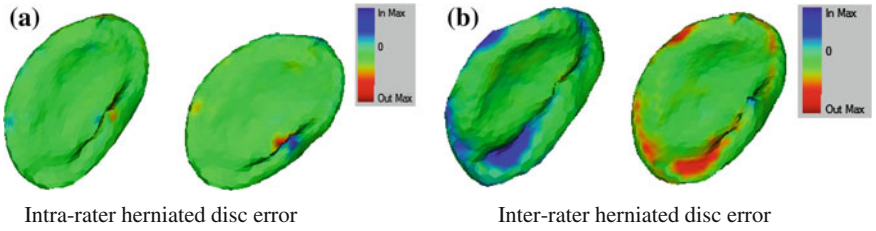

Fig. 7 Maximum out (*red*) and maximum in (*blue*) segmentation error between two sets of herniated disc segmentations performed by **a** the same rater (0.905 mm max. in, 1.214 mm max. out), and **b** different raters (2.290 mm max. in, 2.593 max out). Over and under-segmentation is present at the lateral margins and the pathology where constraint points were required to correct segmentation (Color figure online)

5 Conclusion and Future Work

In our approach, weak shape priors in active surface models are a precursor of application of statistical shape knowledge for segmentation of healthy as well as herniated discs of the lumbar spine. We have reason to believe that incorporation of statistical shape knowledge would reduce the lateral disc segmentation error, as observed in Figs. 5 and 7.

This study describes the application of weak shape priors in active surface models for segmentation of healthy as well as herniated discs. As herniated disc anatomy cannot be faithfully captured by prior shape or intensity features, weak shape prior influence is turned off locally and graceful degradation from these priors is allowed in a user-controlled manner, refining the segmentation result. Our main contribution is a framework for 3D surface-mesh based segmentation and controlled-resolution meshing of healthy and herniated discs of the lumbar spine. In addition, this paper presents a proof of concept of our segmentation method as part of a processing pipeline for anatomical models input to interactive surgery simulation, specially of a discectomy procedure, in conjunction with SOFA-based simulation that captures spine pathology with fidelity.

References

1. Luoma, K., et al.: Low back pain in relation to lumbar disc degeneration. Spine **25**(4), 487–492 (2000)
2. Freemont, A.J., et al.: Current understanding of cellular and molecular events in intervertebral disc degeneration: implications for therapy. J. Pathol. **196**(4), 374–379 (2002)
3. An, H., Anderson, P.: Disc degeneration. Spine **29**, 2677–2678 (2004)
4. Evans, A.C., et al.: 3D statistical neuroanatomical models from 305 MRI volumes. In: Proceedings of the IEEE Nuclear Science Symposium and Medical Imaging Conference, pp. 1813–1817 (1993)
5. Prastawa, M., et al.: A brain tumor segmentation framework based on outlier detection. Med. Img. Anal. **8**(3), 275–283 (2004)
6. Michopoulou, S.K., et al.: IEEE Trans. Biomed. Eng. **56**(9), 2225–2231 (2009)
7. Alomari, R.S., et al.: Lumbar spine disc herniation diagnosis with a joint shape model. In: Proceedings of MICCAI Workshop on Computational Spine Imaging (2013)
8. Klinder, T., et al.: Automated model-based vertebra detection, identification, and segmentation in CT images. Med. Image Anal. **13**, 471–482 (2009)
9. Kelm, M.B., et al.: Spine detection in CT and MR using iterated marginal space learning. Med. Imaging Anal. **17**(8), 1283–1292 (2013)
10. Lalonde, N.M., et al.: Method to geometrically personalize a detailed finite-element model of the spine. IEEE Trans. Biomed. Eng. **60**(7), 2014–2021 (2013)
11. Neubert, A., et al.: Automated detection, 3D segmentation and analysis of high resolution spine MR images using statistical shape models. Phys. Med. Biol. **57**, 8357–8376 (2012)
12. Fardon, D.F., Milette, P.C.: Nomenclature and classification of Lumbar disc pathology. Spine **26**(5), E93–E113 (2001)
13. Perona, P., et al.: Scale-space and edge detection using anisotropic diffusion. IEEE Trans. Pattern Anal. Mach. Intell. **12**(7), pp. 629–639 (1990)
14. Niadich, T.P., et al.: Imaging of the Spine. Elsevier, Philadelphia (2010)
15. Insight Segmentation and Registration Toolkit. www.itk.org
16. Delingette, H.: General object reconstruction based on simplex meshes. Intl. J. Comp. Vis. **32**(2), 111–146 (1999)
17. Gilles, B., et al.: Muskuloskeletal MRI segmentation using multi-resolution simplex meshes with medial representations. Med. Img. Anal. **14**(3), 291–302 (2010)
18. Tejos, C., et al.: Simplex mesh diffusion snakes: Integrating 2d and 3d deformable models and statistical shape knowledge in a variational framework. Intl. J. Comp. Vis. **85**(1), 19–34 (2009)
19. Faure, F., et al.: Soft Tissue Biomechanical Modeling for Computer Assisted Surgery, pp. 283–321. Springer, Heidelberg (2012)
20. Labelle, F., Shewchuk, J.R.: Isosurface stuffing: fast tetrahedral meshes with good dihedral angles. ACM. Trans. Gr. **26**(3), 57 (2007)
21. Malandrino, A., et al.: The effect of sustained compression on Oxygen metabolic transport in the intervertebral disc decreases with degenerative changes. PLoS Comput. Biol. **7**(8), e1002112 (2011)
22. Spilker, R.L.: Mechanical behavior of a simple model of an intervertebral disc under compressive loading. J. Biomech. **13**, 895–901 (1980)
23. Barbieri, E., et al.: A new weight-function enrichment in meshless methods for multiple cracks in linear elasticity. Intl. J. Num. Meth. Eng. **90**(2), 177–195 (2012)
24. Aras, R., Shen, Y., Audette, M.: Point-based methods for medical modeling and simulation. In: MODSIM World conference 2014 (2014)
25. Aras, R.: Meshless mechanics and point-based visualization methods for surgical simulations. Ph.D Dissertation, Old Dominion University (2014)
26. Mesh Valmet: Validation Metric for Meshes. http://www.nitrc.org/projects/meshvalmet/
27. Simulation Open Framework Architecture. http://www.sofa-framework.org/

Part IV
Localization

Localisation of Vertebrae on DXA Images Using Constrained Local Models with Random Forest Regression Voting

P.A. Bromiley, J.E. Adams and T.F. Cootes

Abstract Fractures associated with osteoporosis are a significant public health risk, and one that is likely to increase with an ageing population. However, many osteoporotic vertebral fractures present on images do not come to clinical attention or lead to preventative treatment. Furthermore, vertebral fracture assessment (VFA) typically depends on subjective judgement by a radiologist. The potential utility of computer-aided VFA systems is therefore considerable. Previous work has shown that Active Appearance Models (AAMs) give accurate results when locating landmarks on vertebra in DXA images, but can give poor fits in a substantial subset of examples, particularly the more severe fractures. Here we evaluate Random Forest Regression Voting Constrained Local Models (RFRV-CLMs) for this task and show that, while they lead to slightly poorer median errors than AAMs, they are much more robust, reducing the proportion of fit failures by 68 %. They are thus more suitable for use in computer-aided VFA systems.

1 Introduction

Osteoporosis is a common skeletal disorder defined by a reduction in bone mineral density (BMD) resulting in a T-score of <2.5 (i.e. more than 2.5 standard deviations below the mean in young adults), measured using dual energy X-ray absorptiometry (DXA) images [15]. It significantly increases the risk of fractures, most commonly

P.A. Bromiley (✉) · T.F. Cootes
Imaging Sciences Research Group, University of Manchester,
Manchester, UK
e-mail: paul.bromiley@manchester.ac.uk

T.F. Cootes
e-mail: timothy.f.cootes@manchester.ac.uk

J.E. Adams
Radiology & Manchester Academic Health Science Centre,
Central Manchester University Hospitals NHS Foundation Trust,
Manchester, UK
e-mail: judith.adams@manchester.ac.uk

© Springer International Publishing Switzerland 2015

J. Yao et al. (eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, Lecture Notes in Computational Vision and Biomechanics 20, DOI 10.1007/978-3-319-14148-0_14

occurring in the hips, wrists or vertebrae. Approximately 40 % of postmenopausal caucasian women are affected, increasing their lifetime risk of fragility fractures to as much as 40 % [15]. Osteoporosis therefore presents a significant public health problem for an ageing population.

Accurate identification of vertebral fractures is clinically important in the diagnosis of osteoporosis. Radiological assessment typically uses a semi-quantitative approach [10] requiring subjective judgement by a radiologist. Furthermore, only about one third of vertebral fractures present on images come to clinical attention; they are frequently not noted by radiologists, not entered into medical notes, and do not lead to preventative treatments [6]. Many of these cases involve images acquired for purposes other than VFA. However, a recent multicenter, multinational prospective study [8] has found a false negative rate of 34 % in VFA performed on lateral radiographs of the thoracolumbar spine. The potential utility of computer-aided VFA systems is therefore considerable.

Several authors have investigated the use of methods based on statistical shape models to segment vertebrae in both radiographs and DXA images (e.g. [16]) as a preliminary step for VFA. However, state-of-the-art results achieved using active appearance models (AAMs) [17] exhibit significant numbers of large errors due to fit failures, particularly on the more severely fractured vertebrae. This is the result of two effects. First, osteoporosis patients with vertebral fractures most commonly have only one or two fractures (e.g. [11]). Therefore, models encompassing multiple vertebrae must typically be trained on datasets containing more normal than fractured vertebrae, potentially introducing a bias against the most severe shape changes. Second, work on natural images of faces has shown that holistic methods such as AAMs, which rely on a single model of shape and intensity that covers all landmarks, tend to generalise poorly [7]. An alternative is to use a set of models, each covering an individual landmark. The ambiguity inherent in the use of local image patches may be dealt with by imposing a global shape constraint (e.g. [9]). In particular, regression voting (RV) methods (e.g. [19]), especially those (e.g. [3, 7, 13]) based on Random Forests (RFs) [1] tend to be robust. The RFRV Constrained Local Model (RFRV-CLM) [3, 13], which uses a RF regressor for each point constrained by a global shape model, has been applied successfully to the annotation of landmarks both in facial (e.g. [3]) and clinical (e.g. [13]) images, and shows superior generalisation on facial images compared to the AAM [18].

The hypothesis investigated here is that the superior generalisation capability of the RFRV-CLM will lead to performance improvements, compared to AAMs, in terms of the number of fit failures on DXA spinal images. RFRV-CLMs are applied to annotate vertebral landmarks in a dataset of 320 such images, the first time they have been applied to this task. Extensive experiments were performed to investigate the effect of the free parameters. The results were compared to those from [17] using AAMs on the same dataset, and show that RFRV-CLMs provide a considerable (68 %) reduction in fit failures across all vertebral classifications.

2 Method

The reader is referred to [3, 13] for full details of the RFRV-CLM algorithm; the key points are described below.

Constrained Local Models CLMs [5] build on previous work on Active Shape Models (ASMs) [4] and AAMs [2], providing a method for matching the points of a statistical shape model to an image. They combine global shape constraints with local models of the pattern of intensities. Given a set of training images with manual annotations $\mathbf{x}_l$ of a set of n landmarks $l = 1, \ldots, n$ on each, a statistical shape model is trained by applying principal component analysis (PCA) to the aligned shapes [2]. This yields a linear model of shape variation, which represents the position of each landmark l using

$$\mathbf{x}_l = T_\theta(\bar{\mathbf{x}}_l + \mathbf{P}_l\mathbf{b} + \mathbf{r}_l) \tag{1}$$

where $\bar{\mathbf{x}}_l$ is the mean position of the point in a suitable reference frame, $\mathbf{P}_l$ is a set of modes of variation, $\mathbf{b}$ are the shape parameters, $\mathbf{r}_l$ allows small deviations from the model, and T_θ applies a global transformation (e.g. similarity) with parameters θ.

To match the model to a query image, $\mathbf{I}$, the overall quality of fit Q, of the model to the image is optimised over parameters $\mathbf{p} = \{\mathbf{b}, \theta, \mathbf{r}_l\}$

$$Q(\mathbf{p}) = \Sigma_{l=1}^{n} C_l(T_\theta(\bar{\mathbf{x}}_l + \mathbf{P}_l\mathbf{b} + \mathbf{r}_l)) \quad \text{s.t.} \quad \mathbf{b}^T\mathbf{S}_b^{-1}\mathbf{b} \leq M_t \quad \text{and} \quad |\mathbf{r}_l| < r_t \tag{2}$$

where C_l is a cost image for the fitting of landmark l, $\mathbf{S}_b$ is the covariance matrix of shape model parameters $\mathbf{b}$, M_t is a threshold on the Mahalanobis distance, and r_l is a threshold on the residuals. M_t is chosen using the cumulative distribution function (CDF) of the χ^2 distribution so that 98 % of samples from a multivariate Gaussian of the appropriate dimension would fall within it. This ensures a plausible shape by assuming a flat distribution for model parameters $\mathbf{b}$ constrained within hyper-ellipsoidal bounds [2]. In the original work [2], C_l was provided by normalised correlation with a globally constrained patch model.

RF Regression Voting in the CLM Framework. In RFRV-CLM, C_l in Eq. 2 is provided by voting with a Random-Forest (RF) regressor. To train the RF for a single landmark, the shape model is used to assess the global pose, θ, of the object in each image by minimising $|T_\theta(\bar{\mathbf{x}}) - \mathbf{x}|^2$. Each image is resampled into a standardised reference frame by applying the inverse of the estimated pose. The model is scaled so that the width of the reference frame of the mean shape is a given value, w_{frame}. Sample patches of area w_{patch}^2 are then generated from the resampled images at a set of random displacements from the true point positions. The displacements $\mathbf{d}_j$ are drawn from a flat distribution in the range $[-d_{max}, +d_{max}]$ in x and y. Finally, image features $\mathbf{f}_j$ are extracted from the sample patches. Haar-like features [20] are used, as they have proven effective for a range of applications and can be calculated efficiently from integral images. To allow for inaccurate initial estimates of the pose and to make the detector locally pose-invariant, the process is repeated with random perturbations in scale and orientation of the pose estimate. A RF [1] is then trained,

using a standard, greedy approach, with the feature vectors $\mathbf{f}_j$ as inputs and the displacements $\mathbf{d}_j$ as regression targets. Each tree is trained on a bootstrap sample of N_s pairs $\{(\mathbf{f}_j, \mathbf{d}_j)\}$ from the training data. At each node, a random sub-set of n_{feat} features are chosen from this sample, and a feature f_i and threshold t that best split the data into two compact groups are selected by minimising an entropy measure [3]. Splitting terminates at either a maximum depth, D_{max}, or a minimum number of samples, N_{min}. The process is repeated to generate a forest of size n_{trees}.

RFRV-CLM Fitting Fitting to a query image is initialised via an estimate of the pose of the model e.g. from a small number of manual point annotations or a previous model, providing initial estimates $\mathbf{b}$ and θ (see Sect. 3). Equation 2 is then optimised as follows. The image is resampled in the reference frame using the current pose. Cost images C_l are then computed by evaluating a grid of points in the resampled images over a region of interest around the current estimate of each point; the grid size is defined by a search range $[-d_{search}, +d_{search}]$, and the cost images are calculated for all landmarks independently. At each point $\mathbf{z}_l$ in the grid, the required feature values are extracted and the RF regressor R_l applied. R_l then casts a vote into a cost image C_l using $C_l(\mathbf{z}_l + \delta) \rightarrow C_l(\mathbf{z}_l + \delta) + c$. Each leaf node of the RF contains the mean $\bar{\mathbf{d}}$ and covariance $\mathbf{S}_d$ of the random displacements $\mathbf{d}_i$ from the true point position, in the reference frame, of its training samples. This supports several voting styles (c, δ); a single, unit vote at $\bar{\mathbf{d}}$, or probabilistic voting by weighting with $|\mathbf{S}_d|^{-0.5}$, or by casting a Gaussian spread of votes $N(\bar{\mathbf{d}}, \mathbf{S}_d)$.

The point positions are re-estimated by finding the lowest cost point within a disk of radius r of the current position in each cost image, applying the shape model and moving $\mathbf{b}$ to nearest valid point on the limiting ellipsoid if the shape constraint in Eq. 2 is violated, updating all point positions using $\mathbf{x}_l \rightarrow T_{\theta_r}(\bar{\mathbf{x}}_l + \mathbf{P}_l\mathbf{b} + \mathbf{r}_l)$, and iterating whilst reducing $r \rightarrow k_r r$. The initial disk radius r_{max} was set to the search range d_{search}, the search was terminated at $r_t = 1.5$ pixels (in the reference image), and k_r was set to 0.7. The optimisation is described in full in Algorithm 1.

3 Evaluation

A series of experiments was performed to optimise the various free parameters and options of the RFRV-CLM for application to the task of vertebral localisation in DXA images, and to compare the results to those achieved in [17] using AAMs. To facilitate this comparison, the same dataset and performance metrics were used. The dataset consisted of 320 DXA VFA images scanned on various Hologic (Bedford MA) scanners, obtained from: (a) 44 patients from a previous study [14]; (b) 80 female subjects in an epidemiological study of a UK cohort born in 1946; (c) 196 females attending a local clinic for DXA BMD measurement, and for whom the referring physician had also requested VFA (as approved by the local ethics committee). Manual annotations of 405 landmarks were available for each image, covering the thoracic vertebrae from T7 to T12 and the lumbar vertebrae from L1 to L4. Each of

Algorithm 1 Iterative model matching procedure to estimate the shape and pose parameters in the reference frame, given a set of feature point based cost images C_l.

SHAPE MODEL AND POSE PARAMETER OPTIMISATION

Input: r_{max}, r_t, k_r, $\mathbf{x}_l$ and C_l $\forall 1 \leq l \leq n$

1. Set $r \rightarrow r_{max}$, $\theta_r \rightarrow$ Identity, $\mathbf{x}_l \rightarrow \bar{\mathbf{x}}_l + \mathbf{P}_l\mathbf{b}$, $\mathbf{r}_l = 0$
2. While $r \geq r_t$
   a. For every feature point l, find the best point $\hat{\mathbf{y}}_l$ in a disk of radius r around the current estimate
      $\hat{\mathbf{y}}_l \rightarrow \arg\max_{\mathbf{y}_l:|\mathbf{y}_l-\mathbf{x}_l|<r} C_l(\mathbf{y}_l)$
   b. Fit the shape model to these best points to estimate shape and pose parameters $\{\mathbf{b}, \theta_r\}$ by solving
      $\hat{\mathbf{y}}_l = T_{\theta_r}(\bar{\mathbf{x}}_l + \mathbf{P}_l\mathbf{b})$
   c. If $\mathbf{b}^T\mathbf{S}_b^{-1}\mathbf{b} > M_t$ then move $\mathbf{b}$ to nearest valid point on limiting ellipsoid
   d. Update all feature point positions using $\mathbf{x}_l \rightarrow T_{\theta_r}(\bar{\mathbf{x}}_l + \mathbf{P}_l\mathbf{b} + \mathbf{r}_l)$
   e. Set $r \rightarrow k_r r$ with $0 < k_r < 1$
3. Transform the resulting feature point positions into the image frame using T_θ with $\theta \rightarrow \theta \circ \theta_r$

these vertebrae in each image was also classified by an expert radiologist into one of five groups (normal, deformed but not fractured, and grade 1, 2 and 3 fractures according to the Genant definitions [10]; see Fig. 1).

Following [3, 13], 2-stage, coarse-to-fine RFRV-CLMs were used and, in common with [17], individual models were trained for each vertebra; each covered the target vertebra and its two neighbours, or one neighbour for the top (T7) and bottom (L4) vertebrae. Fitting of the first-stage model was initialised using the approximate centre points of each of the vertebrae covered by the model, calculated from the manual annotations as the centroid of the two central points on the upper and lower vertebral end-plates. This approach was adopted to avoid the significant reduction in random error, compared to individual manual annotations, that would occur if the centre points were calculated as the centroids of all manual annotations on each vertebra. Second-stage fitting was initialised using the results from the first stage. To compensate for the aperture problem present when annotating points on an extended edge, errors on automatic annotations were calculated as the mean, over each vertebra, of the minimum Euclidean distances between the automatic annotations and a Bezier spline through the manual annotations. This was applied to the points on the central vertebra in each model i.e. no use was made of the multiple fits for each point (see Sect. 4).

RFRV-CLM Parameter Optimisation. The free parameters of the RFRV-CLM, as described in Sect. 2, were divided into two sets; RF structure parameters (n_{trees}, n_{feat}, N_{min} and D_{max}), and image parameters (w_{frame}, w_{patch}, d_{max} and d_{search}).[1] These were optimised empirically and, to limit processor time requirements, a sequential approach was applied across both parameters and stages i.e. each first-stage parameter

[1] Parameters w_{patch}, d_{max} and d_{search} were defined in the reference image.

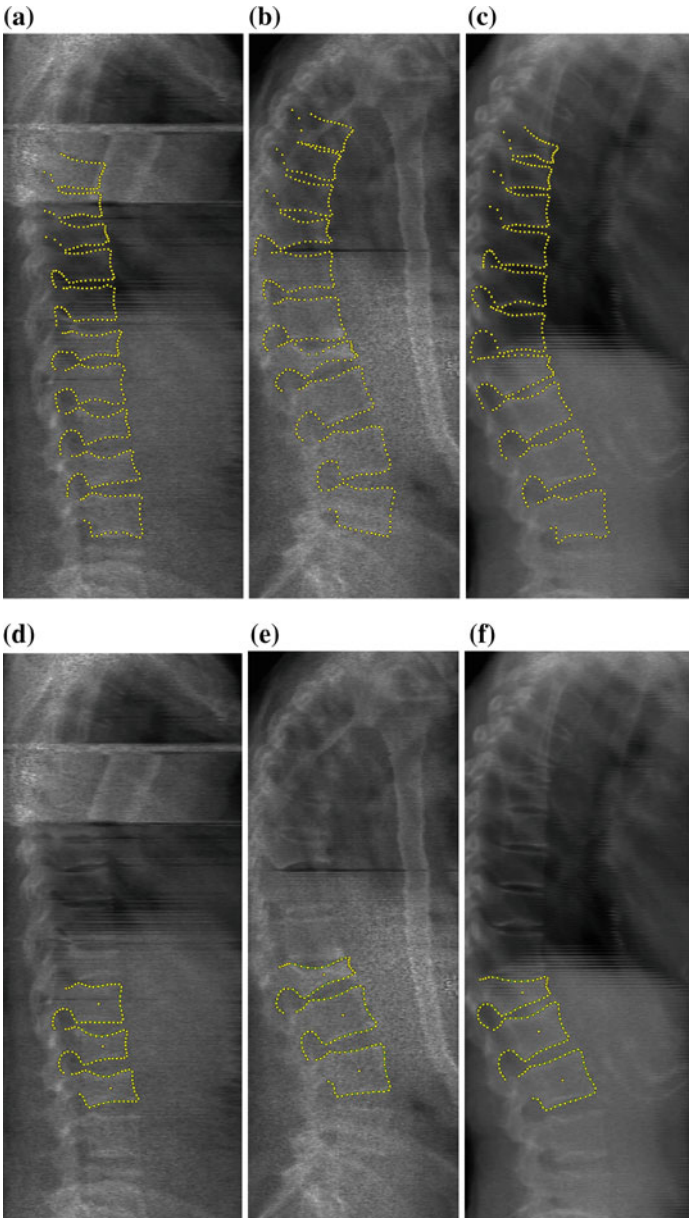

Fig. 1 Example DXA spinal images. **a**–**c** 405-point manual annotations. **d**–**f** Automatic annotation of the L2 vertebra (using the L1-L3 model), using the fully optimised, 2-stage RFRV-CLM. Example (**a**, **d**) shows grade 2 fractures on L2 and L3, (**b**, **e**) show a grade 3 fracture on L1, and (**c**, **f**) show a grade 3 fracture on L1 and a grade 1 fracture on L2

was optimised independently without applying the second-stage model; the optimal values were then fixed, and each second-stage parameter was optimised independently in a two-stage approach. Furthermore, the optimisation experiments were performed only on the L2 vertebra (i.e. the L1-L3 triplet model). L2 was chosen as it was the least obscured by confounding bony structures (ribs, scapulae, the iliac crest etc.) and imaging artefacts, minimising the contamination of the results with fitting failures. The optimisation results were then extended to the other vertebral levels by scaling the optimised w_{frame} using data on mean vertebral heights from [12], such that all image-based parameters were scaled. The data set was divided randomly into halves for training and testing. RF training includes a stochastic element, both in the random selection of data used to train each tree, and the random sub-selection from that data at each node. Therefore, each experiment was repeated five times to evaluate random errors.

For the sake of brevity, complete results are reported only for w_{frame} and w_{patch}, the parameters showing the greatest effect on performance; these are shown in Fig. 2. The graphs show the proportional area under the CDF of mean point-to-curve error on the L2 vertebra across the 160 test images. Performance generally increased with first-stage w_{frame}; however, d_{search} and w_{patch} are defined in the reference frame, and so reducing w_{frame} increases the capture range. Therefore, w_{frame} was set using the point at which the performance increase ceased to be significant, giving 40 and 110 pixels for the first and second stages. Varying w_{patch} had a smaller effect on performance over most of the range tested, but showed more complex behaviour; values of 18 and 21 pixels were selected for the first and second stages, respectively, since these were close to optimal over large portions of the tested ranges of w_{frame}. The remaining parameters were optimised similarly, giving (first stage, second stage): $n_{trees} = 2, 15$; $n_{feat} = 100, 200$; $N_{min} = 1, 1$; $D_{max} = 30, 30$; $w_{frame} = 40, 110$ pixels; $w_{patch} = 18, 21$ pixels; $d_{max} = 15, 15$ pixels; $d_{search} = 15, 10$ pixels (all pixel units except for w_{frame} were defined in the reference frame). In general, the dependence of performance on parameters was weak over large ranges.

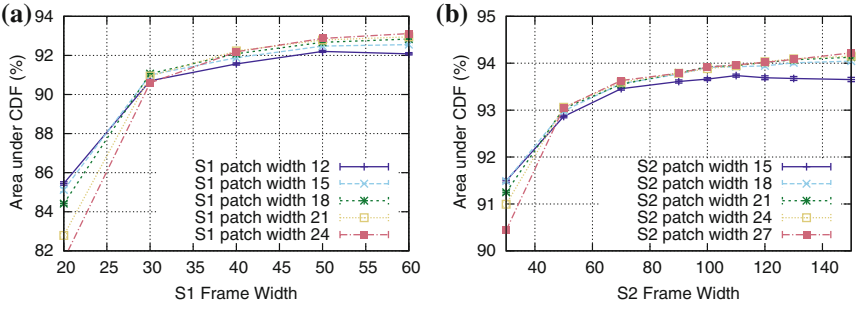

Fig. 2 Optimisation of the patch and frame width for first—(S1, **a**) and second-stage (S2, **b**) RFRV-CLMs on the L2 vertebra. Performance was measured as the percentage area under the CDF of mean point-to-curve error across 160 test images. Error bars are too small to be visible at this scale

Effects of the Shape Model Constraint. A concern with the application of shape models is that the shape constraint may introduce a bias towards the mean of the training data; this is sub-optimal in clinical applications, where the pathological (i.e. outlying) cases are of most interest. To evaluate the effect, experiments were performed using the procedure and optimal parameters described above, both with and without the application of the shape model constraint during the fitting of the second stage of the RFRV-CLM, such that the shape constraint only aided in the approximate location of the global optimum; the final result was based on image information alone. The results are shown in Fig. 3 as the CDF of mean point-to-curve error for the L2 vertebra over the 160 test images. The elimination of the shape model constraint resulted in a small and statistically insignificant change in performance, indicating that any shape model bias had an insignificant effect on the results given the error measure used here.

Effects of the Voting Style. As described in Sect. 2, several methods for voting into the cost images were available. These alternatives were evaluated using the experimental procedure and optimised parameters described above, to determine whether probabilistic voting provided performance enhancements. The results are shown in Fig. 4 as the CDF of mean point-to-curve error for the L2 vertebra over the 160 test images.

The results show that probabilistic voting provided no performance advantage. The performance of Gaussian voting was almost identical to that of single, unit voting. Single, weighted voting resulted in a small decrease in performance; however, these differences were not statistically significant. Similar results have previously been found when applying RFRV-CLMs to facial images [3].

Performance across Multiple Vertebrae. A set of leave-1/4-out experiments was performed to evaluate the RFRV-CLM on all vertebrae between T7 and L4 in all 320 images. The optimised parameters were derived from the L2 vertebra; they were adapted for the other vertebrae by scaling w_{frame} according to the ratio of mean vertebral heights in normal subjects from [12]. Shape model constraints were applied in all stages and single, unit voting was used.

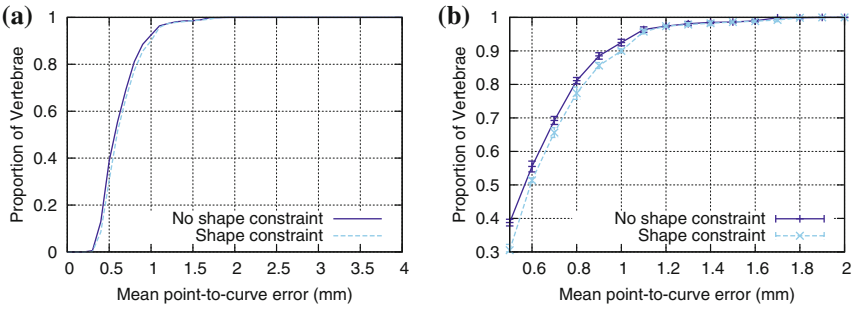

Fig. 3 Evaluation of the effect of the shape constraint in the final fitting stage using the L2 vertebra. Error bars are given as the standard deviation across five repeats; (**b**) is an expanded view of (**a**)

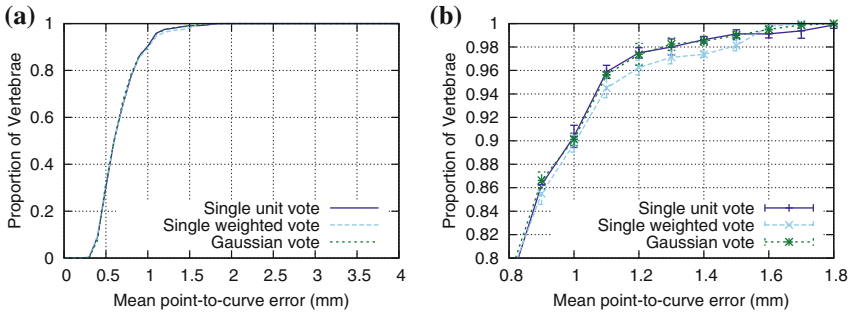

Fig. 4 Evaluation of various voting strategies for the L2 vertebra. Error bars are given as the standard deviation across five repeats; (**b**) is an expanded view of (**a**)

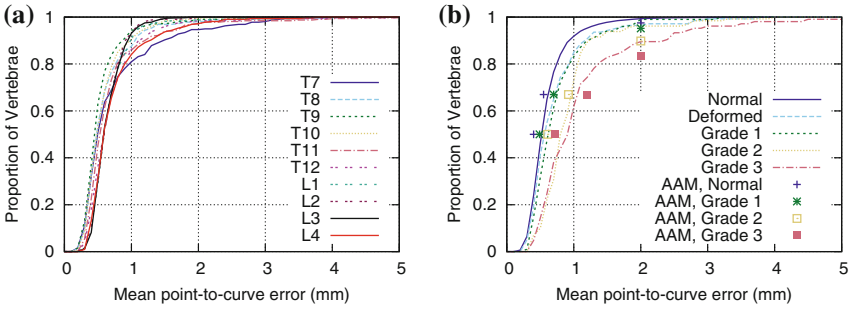

Fig. 5 Evaluation of optimised, two-stage RFRV-CLMs on all vertebrae between T7 and L4. **a** CDF of mean point-to-curve error for each vertebral level. **b** CDF of mean point-to-curve error for each vertebral status. The points show the results achieved by applying an AAM to the same data ([17], Table 1)

The results are shown in Fig. 5, and example fits are shown in Fig. 1. Annotation accuracy for T7 and L4 was lower than that for the other vertebrae, reflecting the fact that the corresponding models covered only two vertebrae, rather than three. The performance also decreased with increasing vertebral level; this may reflect the smaller size of the higher vertebrae and the presence of confounding bony structures (ribs, scapulae; see Fig. 1) in the thoracic region. However, mean errors of less than 2 and 4 mm were achieved for 95 and 99 % respectively of the vertebrae at all levels. The results divided according to vertebral status show that, as expected, performance decreased with increasing severity of fracture i.e. increasing deformation relative to the mean shape. However, mean errors of <4 mm were achieved for 95 % of grade 3 fractures, and 100 % of other classifications.

Table 1 provides numerical performance measures for the RFRV-CLMs, and compares them to the state-of-the-art results reported in [17], which applied AAMs to the same task and dataset. The AAM achieves better performance at the lower end of the CDF, as indicated by lower median errors, indicating smaller random errors on individual points. However, the RFRV-CLM achieves better mean errors for the

Table 1 Statistics of the mean point-to-curve errors on each vertebra for AAM and RFRV-CLM; bold numbers are the best results for each statistic/status

Vertebra status	AAM [17]				RFRV-CLM			
	Percentage of sample (%)	Mean (mm)	Median (mm)	Percentage errors >2 mm (%)	Percentage of sample (%)	Mean (mm)	Median (mm)	Percentage errors >2 mm (%)
Normal	84.9	**0.55**	**0.40**	2.5	83.25	0.59	0.52	**0.68**
Deformed	–	–	–	–	4.38	0.72	0.57	3.57
Grade 1	5.9	**0.70**	**0.49**	4.8	3.16	0.73	0.60	**0.99**
Grade 2	5.1	0.92	**0.61**	10.2	4.06	**0.91**	0.80	**3.84**
Grade 3	4.1	1.19	**0.72**	16.5	3.28	**1.11**	0.90	**10.48**

A total of 60 vertebrae were unclassified by the expert, due to one of the end-plates not being visible, and so 3,140 vertebrae were included in the RFRV-CLM results. Roberts et al. [17] also excluded the deformed class, hence the differences between columns 2 and 6

more severely fractured vertebrae, and substantially lower numbers of vertebrae with mean errors >2 mm regardless of classification, indicating better performance at the higher end of the CDF i.e. a smaller number of fit failures (errors >2 mm on 3.6 % of all vertebrae for AAMs versus 1.2 % for RFRV-CLMs). The median and mean errors across all 3,200 vertebrae were 0.60 and 0.65 mm respectively, compared to 0.43 and 0.60 mm from [17]. Mean search time was 366 ms per triplet per image on a Dell Precision workstation with 2 Intel Xeon 5670 processors and 24 GB RAM, running OpenSuse 11.3 × 64 (Linux kernel 2.6.34), using a single core. Mean search time per image (i.e. for all ten triplets) was 3.7 s.

4 Discussion and Conclusions

This paper has compared the performance of multi-stage RFRV-CLMs to that of AAMs in the task of vertebral landmark annotation on DXA spinal images. Several preliminary experiments were performed to optimise the various free parameters and options of the algorithm. In particular, no significant performance differences were observed, for the error metrics used, either when implementing fully probabilistic regression voting or when eliminating the shape model constraint in the final stage of fitting, such that the result was driven by image information alone.

Comparison of the errors on automatic landmarks from AAMs and RFRV-CLMs can be divided into two components; the random errors on landmarks from successful fits, best represented by the median of the error distribution due to its non-Gaussian shape, and the number of fit failures. Application of fully optimised models to ten vertebral levels in 320 DXA spinal images showed that, whilst the AAM produced smaller median errors, the difference was small at less than 0.2 mm regardless of classification. For comparison, the Genant method for vertebral fracture classification [10] defines grade 1, 2 and 3 fractures as vertebral height reductions of 20–25, 25–40, and >40 %, respectively, and [12] measured mean vertebral heights varying from 22.97 ± 1.52 mm for T7 to 35.62 ± 2.21 mm for L4 in a sample of 108 normal women. Therefore, vertebral fractures are defined via height reductions of $\gtrsim$5 mm regardless of grade or level. The more significant difference between the two techniques is in terms of the number of fit failures, since these represent cases where accurate diagnosis of the vertebral status using the automatic landmarks would not be possible. Fit failures were identified using a threshold of 2 mm, for ease of comparison to the results presented in [17]. The RFRV-CLM produced lower numbers of vertebrae with errors >2 mm for all classifications; a reduction of 68 %. Therefore, in the region of the CDF important for computer-aided VFA, RFRV-CLMs out-perform AAMs.

In this work, no use was made of the multiple fits to each vertebra provided by the overlaps of the models; only the points on the central vertebra in each were used. In future work, we intend to explore the combination of the multiple fits with goodness-of-fit measures both to improve the accuracy of the automatic annotation, in terms of random error, and to detect instances of fit failures i.e. systematic errors in individual fits. Furthermore, we intend to extend the work to include both radiographs

and mid-line sagittal CT images. Finally, we intend to investigate the use of automatic landmarks for vertebral classification, comparing the accuracy of approaches based on the Genant height ratios to classifiers applied both to the point locations themselves, and to the shape parameters generated during fitting.

Acknowledgments This publication presents independent research supported by the Health Innovation Challenge Fund (grant no. HICF-R7-414/WT100936), a parallel funding partnership between the Department of Health and Wellcome Trust. The views expressed in this publication are those of the authors and not necessarily those of the Department of Health or Wellcome Trust.

References

1. Breiman, L.: Random forests. Mach. Learn. **45**, 5–32 (2001)
2. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. IEEE Trans. Pattern Anal. Mach. Intell. **23**, 681–685 (2001)
3. Cootes, T.F., Ionita, M.C., Lindner, C., Sauer, P.: Robust and Accurate Shape Model Fitting Using Random Forest Regression Voting. In: Proc. ECCV, pp. 278–291. Springer, Berlin (2012)
4. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models—their training and application. Comput. Vis. Image Underst. **61**(1), 38–59 (1995)
5. Cristinacce, D., Cootes, T.: Automatic feature localisation with constrained local models. J. Pattern Recognit. **41**(10), 3054–3067 (2008)
6. Cummings, S.R., Melton, J.: Epidemiology and outcomes of osteoporotic fractures. Lancet **359**(9319), 1761–1767 (2002)
7. Dantone, M., Gall, J., Fanelli, G., Van Gool, L.: Real-time facial feature detection using conditional regression forests. In: Proc. CVPR, pp. 2578–2585 (2012)
8. Delmas, P.D., van de Langerijt, L., Watts, N.B., Eastell, R., Genant, H.K., Grauer, A., Cahall, D.L.: Underdiagnosis of vertebral fractures is a worldwide problem: the IMPACT study. J. Bone Miner. Res. **20**(4), 557–563 (2005)
9. Donner, R., Menze, B., Bischof, H., Langs, G.: Fast anatomical structure localization using top-down image patch regression. In: Medical Computer Vision, Recognition Techniques and Applications in Medical Imaging, Lecture Notes in Computer Science, vol. 7766, pp. 133–141 (2013)
10. Genant, H.K., Wu, C.Y., Kuijk, C.V., Nevitt, M.C.: Vertebral fracture assessment using a semiquantitative technique. J. Bone Miner. Res. **8**(9), 1137–1148 (1993)
11. Leech, J.A., Dulberg, C., Kellie, S., Pattee, L., Gay, J.: Relationship of lung function to severity of osteoporosis in women. Am. Rev. Respir. Dis. **141**(1), 68–71 (1990)
12. Leidig-Bruckner, G., Minne, H.W.: The spine deformity index (SDI); a new approach to quantifying vertebral crush fractures in patients with osteoporosis. Vertebral Fracture in Osteoporosis, pp. 235–252. Osteoporosis Research Group, University of California, California (1995)
13. Lindner, C., Thiagarajah, S., Wilkinson, J.M., arcOGEN Consortium, T., Wallis, G.A., Cootes, T.F.: Fully automatic segmentation of the proximal femur using random forest regression voting. IEEE Trans. Med. Imaging **32**(8), 1462–1472 (2013)
14. McCloskey, E., Selby, P., de Takats, D., Bernard, J., Davies, M., Robinson, J., Francis, R., Adams, J., Pande, K., Beneton, M., Jalava, T., Loyttyniemi, E., Kanis, J.A.: Effects of clodronate on vertebral fracture risk in osteoporosis: a 1-year interim analysis. Bone **28**(3), 310–315 (2001)
15. Rachner, T.D., Khosla, S., Hofbauer, L.C.: Osteoporosis: now and the future. Lancet **377**(9773), 1276–1287 (2011)
16. Roberts, M.G., Cootes, T.F., Adams, J.E.: Vertebral morphometry: semi-automatic determination of detailed shape from DXA images using active appearance models. Investig. Radiol. **41**(12), 849–859 (2006)

17. Roberts, M.G., Cootes, T.F., Adams, J.E.: Automatic Location of Vertebrae on DXA Images Using Random Forest Regression. In: Proc. MICCAI 2012, LNCS, vol. 7512, pp. 361–368. Springer-Verlag, Berlin (2012)
18. Sauer, P., Cootes, T., Taylor, C.: Accurate Regression Procedures for Active Appearance Models. In: J. Hoey, S. McKenna, E. Trucco (eds.) Proc. BMVC, pp. 30.1–30.11 (2011)
19. Valstar, M.F., Martinez, B., Binefa, X., Pantic, M.: Facial point detection using boosted regression and graph models. In: Proc. CVPR, pp. 2729–2736 (2010)
20. Viola, P., Jones, M.: Rapid object detection using a boosted cascade of simple features. In: Proc. CVPR, pp. 511–518. IEEE Computer Society (2001)

Bone Profiles: Simple, Fast, and Reliable Spine Localization in CT Scans

Jiří Hladůvka, David Major and Katja Bühler

Abstract Algorithms centered around spinal columns in CT data such as spinal canal detection, disk and vertebra localization and segmentation are known to be computationally intensive and memory demanding. The majority of these algorithms need initialization and try to reduce the search space to a minimum. In this work we introduce *bone profiles* as a simple means to compute a tight ROI containing the spine and seed points within the spinal canal. Bone profiles rely on the distribution of bone intensity values in axial slices. They are easy to understand, and parameters guiding the ROI and seed point detection are straight forward to derive. The method has been validated with two datasets containing 52 general and 242 spine-focused CT scans. Average runtimes of 1.5 and 0.4 s are reported on a single core. Due to its slice-wise nature, the method can be easily parallelized and fractions of the reported runtimes can be further achieved. Our memory requirements are upper bounded by a single CT slice.

1 Introduction

Image analysis methods focussing on spinal columns in CT scans such as detection, localization and segmentation of disks, vertebrae and spinal canal need high computational power and memory. This is particularly true for scans containing lower extremities or even full body scans. In such cases it turns out that for any spine related image processing task a significant portion of the CT data is unnecessary to deal with. Slices containing legs, which often make up the half of the data, can be left

J. Hladůvka (✉) · D. Major · K. Bühler
VRVis Center for Virtual Reality and Visualization, Donau-City-Strasse 1,
1220 Vienna, Austria
e-mail: jiri.hladuvka@vrvis.at

D. Major
e-mail: david.major@vrvis.at

K. Bühler
e-mail: katja.buehler@vrvis.at

© Springer International Publishing Switzerland 2015

J. Yao et al. (eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, Lecture Notes in Computational Vision and Biomechanics 20, DOI 10.1007/978-3-319-14148-0_15

out already during volume reconstruction from the DICOM input. Parts too far from the spinal canal can be ignored during the search for vertebrae and intervertebral disks.

In this work we present a fast, memory efficient and easy to implement method to find bounding boxes around spinal columns and to propose reliable initialization seeds for subsequent algorithms. We propose a one-pass, slice-wise method based on simple bone distribution signatures. It is:

slice-wise, i.e. memory-efficient. Every DICOM slice is processed independently of the others. Memory requirements shrink to allocation of one image.

one-pass, i.e. fast. Each slice needs to be processed only once. The runtime therefore scales linearly with the total number of slices and the slice resolution. Being slice-wise the algorithm is furthermore suitable for parallelization, e.g. a map-reduce paradigm.

easy to implement. Contrary to machine learning approaches, we propose simple, threshold-based, bone-distribution descriptors and map-reduce them into 1D arrays referred to as *bone profiles*. The algorithm has 6 easy to understand parameters. We list them together with the values used for the validation.

The rest of this paper is organized as follows: We briefly describe state-of-the-art in Sect. 2. In Sect. 3 we propose simple bone distribution descriptors and introduce their aggregation in bone profiles. In Sect. 4 we show how the profiles help localizing the spine. The proposed method is evaluated in Sect. 5. This paper ends with discussion and conclusions in Sect. 6.

2 Related Work

Spine ROI localization is conducted as an initial step in the literature in order to accelerate the detection and segmentation of the spinal parts. The authors of [3] start with a slice-wise detection of candidate positions located on bones and extend the positions to regions where features are extracted. The features are then compared to a sample set of previously annotated vertebra regions to find the best candidate which surrounds the vertebra. The approach of [7] extracts disk clue points and fits polynomials to them on every sagittal slice of an MR scan. Vertebra height statistics are computed along the fitted polynomials and the polynomial with the minimum vertebra height variance indicates the slice for further processing. Stern et al. [8] work with 3D CT and MR scans. They extract the spinal centerline by using gradient vectors of opposite pairs of vertebral body edge points. In [10] at first spine regions are localized slice-wise through thresholding and connected component analysis. The watershed algorithm is then used to extract spinal canal candidates. Graph search helps to find the spinal canal. Klinder et al. [5] apply Generalized Hough Transform using vertebral foramen meshes. The global maximum in the Hough space corresponds to a position in the spinal cord. Kelm et al. [4] localize

spine regions roughly as a first step such as we do. However, they apply the Marginal Space Learning algorithm on machine learned position candidates in CT and MR scans. In summary and in contrast to our work, the aforementioned algorithms for spine localization use either more costly features, or more complex machine-learned models, or a combination of both.

3 Bone Distribution Descriptors and Profiles

We assume sequences of axial CT slices with known z coordinate and pixel spacing, both given in millimeters. We denote the pixel positions in the patient coordinates as $\mathbf{p} = (p^x, p^y, p^z)$. For sake of simplicity we also assume the feet-to-head, face-up (supine) orientation of the patient. The pixel gray values will be denoted as $g = g(\mathbf{p})$ and we assume them to be in Hounsfield units.

Our approach is based on bone distribution signatures within slices. For a fixed slice at location z we are therefore first interested in a *rough* segmentation $\mathscr{B}_z$ of the bones. Given the Hounsfield intensities g, this can be achieved by an interval threshold, i.e., using two constants:

$$\mathscr{B}_z = \{\mathbf{p} = (p^x, p^y, p^z) \quad | \quad p^z = z \ \wedge \ g(\mathbf{p}) \in [400, 1050]\} \tag{1}$$

3.1 Centroids and Deviations

The simplest features are based on the centroid of segmentations $\mathscr{B}_z$

$$\boldsymbol{\mu_z} \ = \ \frac{1}{|\mathscr{B}_z|} \sum_{\mathbf{p} \in \mathscr{B}_z} \mathbf{p} \ = \ (\mu_z^x, \mu_z^y, z) \tag{2}$$

and on the length σ_z of the associated standard deviation vector

$$\sigma_z = \sqrt{\frac{1}{|\mathscr{B}_z| - 1} \sum_{\mathbf{p} \in \mathscr{B}_z} (p^x - \mu_z^x)^2 + (p^y - \mu_z^y)^2}. \tag{3}$$

Centroids $\boldsymbol{\mu_z}$ correlate with the spine reliably in the lumbar slices where pelvis, ribs, or head do not contribute to it.

The lumbar part can be characterized by deviation lengths σ_z related to size of a vertebra seen in an axial slice [9]. Values of σ_z larger than 40 mm indicate presence of non-vertebra bones.

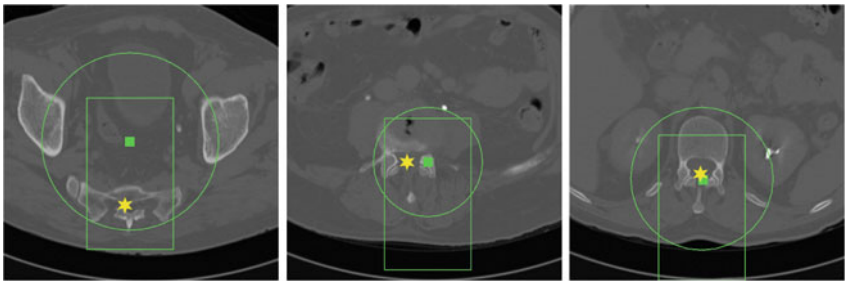

Fig. 1 Examples showing centroids $\boldsymbol{\mu}_\mathbf{z}$ (*square*), a *circle* of radius σ_z, the 80×140 mm refinement window (*rectangle*), and refined center $\boldsymbol{\nu}_\mathbf{z}$ (*star*)

Centroid refinement. While reliable in the lumbar area, the centroids $\boldsymbol{\mu}_\mathbf{z}$ may drift remarkably from the spine if the pelvis or ribs contribute by their pixels (cf. Fig. 1).

To avoid this we refine the centroids $\boldsymbol{\mu}_\mathbf{z}$ within rectangular 80×140 mm windows, asymmetrically spanned around them:

$$\mathscr{W}_z = \{\mathbf{p} \in \mathscr{B}_z \mid -40 \le p^x - \mu_z^x \le 40 \ \wedge \ -40 \le p^y - \mu_z^y \le 100\} \tag{4}$$

The size of the windows is set to be sufficiently big to accommodate any vertebra in an axial view [9] and to account for relative positions of the centroids and vertebrae in pelvis slices. The centroids $\boldsymbol{\mu}_\mathbf{z}$ are refined to the center $\boldsymbol{\nu}_\mathbf{z}$ of bone pixels in this window:

$$\boldsymbol{\nu}_\mathbf{z} = \frac{1}{|\mathscr{W}_z|} \sum_{\mathbf{p} \in \mathscr{W}_z} \mathbf{p} \tag{5}$$

3.2 Shape Histograms: AP Versus LR Distribution

To identify leg slices, we propose to discriminate slices with bone distributions dominant in the left-to-right direction and zero contributions in the anterior-posterior direction (see Fig. 2).

We construct 4-bin histograms located in the refined centers $\boldsymbol{\nu}_\mathbf{z}$. Putting $\delta = \mathbf{p} - \boldsymbol{\nu}_\mathbf{z}$ we define the following four quantities:

$$h_z^A = |\{\mathbf{p} \in \mathscr{B}_z \mid \delta^y < -|\delta^x| \le 0\}| \tag{6}$$

$$h_z^P = |\{\mathbf{p} \in \mathscr{B}_z \mid \delta^y > \ \ |\delta^x| \ge 0\}| \tag{7}$$

$$h_z^R = |\{\mathbf{p} \in \mathscr{B}_z \mid \delta^x < -|\delta^y| < 0\}| \tag{8}$$

$$h_z^L = |\{\mathbf{p} \in \mathscr{B}_z \mid \delta^x > \ \ |\delta^y| > 0\}| \tag{9}$$

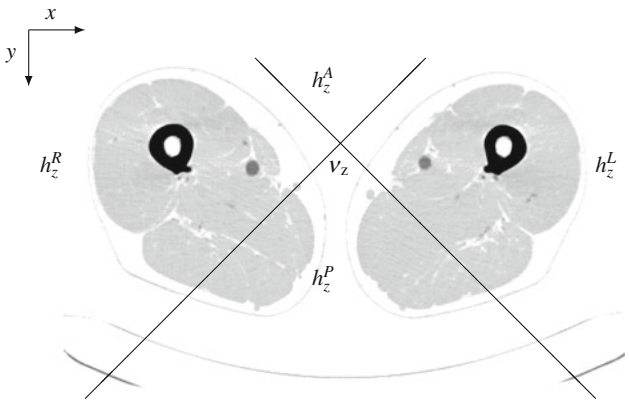

Fig. 2 *Right*, *left*, ante, and poste histogram bins centered at $\boldsymbol{v_z}$ overlaid over a negative of a CT slice at position z showing legs and table

With the AP/LR histograms we reformulate the leg detection as a search for slices, where ante-poste bone contributions vanish. We introduce scalars λ_z and a threshold to yield this:

$$0 \leq \lambda_z = \frac{h_z^A}{h_z^L + h_z^R} < 0.04 \tag{10}$$

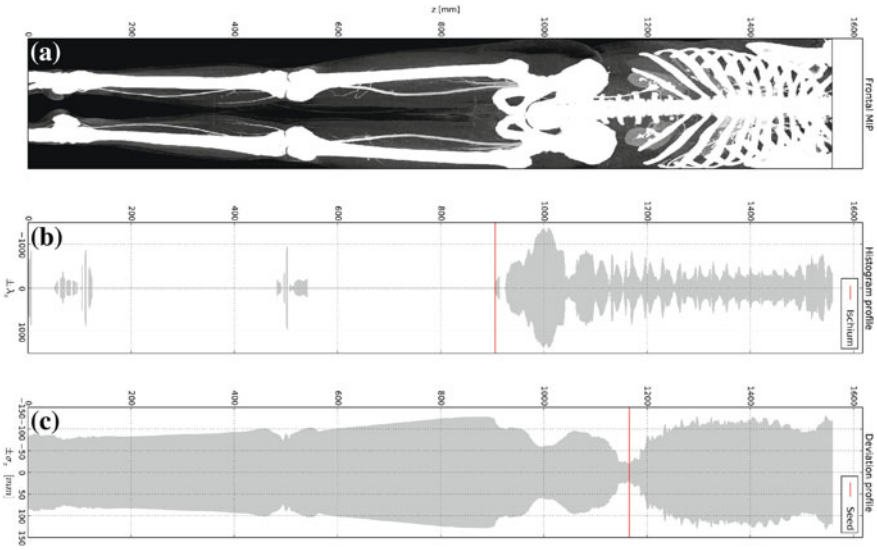

Fig. 3 An example CT scan in a frontal maximum intensity projection (MIP) (**a**), its histogram profile (**b**) and deviation profile (**c**)

Note that we exclude the posterior voxels h_z^P from (10) in order to ignore eventual contribution of a CT table.

3.3 Bone Profiles

In the previous section we have introduced two bone distribution descriptors, i.e. scalars λ_z and σ_z for every slice z. Next we aggregate them into two 1D arrays indexed by z and refer to as the *bone profiles*. A symmetric plot of bone profiles along the z-axis is shown in Fig. 3.

4 Applications to Spinal Column Localization

In this section we show how the bone profiles and the refined centers $\boldsymbol{v}_\mathbf{z}$ can be used to bound the spinal column and to identify a reliable initialization seed for subsequent computations.

4.1 Discarding the Slices up to the Ischium

When dealing with spines, leg slices should be taken out of consideration. We observed that the first occurrence of vanishing λ_z in the top-to-bottom order may correspond either to the bottom of sacrum or the bottom of the pelvis—the ischium (cf. Fig. 3b).

In order to have a security margin between the spine and the slices to drop we suggest to identify *ischium* slices. We identify them by the first 65 mm long segment of zeros in the histogram profile, i.e. a sequence longer than the average distance from ischium to bottom of sacrum.

4.2 Seeding a Spinal Canal Search

Algorithms using incremental/propagated search need to be initialized [5]. To obtain a reliable seed point near the spinal canal we consider the refined center $\boldsymbol{v}_{\mathbf{z}^\star}$ in a slice with minimal deviation σ_z (cf. Eq. 11). In this case no other bones except for vertebra contribute to the signatures and the point $\boldsymbol{v}_{\mathbf{z}^\star}$ yields an estimate of the spinal canal. Such slices are predominantly found either in the lumbar area between pelvis and the first rib (cf. Fig. 3c) or in the neck area.

$$\boldsymbol{v}_{\mathbf{z}^\star} \quad | \quad z^\star = \operatorname{argmin}\{\sigma_z\} \tag{11}$$

4.3 Bounding the Spinal Column

Machine learning methods need to compute a vector of features at every voxel. Reducing the amount of voxels to be classified to a minimum can therefore significantly speed up such algorithms. After the leg slices have been discarded we wish to further prune the space by setting coronal and sagittal bounding planes.

For healthy spines the previously found seed $\boldsymbol{v}_{z^\star}$ could be reused to set up a bounding box of a predefined size. Such an approach would, however, fail for scolioses and other spine curvature related disorders.

To deal with such cases we derive coronal and sagittal planes from the bounding box of the 80×140 mm windows (cf. Eq. 4) spanned symmetrically around a *subset* of centers $\boldsymbol{v}_z$. The refined centers $\boldsymbol{v}_z$ are first sorted by a *drift reliability* $\Delta_z = ||\boldsymbol{v}_z - \boldsymbol{\mu}_z||$ from the original centroids $\boldsymbol{\mu}_z$: the smaller the drift the more reliable the center. A fraction of sorted $\boldsymbol{v}_z$ involved in spanning bounding planes balances the tightness of bounding around spinal column and the data reduction. It is the last and the only free parameter in our method.

5 Results

Data and timing. All experiments were performed on two datasets. Dataset 1 contained 52 diversely cropped CT scans from three different vendors including 18 instances with legs up to abdomen and 34 scans with torso and/or head with neck. 29 out of the 52 were CTA scans containing contrast-enhanced vessels. The pixel sizes of the 512×512 axial images range from 0.26 to 1 mm, and the slice distances vary from 0.1 to 3 mm. The smallest and largest scan comprised of 103 and 5,966 slices respectively. Ground truth (ischium slice, spinal canal centers and vertebral body centers) was generated by a medical expert. For Dataset 2 we used the annotated spine CT database for benchmarking of vertebrae localization and identification with 242 spine-focused CT scans containing varying pathologies [2] (publicly available at http://spineweb.digitalimaginggroup.ca). Vertebral body centers were present for this dataset, ischium and spinal canal centers were added. The time performance has been measured single threaded on an Intel Core i7 2.6 GHz machine. The computation of both profiles took on average 1.5 s for Dataset 1 and 0.4 s for Dataset 2.

Discarding the leg slices. The ischium identification quality was measured in the 18 CT scans of Dataset 1 where legs were present. The mean error from the true ischium slice was 16.5 ± 13.2 mm which yields a sufficiently tall margin to the bottom of the spine, i.e. no cropping of the spine was observed. The amount of voxels has been reduced by a factor of 2.5 on average. For the remaining 276 scans without leg slices we counted the false positive occurrence of an ischium slice which was 0.

Seed detection. To assess the quality of the seeds $\boldsymbol{v}_{z^\star}$ we evaluated their distribution w.r.t. the associated spinal canal centers (cf. Fig. 4). Table 1 summarizes the seed detection results: the average x- and y-deviations of $\boldsymbol{v}_{z^\star}$ from spinal canal centers

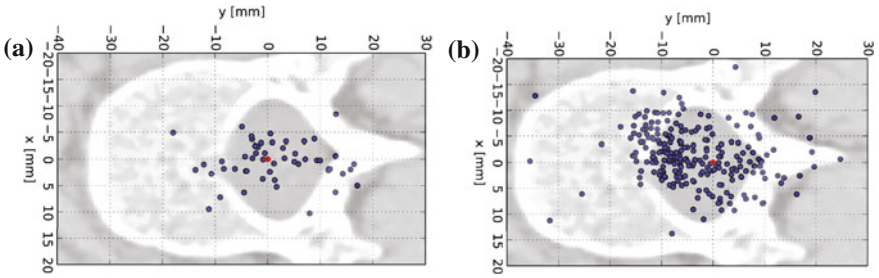

Fig. 4 Distribution of seeds vz for all instances of Dataset 1 (**a**) and Dataset 2 (**b**) relative to the ground truth spinal canal centers (located at origin) overlaid on an example slice (Color figure online)

Table 1 x- and y-deviations of detected spinal canal seeds from ground truth annotations and spinal canal identification rate (CIDR) for Dataset 1 and 2

	x-deviation (mm)	y-deviation (mm)	CIDR (%)
Dataset 1	2.8 ± 2.4	6.8 ± 5.0	75
Dataset 2	4.0 ± 3.3	8.0 ± 5.6	79

and the spinal canal identification rates (CIDR) for both datasets. The spinal canal identification rate indicates how often the seeds hit the spinal canal. For data where the seed was not detected within the spinal canal (25 % in Dataset 1 and 21 % in Dataset 2), it was placed either on the vertebral arch or on the vertebra body.

Bounding the spinal column. In order to evaluate the accuracy and utility of the spinal column bounding we analyzed how well spine ground truth was covered by the planes and to what amount data was reduced. As for the spine coverage, all expert-annotated vertebral body centers were contained within the coronal and sagittal bounding planes. To see the accuracy of the spine coverage of whole vertebrae we measured the minimum distances between the vertebral body center annotations and the left/right sagittal planes (Δx), the anterior coronal plane (Δy^A) and the posterior coronal plane (Δy^P) for each scan. Two distinct distances in the y-direction were necessary because of the sagittal asymmetry of vertebrae around the vertebral body center. In order to see if our bounding planes crop into vertebrae, we compared Δx to the half of the average midtransverse diameter ($l_x = 22.5\,\text{mm}$) of lumbar vertebrae [9], Δy^A to the half of the average inferior width of lumbar vertebral bodies ($l_y^A = 17\,\text{mm}$) and Δy^P to the average sagittal distance from vertebral body center to spinous process ($l_y^P = 65\,\text{mm}$) of lumbar vertebrae [1]. As higher percentages of reliable seeds stretched the bounding planes away from body centers (Sect. 4.3 and Fig. 6b), we investigated two percentage values, 40 and 90 %. For Dataset 1, 3 %(0 %) of the cases had slightly smaller Δx than l_x when using 40 %(90 %) of the most reliable seeds for bounding plane computation. l_x was not completely covered by 1.5 %(1 %) of the scans in Dataset 2 with 40 %(90 %) of the reliable seeds. These outliers were due to scoliosis in both datasets. Examples are shown in Fig. 5a, b.

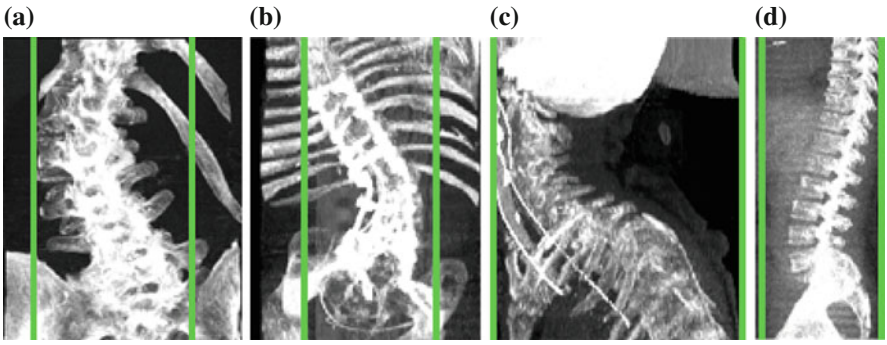

Fig. 5 MIPs of example data with bounding planes (*green lines*) generated by 40 % of the most reliable seeds. **a, b** Frontal MIPs of highly scoliotic spines. **c, d** Sagittal MIPs of cases where the input volume is cropped either on the anterior or on the posterior side near to vertebral body center annotations (Color figure online)

As for the Δy^A values, all of the cases had larger Δy^A values than l_y^A for both reliability percentages in Dataset 1. 1.6 % of the cases had smaller Δy^A values than l_y^A for both percentages in Dataset 2. These cases, however, were already too tightly cropped scans where expert annotations of vertebra centers were too close to the anterior volume border (see Fig. 5c) and our bounding planes did not further crop into the volume. The Δy^P values were in 5 % of the cases smaller than l_y^P for both reliability percentages in Dataset 1 which was due to tightly cropped input scans in the posterior direction with annotations close to the volume border (see Fig. 5d). In Dataset 2, 27 %(18 %) of the scans had smaller Δy^P values than l_y^P for 40 %(90 %) of the most reliable seeds. 4 CT scans out of the 27 % were slightly cropped. The remaining scans out of the 27 % and all cases out of the 18 % were handled correctly (full vertebra coverage) and either had annotations close to volume borders or contained vertebrae with smaller sagittal vertebral lengths than those of lumbar vertebrae (l_y^P).

To assess the utility of the spinal bounding planes, the average data reduction factor (DRF) was measured for both datasets after discarding the leg slices. It was 4.5(3.3) on average for Dataset 1 and 1.7(1.5) for Dataset 2 when considering 40 %(90 %) of the most reliable seeds for bounding plane construction.

6 Discussion and Conclusion

To the best of our knowledge, this is the first time a machine-learning-free method yields fast and reliable spine localization in such large (52 + 242) and diverse population of both normal and abnormal CT scans. Other machine-learning free related works (such as [3, 5]) achieve results with comparable accuracy to ours but they are slower and tested on much less data. Graf et al. [3] tested on 34 CT scans within an

average time of approximately 1 minute per scan and Klinder et al. [5] needs 7.6 s to find an initial spine seed in a collection of 64 CT scans.

Machine-learning-based approaches [2, 4] are not only competitive enough based on accuracy to the method we proposed, but they were tested with similarly high amounts of data and achieved good time performances. Glocker et al. [2] used 242 spine CT scans (our Dataset 2) for testing and Kelm et al. [4] needed only 11.5 s on average for intervertebral disk detection and labeling. In the following we therefore discuss our work from the perspective of machine learning methods.

First, setting-up a machine learning method requires ground truth annotations which is a labor intensive and error-prone task.

Second, the state-of-the-art machine learning algorithms (i.e., decision trees, random forests) combine image features in unintuitive and hard to interpret ways [6]. Potential algorithm failures are consequently impossible to explain and fix. On the contrary we proposed six parameters with a clear interpretation and justification (see Table 2): two Hounsfield unit thresholds for rough bone segmentation (Eq. 1), width and height of the refinement and spanning rectangles (Eq. 4, Sect. 4.3) deduced from morphometry of vertebrae and pelvis. The threshold to distinguish between legs and the rest of the slices λ_z (Eq. 10) was set empirically and validated after plurality of ground annotations was available (see Fig. 6a). The amount of reliable seeds involved in bounding the spine (Sect. 4.3) yields a parameter to control the bounding tightness (see Fig. 6b). Increasing this amount increases the overall minimal distance of the planes to the vertebral body center annotations, i.e. it decreases the DRF (see Sect. 5). This is the only free parameter of our method and users may tailor it to own applications: 10 % assures bounding of all vertebral body centers, 40 % bounds whole vertebrae but some cases may get slightly cropped, 90 % contains only a small set of vertebrae not covered completely and 100 % guaranties full coverage of the spinal column.

The six parameters make our method easy to implement. In contrast, parameters inherent to machine learning based methods (e.g., tree depth, number of trees) need to be found in a cross-validation scheme [6]. Such parameter setups pose difficulties for researchers who desire to re-implement the algorithms for own purposes.

Finally, learning methods need to compute high dimensional feature vectors at every voxel to yield a class or a regression value *after* the volume has been

Table 2 Summary of parameters and their values used within this work

Parameter name	Parameter value
Bone segmentation threshold, low	400 HU
Bone segmentation threshold, high	1050 HU
Spanning rectangle, width	80 mm
Spanning rectangle, height	140 mm
Ischium detection threshold	0.04
Amount of most reliable seeds	10 %,40 %, 90 %,100 %

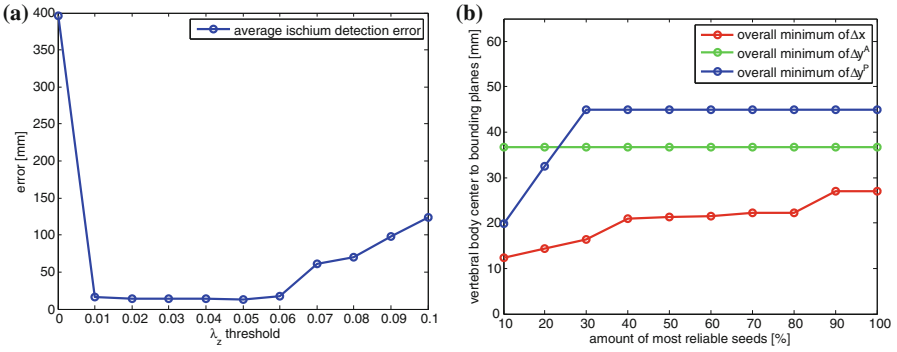

Fig. 6 **a** Empirical selection of parameter values for λ_z based on average ischium detection errors of Dataset 1. λ_z yields the least average error at 0.04. **b** Empirical parameter selection for the tightness of bounding planes based on Dataset 1. The values are minima of the 52 minimal distances (Δx, Δy^A and Δy^P).

reconstructed from the slices and loaded into memory. Moreover, data structures often involved in feature extraction (e.g., pyramids or integral volumes) demand additional memory. From perspective of machine learning methods, we thus introduced a computationally inexpensive tool to prune the search space. It is not clear how machine learning methods would perform with really big data. While the biggest of the currently available volumes of Dataset 2 is 256 MB, we were able to process a 3 GB scan in Dataset 1 within a memory required by one axial slice. We agree that the memory issue is irrelevant during research. It may, however, become a crucial argument for radiological departments who often pose tight memory constraints.

We have shown that our method delivers reliable results for spine localization within average times of 1.5 and 0.4 s for two sets of CT scans. It is faster compared to other methods such as [3] taking 1 min and [5] taking 7.6 s.

Possible weakness of our algorithm is related to the influence of artificial materials and artifacts with Hounsfield values similar to bones (e.g., contrast agents, implants, imaging artifacts). However, our results on CTA scans, pathological cases and scans from multiple vendors showed that our method is able to handle data with abnormalities and different imaging parameter settings correctly. A detailed study on this topic is left to future work.

Acknowledgments The competence center VRVis with the grant number 843272 is funded by BMVIT, BMWFW, and ZIT—The Technology Agency of the City of Vienna within the scope of COMET—Competence Centers for Excellent Technologies. The program COMET is managed by FFG. Thanks go to our project partner AGFA Healthcare for providing data and valuable input.

References

1. Gilad, I., Nissan, M.: Sagittal evaluation of elemental geometrical dimensions of human vertebrae. J. Anat. **143**, 115–120 (1985)
2. Glocker, B., Zikic, D., Konukoglu, E., Haynor, D.R., Criminisi, A.: Vertebrae localization in pathological spine CT via dense classification from sparse annotations. In: Medical Image Computing and Computer-Assisted Intervention, pp. 262–270. Springer, Berlin (2013)
3. Graf, F., Kriegel, H.P., Schubert, M., Strukelj, M., Cavallaro, A.: Fully automatic detection of the vertebrae in 2D CT images. In: SPIE Medical Imaging, Vol. 7962 (2011)
4. Kelm, B.M., Wels, M., Zhou, S.K., Seifert, S., Suehling, M., Zheng, Y., Comaniciu, D.: Spine detection in CT and MR using iterated marginal space learning. Med. Image Anal. **17**(8), 1283–1292 (2013)
5. Klinder, T., Ostermann, J., Ehm, M., Franz, A., Kneser, R., Lorenz, C.: Automated model-based vertebra detection, identification, and segmentation in CT images. Med. Image Anal. **13**(3), 471–482 (2009)
6. Kotsiantis, S.: Supervised machine learning: A review of classification techniques. Informatica **31**, 249–268 (2007)
7. Peng, Z., Zhong, J., Wee, W., Lee, J.h.: Automated vertebra detection and segmentation from the whole spine MR images. In: 27th Annual International Conference of the Engineering in Medicine and Biology Society, pp. 2527–2530 (2005)
8. Štern, D., Likar, B., Pernuš, F., Vrtovec, T.: 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 (2010)
9. Van Schaik, J., Verbiest, H., Van Schaik, F.: Morphometry of lower lumbar vertebrae as seen on CT scans: newly recognized characteristics. Am. J. Roentgenol. **145**(2), 327–335 (1985)
10. Yao, J., O'Connor, S.D., Summers, R.M.: Automated spinal column extraction and partitioning. In: 3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro, pp. 390–393 (2006)

Part V
Modeling

Area- and Angle-Preserving Parameterization for Vertebra Surface Mesh

Shoko Miyauchi, Ken'ichi Morooka, Tokuo Tsuji, Yasushi Miyagi, Takaichi Fukuda and Ryo Kurazume

Abstract This paper proposes a parameterization method of vertebra models by mapping them onto the parameterized surface of a torus. Our method is based on a modified Self-organizing Deformable Model (mSDM) [1], which is a deformable model guided by competitive learning and an energy minimization approach. Unlike conventional mapping methods, the mSDM finds the one-to-one mapping between arbitrary surface model and the target surface with the same genus as the model. At the same time, the mSDM can preserve geometrical properties of the original model before and after mapping. Moreover, users are able to control mapping positions of the feature vertices in the model. Using the mSDM, the proposed method maps the vertebra model onto a torus surface through an intermediate surface with the approximated shape of the vertebra. The use of the intermediate surface results in the stable mapping of the vertebra to a torus compared with the direct mapping from the model to the torus.

S. Miyauchi (✉) · K. Morooka · T. Tsuji · R. Kurazume
Graduate School of Information Science and Electrical Engineering,
Kyushu University, Fukuoka 819-0395, Japan
e-mail: miyauchi@irvs.ait.kyushu-u.ac.jp

K. Morooka
e-mail: morooka@ait.kyushu-u.ac.jp

T. Tsuji
e-mail: tsuji@ait.kyushu-u.ac.jp

R. Kurazume
e-mail: kurazume@ait.kyushu-u.ac.jp

Y. Miyagi
Department of Neurosurgery, Kaizuka Hospital, Fukuoka 812-0053, Japan
e-mail: yamiyagi@digital.med.kyushu-u.ac.jp

T. Fukuda
Graduate School of Medical Sciences, Kumamoto University,
Kumamoto 860-8556, Japan
e-mail: tfukuda@kumamoto-u.ac.jp

© Springer International Publishing Switzerland 2015

J. Yao et al. (eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, Lecture Notes in Computational Vision and Biomechanics 20, DOI 10.1007/978-3-319-14148-0_16

1 Introduction

Recent medical imaging devices provide high-resolution medical images. Computer aided diagnosis and therapy systems have been developed that uses the tissue models obtained by the images [2, 3]. To generate the model, the tissues are segmented from the images. The variability of the tissue shapes makes it difficult to segment the tissues automatically. One solution for this problem is to use Statistical Shape Model (SSM), which identifies the considerable natural variability and commonality of the tissue by statistical analysis. Because of the prior information about the tissue shape provided by SSM, the SSM-based techniques have obtained considerable success in the tissue segmentation [4]. Here, in this paper, we focus on the 3D mesh model of a tissue surface with triangular patches.

The fundamental process for building SSM is to establish the correspondence between the tissue models of different individuals. Generally, the models have different number of vertices. This increases the difficulty of directly finding the relationship between the models. One approach to the challenging correspondence problem is to map the models onto a parametric target surface with simple shapes such as a plane and a spherical surface. Since each vertex in the mapped model is assigned to the parametric coordinate on the target surface, parameterization allows to easily determine the model correspondences on the target surface. The mapping, called parameterization, is a useful tool for applications using models, including morphing, mesh-editing, remeshing, shape-analysis [5].

When applying the parameterization to vertebrae models with non-zero genus, there are two patterns of the target surface, that is, a target surface with the different or same genus of the vertebrae to be mapped. The former is a plane or a zero-genus surface while the latter includes a torus surface. Almost all previous parameterization methods have employed a different genus target surface of the vertebrae. Spherical harmonic [6] and spherical Wavelets [7] are often used for building SSM. In these methods, spherical mapping is applied to represent the model with polar coordinates. To achieve such mapping, the topology of the vertebra model needs to be changed. Becker et al. [8] solved the correspondence problem of the vertebrae models by cutting approximately the models along the same anatomical paths, and mapping them onto a plane. Like the method in [6–8], the vertebrae parameterization methods using zero genus surface need to change the topology of the model. The topological change process sometimes causes local distortions in the parameterization results. Therefore, the quality of the correspondence depends on the cut paths. On the other hand, the topological change process is not needed for the method using the target surface with the same genus as vertebrae. However, there were few methods which parameterize models by using the target surface with non-zero genus.

In this paper, we propose a new parameterization method of vertebrae by mapping it onto a torus. We have developed a modified Self-organizing Deformable Model (mSDM) [1], which finds the one-to-one mapping between arbitrary surface model and the target surface with the same genus as the model. At the same time, the mSDM

can both preserve geometrical properties of the original model before and after mapping, and control mapping positions of the feature vertices in the model. Using the mSDM, the proposed method maps the vertebra model onto a torus surface. It is simple to map the vertebra to a torus directly. However, this direct mapping is difficult since the deformation of the shape is large. Such large deformation sometimes results in the incomplete one-to-one mapping. To realize the stable mapping, we map the vertebra onto a torus through an intermediate surface, which has the approximated shape of the vertebra. Once a tissue model and its target surface are given, our mSDM is performed automatically with no manual intervention. Additionally, users can select feature vertices and their mapped positions.

2 Parameterization

The surface model $\mathscr{M}$ of a vertebra with one genus is represented by triangular patches. For each vertex $\boldsymbol{v}_i$ $(i = 1, 2, \ldots, N_{\mathscr{M}})$ in $\mathscr{M}$, its 1-ring region R_i consists of $N_R^{(i)}$ patches containing the vertex $\boldsymbol{v}_i$.

2.1 Modified Self-organizing Deformable Model

The mSDM is a deformable model whose shape is deformed by using a competitive learning and an energy minimization approach. Regrading a tissue model as the mSDM, the model is mapped onto a target surface by fitting the model to the target surface. The overview of the mSDM algorithm is as follows.

Step. M-1 A vertebra model $\mathscr{M}$ is deformed to fit to the target surface by the original SDM algorithm [9].

Step. M-2 The deformed model may contain foldovers. To realize the one-to-one correspondence, the foldover is removed by moving the vertices to within their 1-ring region.

Step. M-3 After the Step. M-2, the feature vertices move away from their corresponding points. Free-Form Deformation (FFD) [10] is applied to the mapped model to correct the positions of the feature vertices.

Step. M-4 To compare with the models, it is desirable that the mapped model preserves the original geometric properties as far as possible. Such deformation is achieved by an area- and/or angle-preserving mapping Φ. Practically, the model obtained after Step. M-3 is deformed by minimizing an objective function $E(\Phi)$ that is a weighted linear combination of area error term E_{area} and angle error term E_{angle}:

$$E(\Phi) = \mu E_{area} + (1 - \mu) E_{angle} \simeq \sum_i e(\Phi); \tag{1}$$

$$e(\Phi) = \mu\psi e_{area}(\Phi) + (1-\mu)e_{angle}(\Phi); \tag{2}$$

$$e_{area}(\Phi) = \sum_{p_k^{(i)} \in R_i} \left| \frac{\Phi(A_{p_k^{(i)}})}{\sum_{p_k^{(i)} \in R_i} \Phi(A_{p_k^{(i)}})} - \frac{A_{p_k^{(i)}}}{\sum_{p_k^{(i)} \in R_i} A_{p_k^{(i)}}} \right|; \tag{3}$$

$$e_{angle}(\Phi) = \sum_{p_k^{(i)} \in R_i} \sum_{d=0}^{3} |\Phi(\theta^d_{p_k^{(i)}}) - \theta^d_{p_k^{(i)}}|, \tag{4}$$

where ψ is a scaling factor to adjust the ranges of the two error terms. $A_{i,k}$ and $\theta^d_{i,k}$ are the area and one angle of the patch $p_{i,k} (k = 1, 2, \ldots, N_R^{(i)})$ included in the 1-ring region R_i of the vertex $\boldsymbol{v}_i$. $\Phi(A)$ and $\Phi(\theta)$ is the area and angle of the patch in $\Phi(\mathscr{M})$. The area errors in Eq. (3) is obtained by the total difference between the area ratio of a patch before and after the mapping. Here, the area ratio of the patch $p_k^{(i)}$ is defined as the ratio of the area $A_{p_k^{(i)}}$ of $p_k^{(i)}$ to the whole area of the 1-ring region R_i of $p_k^{(i)}$. In the same way, the angle error in Eq. (4) is obtained by the total difference between the angles of the patch $p_k^{(i)}$ before and after the mapping. Changing the weighting factor μ from 0 to 1, the mapping becomes from angle- to area-preserving mapping. The setting of μ determines the kinds of geometrical features preserved after the mapping. The discussion about the setting of μ will be found in [1]. From Eqs. (1) and (2), the minimization of Eq. (1) is replaced as the optimization problem of positioning the vertices in the 1-ring region by moving them repeatedly. The optimal mapping is found by applying a greedy algorithm with Eq. (2). Then, the vertices are not on the target surface completely. Therefore, they are mapped onto the nearest patch of the target surface after the movement.

Once a tissue model and its target surface are given, our mSDM is performed automatically with no users' manual intervention. Moreover, the mSDM framework controls the movement of several feature vertices of the model. Practically, when users manually specify the feature vertices and their corresponding locations on the target surface, the SDM algorithm maps the tissue model onto the target surface while moving the feature vertices toward their corresponding locations. This characteristic of the mSDM makes it easy to find the correspondence between tissue models by mapping the anatomical feature vertices of the tissue model onto their specific locations on the target surface.

2.2 Vertebra Parameterization Using mSDM

Our parameterization method uses an intermediate surface $\mathscr{I}$ whose shape is close to vertebrae shape. Here, the intermediate surface is determined as a surface model which has intermediate shape between the vertebra and the torus. Figure 1 shows

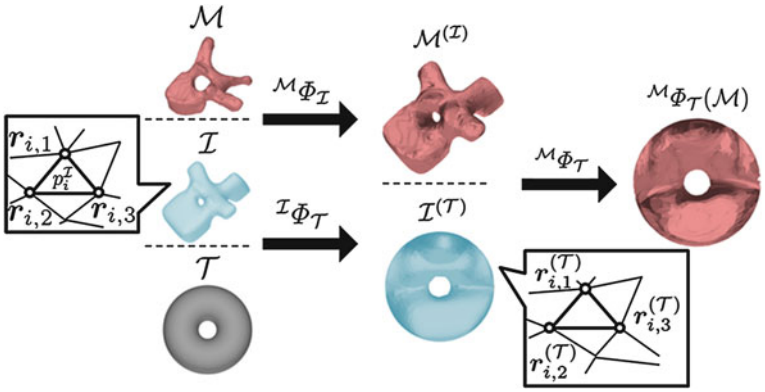

Fig. 1 Vertebra parameterization using mSDM

the overview of the mapping a vertebra $\mathscr{M}$ onto a torus $\mathscr{T}$. When the vertebra is mapped directly onto the torus whose shape is far from the vertebra, users need to tune the parameters in the mSDM algorithm and the scale of the torus. Especially, the latter tuning is to change the size of the torus and its hole according to the hole size of each individual vertebra. If the user fails the initial settings, we sometimes obtain the incomplete one-to-one mapping between the vertebra and the torus. Figure 2a, b show the examples of the vertebra models obtained by mapping the vertebra onto the torus directly. As shown in Fig. 2, the accuracy of the direct mapping depends on the parameters and the scale tuning heavily. On the other hand, through our experiments, the use of the intermediate surface results in the stable mapping of the vertebra models independent of the tuning for individual models. At first, we find one-to-one mapping ${}^{\mathscr{M}}\Phi_{\mathscr{I}}$ between $\mathscr{M}$ and $\mathscr{I}$. Similarly, one-to-one mapping ${}^{\mathscr{I}}\Phi_{\mathscr{T}}$ from $\mathscr{I}$ to $\mathscr{T}$ is found. Combining the two mappings ${}^{\mathscr{M}}\Phi_{\mathscr{I}}$ and ${}^{\mathscr{I}}\Phi_{\mathscr{T}}$, we obtain the mapping ${}^{\mathscr{M}}\Phi_{\mathscr{T}}$ from $\mathscr{M}$ to $\mathscr{T}$.

Our parameterization method consists of four steps.

Step. P-1 Select an intermediate surface $\mathscr{I}$.

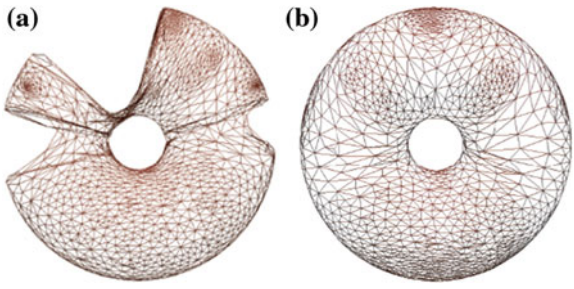

Fig. 2 Mapped vertebra models onto the torus by applying direct mapping **a** before and **b** after parameter and scale tuning of the SDM and the torus

Step. P-2 Find a mapping ${}^{\mathscr{I}}\Phi_{\mathscr{T}}$ by mapping from $\mathscr{I}$ to $\mathscr{T}$ with mSDM.
Step. P-3 Find a mapping ${}^{\mathscr{M}}\Phi_{\mathscr{I}}$ by mapping from $\mathscr{M}$ to $\mathscr{I}$ with mSDM.
Step. P-4 Obtain the direct mapping ${}^{\mathscr{M}}\Phi_{\mathscr{T}}$ by combining ${}^{\mathscr{M}}\Phi_{\mathscr{I}}$ and ${}^{\mathscr{I}}\Phi_{\mathscr{T}}$.

The following describes the detail of Step. P-4. Let us denote as $\mathscr{M}^{(\mathscr{I})} = {}^{\mathscr{M}}\Phi_{\mathscr{I}}(\mathscr{M})$ the model $\mathscr{M}$ mapped onto the intermediate surface $\mathscr{I}$. For each vertex $\boldsymbol{v}_i^{(\mathscr{I})} = {}^{\mathscr{M}}\Phi_{\mathscr{I}}(\boldsymbol{v}_i)$ in $\mathscr{M}^{(\mathscr{I})}$, we find a closest patch $p_i^{\mathscr{I}}$ in $\mathscr{I}$ to the vertex $\boldsymbol{v}_i^{(\mathscr{I})}$. When, $p_i^{\mathscr{I}}$ consists of three vertices $\boldsymbol{r}_{i,1}, \boldsymbol{r}_{i,2}, \boldsymbol{r}_{i,3}$ (Fig. 1), the coordinate of $\boldsymbol{v}_i^{(\mathscr{I})}$ is represented by

$$\boldsymbol{v}_i^{(\mathscr{I})} = \boldsymbol{r}_{i,1} + \alpha_1(\boldsymbol{r}_{i,2} - \boldsymbol{r}_{i,1}) + \alpha_2(\boldsymbol{r}_{i,3} - \boldsymbol{r}_{i,1}), \tag{5}$$

where α_1 and α_2 ($0 \leq \alpha_1, \alpha_2 \leq 1$) are real number parameters. Similarly, using the mapping ${}^{\mathscr{I}}\Phi_{\mathscr{T}}$, the vertices $\boldsymbol{r}_{i,*}$ in $\mathscr{I}$ are represented by $\boldsymbol{r}_{i,*}^{(\mathscr{T})} = {}^{\mathscr{I}}\Phi_{\mathscr{T}}(\boldsymbol{r}_{i,*})$. Therefore, the vertices $\hat{\boldsymbol{v}}_i^{(\mathscr{T})} = {}^{\mathscr{M}}\hat{\Phi}_{\mathscr{T}}(\boldsymbol{v}_i)$ of the model mapped on the torus are obtained by replacing the coordinates of vertices in $\mathscr{I}$ with that in $\mathscr{I}^{(\mathscr{T})}$ in Eq. (5):

$$\begin{aligned}\hat{\boldsymbol{v}}_i^{(\mathscr{T})} &\simeq {}^{\mathscr{I}}\Phi_{\mathscr{T}}{}^{\mathscr{M}}\Phi_{\mathscr{I}}(\boldsymbol{v}_i) \\ &= \boldsymbol{r}_{i,1}^{(\mathscr{T})} + \alpha_1(\boldsymbol{r}_{i,2}^{(\mathscr{T})} - \boldsymbol{r}_{i,1}^{(\mathscr{T})}) + \alpha_2(\boldsymbol{r}_{i,3}^{(\mathscr{T})} - \boldsymbol{r}_{i,1}^{(\mathscr{T})}).\end{aligned} \tag{6}$$

The vertices obtained from Eq. (6) may not be located on the torus surface completely. Therefore, we project the vertices and estimate the vertices position $\boldsymbol{v}_i^{(\mathscr{T})} = {}^{\mathscr{M}}\Phi_{\mathscr{T}}(\boldsymbol{v}_i)$ on the torus by

$$\boldsymbol{v}_i^{(\mathscr{T})} \simeq \boldsymbol{O}_i + B\frac{(\hat{\boldsymbol{v}}_i^{(\mathscr{T})} - \boldsymbol{O}_i)}{\|\hat{\boldsymbol{v}}_i^{(\mathscr{T})} - \boldsymbol{O}_i\|}; \tag{7}$$

$$\boldsymbol{O}_i = \frac{A}{\sqrt{\hat{\boldsymbol{v}}_{i,x}^{(\mathscr{T})2} + \hat{\boldsymbol{v}}_{i,y}^{(\mathscr{T})2}}}\begin{bmatrix}\hat{\boldsymbol{v}}_{i,x}^{(\mathscr{T})} \\ \hat{\boldsymbol{v}}_{i,y}^{(\mathscr{T})} \\ 0\end{bmatrix}, \tag{8}$$

where A and B are major and small radius in the torus. $\hat{\boldsymbol{v}}_{i,x}^{(\mathscr{T})}$ and $\hat{\boldsymbol{v}}_{i,y}^{(\mathscr{T})}$ are x and y coordinates of $\hat{\boldsymbol{v}}_i^{(\mathscr{T})}$.

2.3 Remeshing of the Original Tissue Model

One application using the proposed method is to remesh the original tissue models $\mathscr{M}$ by inversely mapping their target surface $\mathscr{S} = \{\mathscr{I}, \mathscr{T}\}$ onto the original model. By the back mapping, the original models are described by the mesh structure of their target surfaces.

First, each point $\boldsymbol{r}_j$ of the target surface is mapped onto the nearest patch $\delta^{(\mathscr{S})} = {}^{\mathscr{M}}\Phi_{\mathscr{S}}(\delta)$ of the mapped tissue model on the target surface. Here, δ is a patch of the original tissue model, and consists of $\boldsymbol{v}_{j,1}^{(\mathscr{S})}$, $\boldsymbol{v}_{j,2}^{(\mathscr{S})}$ and $\boldsymbol{v}_{j,3}^{(\mathscr{S})}$. The mapped point $\tilde{\boldsymbol{r}_j}$ of $\boldsymbol{r}_j$ is represented by

$$\tilde{\boldsymbol{r}_j} = \boldsymbol{v}_{j,1}^{(\mathscr{S})} + \beta_1(\boldsymbol{v}_{j,2}^{(\mathscr{S})} - \boldsymbol{v}_{j,1}^{(\mathscr{S})}) + \beta_2(\boldsymbol{v}_{j,3}^{(\mathscr{S})} - \boldsymbol{v}_{j,1}^{(\mathscr{S})}), \tag{9}$$

where β_1 and β_2 ($0 \leq \beta_1, \beta_2 \leq 1$) are real number parameters. The coordinate $\boldsymbol{r}_j^{\mathscr{M}} = {}^{\mathscr{M}}\Phi_{\mathscr{S}}^{-1}(\tilde{\boldsymbol{r}}_j)$ is calculated by using β_1, β_2 and the original model's vertices $\boldsymbol{v}_{j,1}, \boldsymbol{v}_{j,2}, \boldsymbol{v}_{j,3}$ which forming $\delta^{(\mathscr{S})}$:

$$\boldsymbol{r}_j^{\mathscr{M}} = \boldsymbol{v}_{j,1} + \beta_1(\boldsymbol{v}_{j,2} - \boldsymbol{v}_{j,1}) + \beta_2(\boldsymbol{v}_{j,3} - \boldsymbol{v}_{j,1}). \tag{10}$$

3 Experimental Results

3.1 Vertebra Model Parameterization

To verify the applicability of our proposed method, we made the experiments using six vertebrae models: three thoracic and three lumbar vertebrae. These models are generated from CT images of three healthy persons. Figure 3a shows the examples of the thoracic and lumbar models. The intermediate surfaces for thoracic and lumbar have their simplified shapes (Fig. 3b).

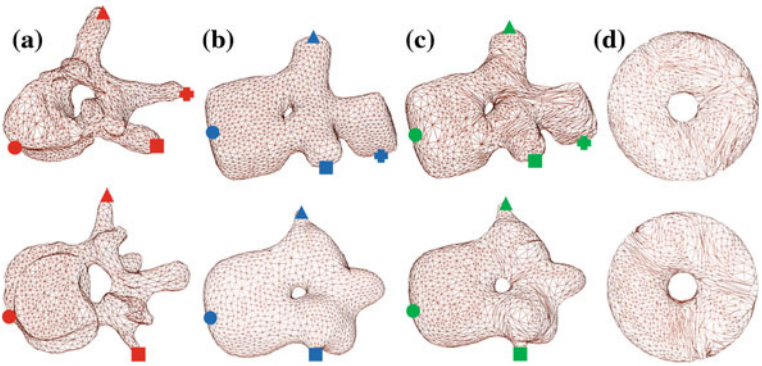

Fig. 3 Models of thoracic (*upper*) and lumbar (*lower*) in each parameterization step: **a** original models; **b** intermediate surfaces; **c** mapped models onto the intermediate surface; **d** final parameterization results. *Red* and *blue* points are feature vertices and their target points. *Green* points denote the feature vertices completely coincided with their corresponding points. Corresponding points have same shape (Color figure online)

The mSDM is applied to find the correspondence between each intermediate surface and the torus surface (in Step. P-2). In the same way, the vertebra model is mapped onto the intermediate surface by mSDM (in Step. P-3). In the experiment, we use four and three feature vertices of thoracic and lumbar. The feature vertices and their target points are manually selected from the vertebra, the intermediate surface and the torus. In Fig. 3a, b, red and blue points are the feature vertices and their target points of thoracic and lumbar. Corresponding points are represented by the marks with same shapes including a circle, a triangle, a square and a plus. The vertebra model mapped onto the intermediate surface $\mathscr{I}$ (Fig. 3c) has no foldovers. In addition, green points in Fig. 3c means the feature vertices completely coincided with their corresponding points. From the figures, feature vertices are located at their target positions correctly.

As shown in Fig. 3d, the mapping from the vertebra to the torus is achieved by combining the two mapped models described in Sect. 2.2. The final vertebrae models mapped on the torus have no foldovers, and feature vertices of the vertebra are located respectively at same positions. The mapping of the intermediate surface onto the torus is independent of the vertebra models to be mapped. Therefore, once the feature vertices are correctly mapped onto the target positions of the intermediate surface, the feature vertices of the models are mapped onto the same position of the torus. From the experimental results, our proposed method enables to represent vertebrae with two parameters without changing the topology of the vertebrae.

mSDM realizes three types of the geometrical feature preserving mappings: the area-, the angle-, and the area and angle-preserving mappings. Each mapping result is evaluated by its specific error distributions. The functions e_{area} (Eq. 3) and e_{angle} (Eq. 4) evaluate the quality of the area- or the angle- preserving mappings, respectively. When both the areas and angles of the patches in the mapped model are preserved simultaneously, the edge lengths of the patches are also preserved before and after mapping [11]. In order to evaluate the area- and angle-preserving mapping, the length error of the edge composed of two vertices $\boldsymbol{v}_s$ and $\boldsymbol{v}_t$ is defined by

$$e_{edge}(\boldsymbol{v}_s, \boldsymbol{v}_t, \Phi) = \left| \sqrt{\frac{\sum_{p \in \mathscr{M}} A_p}{\sum_{p \in \mathscr{M}} \Phi(A_p)}} |\Phi(\boldsymbol{v}_t) - \Phi(\boldsymbol{v}_s)| - |\boldsymbol{v}_t - \boldsymbol{v}_s| \right|, \tag{11}$$

where $\Phi(\boldsymbol{v})$ is the vertex obtained by applying the mapping Φ with the vertex $\boldsymbol{v}$. When the values of the error functions are close to zero, the mappings preserve their corresponding geometrical features.

Figure 4a–c show the distributions of e_{area}, e_{angle} and e_{edge} in the parameterization. The value of μ is set to real number ranging from 0 to 1. In this experiment, we used three values of μ, that is, $\mu = 0, 0.5$ and 1.0. In the figures, three lines correspond to the cases of $\mu : \mu = 0.0$(red line), 0.5(blue line), and 1.0(green line). Moreover, the black line in Fig. 4 shows the distribution of the mapped vertebra model $\mathscr{M}^-$ onto the torus by applying mSDM without Step. M-4. Compared with $\mathscr{M}^-$, our proposed method reduces the errors on areas, angles and edges. In the case of the area-preserving mapping ($\mu = 1$), the distribution of e_{area} has the highest

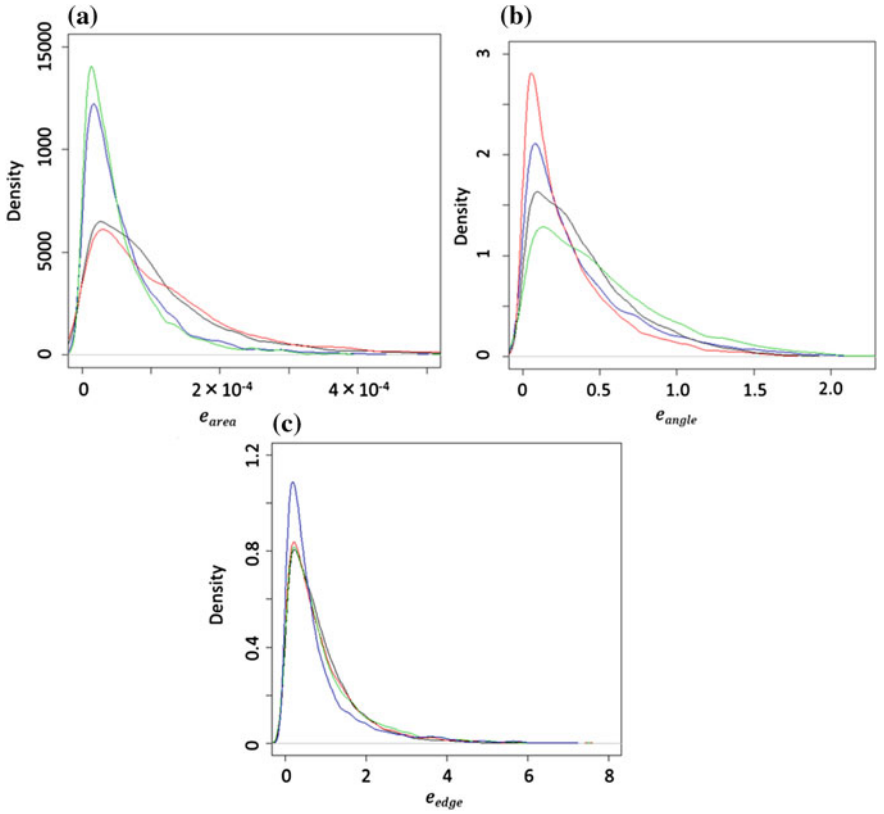

Fig. 4 Error distributions of the mapped models of a vertebra to the torus thorough the intermediate surface: **a** area; **b** angle; and **c** edge length. Each color line corresponds to each value of μ: *red* is 0.0; *blue* is 0.5; *green* is 1.0 (Color figure online)

peak at zero. The distribution of e_{angle} becomes narrow with decreasing the value of μ. In the case of the angle-preserving mapping ($\mu = 0$), the distribution of e_{angle} is steep, and has the single peak at zero. In the case of the area- and angle-preserving mapping ($\mu = 0.5$), the distribution of e_{edge} has the highest peak at zero. From these results, our method can achieve the preservation of the geometrical features.

To verify geometrical properties in each parameterization step, we calculate a rate of patches preserving geometrical properties (Table 1). The ratio means the percentage of the amount of the patches whose each geometrical error is less than a given threshold. When average angle and area of all patches in $\mathscr{M}$ is denoted as $\bar{\theta}$ and $\bar{S}$, and average edge length in $\mathscr{M}$ as $\bar{L}$, the thresholds of e_{angle} and e_{area} in the experiment are set to $0.3\bar{\theta}$, $0.3\bar{S}$ and $0.3\bar{L}$. Table 1 shows the average rate in the mapped models of thoracic ("T" in Table 1) and lumber ("L"). Comparing the model $\mathscr{M}^-$, our method maps the model onto the intermediate surface and the torus while keeping almost geometrical properties of the models. Moreover, the rate in Step. P-3

Table 1 A rate of patches preserving geometrical properties in mapping results (T: Thoracic, L: Lumber)

		before the process			$\mu = 0$	$\mu = 0.5$	$\mu = 1$
		e_{angle}	e_{area}	e_{edge}	e_{angle}	e_{edge}	e_{area}
Target	T	70.8	49.4	62.9	80.0	67.4	81.5
	L	59.1	61.1	64.8	75.4	79.8	90.2
P-3	T	57.2	56.4	65.8	81.5	80.5	90.0
Average	L	61.3	59.6	69.3	83.1	81.9	96.6
P-4	T	48.3	44.5	52.1	56.2	62.8	71.0
Average	L	44.2	49.4	53.8	58.0	65.8	70.9
Direct mapping	T	50.6	37.4	48.3	76.4	63.6	84.6
Average	L	48.7	42.2	51.8	74.3	64.1	80.0

The ratio means the percentage of the amount of the patches whose each geometrical error is less than a given threshold. When average angle and area of all patches in $\mathscr{M}$ is denoted as $\bar{\theta}$ and $\bar{S}$, and average edge length in $\mathscr{M}$ as $\bar{L}$, the thresholds of e_{angle} and e_{area} in the experiment are set to $0.3\bar{\theta}$, $0.3\bar{S}$ and $0.3\bar{L}$. The values in P-3 and P-4 are average of three thoracic and lumber vertebrae respectively

are higher than those in Step. P-4. This shows that the use of a target surface whose shape is close to $\mathscr{M}$ enables to keep the geometrical informations of the original vertebrae model. This is an advantage on using intermediate surface. On the other hand, vertebrae are dealt with in simpler parametric space by mapping the vertebrae onto a torus. Therefore, we had better select the two mapping results depending on the purpose of using the mapping. In the experiment, the error ratios are used to compare our mSDM with and without Step. M-4. The acceptance error ratios need to be defined to evaluate the mapping result objectively. The definition of the acceptance rate is one of our future works.

Moreover, we mapped the vertebrae onto the torus directly. The fourth row in Table 1 shows the average preserving rate of the directly mapped models of the vertebrae to the torus. The rates of the directly mapped vertebrae are sometimes higher than those of the mapped vertebrae through the intermediate surface. However, the direct mapping needs parameter tuning of the original SDM and scale tuning of the torus for each vertebra before the mapping. As shown in Fig. 2a, the vertebra model is not mapped onto the torus correctly without the scale tuning of the torus for the model. The improvement of the preserving ratio is one of our future works.

3.2 Vertebra Model Remeshing

Figure 5a–d show the remeshed models of the thoracic and the lumber obtained by the mapped models in Step. P-3 and 4 respectively. The remeshed models (Fig. 5a, c) with Step. P-3, using target surfaces whose shape is close to vertebrae

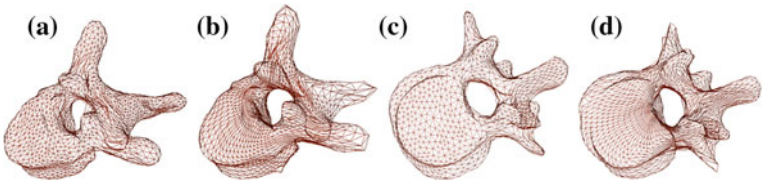

Fig. 5 Remeshing results of a thoracic and a lumber vertebra: **a**, **c** with Step. P-3 results; **b**, **d** with Step. P-4 results

Fig. 6 SSMs constructed by using the three remeshed thoracic models obtained in Step. P-3. The shape parameter of the primary ingredient is **a** −1.0 and **c** 1.0. **b** is the average of the three models

shape, have more uniform vertex density compared with the models (Fig. 5b, d) obtained by the mapping in Step. P-4.

These remeshed models enable to construct a vertebra SSM (Fig. 6) by the three remeshed thoracic models obtained in Step. P-3. Figure 6b show the average shape of the three models. Figure 6a, c show the SSM when the primary shape parameter is set to −1.0 and 1.0. Since models have smooth shape, our mapping enables to find the correspondence between the models correctly.

4 Conclusion

We proposed the method of parameterization vertebrae models onto torus. Our method can map the models without foldovers while both controlling feature vertices positions and keeping geometrical features. Vertebrae can be parametrized to torus which has only two parameter without changing topology, so the parameterization result has less distortion on the parameter space. Moreover, we confirmed that our method can apply to remeshing of tissue models.

Acknowledgments This work was supported by Foundation of Kyushu University, JSPS KAKENHI Grant Number 26560262, 24390345 and the Ministry of Health, Labour and Welfare(201024171A).

References

1. Miyauchi, S., Morooka, K., Miyagi, Y., Fukuda, T., Tsuji, T., Kurazume, R.: Tissue surface model mapping onto arbitrary target surface based on self-organizing deformable model. In: Workshop on Computer Vision at the 4th International Conference on Emerging Security Technologies, pp. 79–82 (2013)
2. Kobatake, H.: Future cad in multi-dimensional medical images:-project on multi-organ, multi-disease cad system. Comput. Med. Imaging Graph **31**(4), 258–266 (2007)
3. Morooka, K., Nakamoto, M., Sato, Y.: A survey on statistical modeling and machine learning approaches to computer assisted medical intervention: Intraoperative anatomy modeling and optimization of interventional procedures. Trans. Inf. Syst. **E96-D**(4), 784–797 (2013)
4. Heimann, T., Meinzer, H.P.: Statistical shape models for 3D medical image segmentation: a review. Med. Image Anal. **13**(4), 543–563 (2009)
5. Hormann, K., Levy, B., Sheffer, A.: Mesh parameterization:theory and practice. (2007)
6. Styner, M., Lieberman, J.A., Pantazis, D., Gerig, G.: Boundary and medial shape analysis of the hippocampus in schizophrenia. Med. Image Anal. **8**(3), 197–203 (2004)
7. Yu, P., Grant, P.E., Qi, Y., Han, X., Segonne, F., Pienaar, R., et al.: Cortical surface shape analysis based on spherical wavelets. IEEE Trans. Med. Imaging **26**(4), 582–597 (2007)
8. Becker, M., Kirschner, M., Fuhrmann, S., Wesarg, S.: Automatic construction of statistical shape models for vertebrae. In: Medical Image Computing and Computer-Assisted Intervention-MICCAI 2011, pp. 500–507. Springer, Heidelberg (2011)
9. Morooka, K. I., Nagahashi, H.: Self-organizing deformable model: a new method for fitting mesh model to given object surface. In: Advances in Visual Computing, pp. 151–158. Springer, Heidelberg (2005)
10. Sederberg, T. W., Parry, S. R.: Free-form deformation of solid geometric models. In: ACM Siggraph Computer Graphics, vol. 20, no. 4, pp. 151–160. ACM (1986)
11. Floater, M.S., Hormann, K.: Surface parameterization: a tutorial and survey. In: Advances in multiresolution for Geometric Modelling, pp. 157–186. Springer, Heidelberg (2005)

Contour Models for Descriptive Patient-Specific Neuro-Anatomical Modeling: Towards a Digital Brainstem Atlas

Nirmal Patel, Sharmin Sultana and Michel A. Audette

Abstract This paper describes on-going work on the transposition to digital format of 2D images of a printed atlas of the brainstem. In MRI-based anatomical modeling for neurosurgery planning and simulation, the complexity of the functional anatomy entails a digital atlas approach, rather than less descriptive voxel or surface-based approaches. However, there is an insufficiency of descriptive digital atlases, in particular of the brainstem. Our approach proceeds from a series of numbered, contour-based sketches coinciding with slices of the brainstem featuring both closed and open contours. The closed contours coincide with functionally relevant regions, in which case our objective is to fill in each corresponding label, which is analogous to painting numbered regions in a paint-by-numbers kit. The open contours typically coincide with cranial nerve tracts as well as symbols representing the medullary pyramids. This 2D phase is needed in order to produce densely labeled regions that can be stacked to produce 3D regions, as well as identifying embedded paths and outer attachment points of cranial nerves. In future work, the stacked labeled regions will be resampled and refined probabilistically, through active contour and surface modeling based on MRI T1, T2 and tractographic data. The relevance to spine modeling of this project is two-fold: (i) this atlas will fill a void left by the spine and brain segmentation communities, as no digital atlas of the brainstem exist, and (ii) this atlas is necessary to make explicit the attachment points of major nerves having both cranial and spinal origin, specifically nerves X and XI, as well all the attachment points of cranial nerves other than I and II.

N. Patel
Department of Electrical and Computer Engineering, Old Dominion University, Norfolk, USA
e-mail: npate012@odu.edu

S. Sultana · M.A. Audette (✉)
Department of Modeling, Simulation and Visualization Engineering, Old Dominion University, Norfolk, USA
e-mail: maudette@odu.edu

S. Sultana
e-mail: ssult003@odu.edu

© Springer International Publishing Switzerland 2015

J. Yao et al. (eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, Lecture Notes in Computational Vision and Biomechanics 20, DOI 10.1007/978-3-319-14148-0_17

1 Introduction

1.1 Neuroanatomy

The brainstem is the point of attachment of ten of the twelve pairs of cranial nerves. It is also the pathway for all fiber tracts passing up and down from peripheral nerves and spinal cord to the upper areas of the brain. It consists of three components: the medulla oblongata, the midbrain and the pons. The medulla oblongata acts as a relay station for motor tracts crossing between the spinal cord and the brain. It exerts control on the respiratory, motor and cardiac functions, as well as on several mechanisms of reflex activities such as swallowing and vomiting. The midbrain is the neural pathway of the cerebral hemispheres and contains auditory and visual reflex centers. The pons links different parts of the brain and is the relay station from the medulla to the higher cortical structures of the brain; it also contains the respiratory control center.

The relevance of the brainstem to the spinal anatomy is summarized in Fig. 1. The brainstem consists of the midbrain, pons and medulla oblongata; while it lies in the inferior part of the brain, *the brainstem is structurally continuous with the spinal cord*. Also, surgery planning and simulation of the spine require adequate modeling of critical tissues, the most important of which are spinal nerves, and in *the cervical portion of the spine, three of these major nerves, the glossopharyngeal (IX), vagus (X) and accessory nerves (XI) run near spinal nerves, while nerve XI has both a cranial and a spinal portion with roots in the brainstem and spine*, as depicted in Fig. 1.

Just as importantly, as seen in Fig. 2, the *brainstem is extremely relevant to modeling the neuroanatomy of the skull base.* In particular, the brainstem imbeds the nuclei and tracts of ten pairs of cranial nerves, nerves III to XII, and can be thought of as a twenty-legged spider. The olfactory (I) and ocular (II) nerves, emerge from the forebrain anteriorly and superiorly to the brainstem. So far, modeling that is conducive to the minimally supervised segmentation of cranial nerves has been neglected by the computer assisted surgery community.

1.2 Objectives and Clinical Motivation

The long-term clinical objective of this project is to provide a minimally supervised segmentation tool for patient-specific modeling of the brainstem. *There are a number of digital atlases of the brain in open source, but these atlases neglect the brainstem*: these atlases typically represent the whole brain, with little detail provided on the brainstem, or they concentrate on subcortical structures or the cerebellum, or alternately represent tractographic connectivity. The full-brain atlases include Harvard's Surgical Planning Laboratory (SPL) Atlas, atlases available with Oxford's FMRIB Software Library (FSL), which are listed on FSL's wiki (FSL), and the Freesurfer

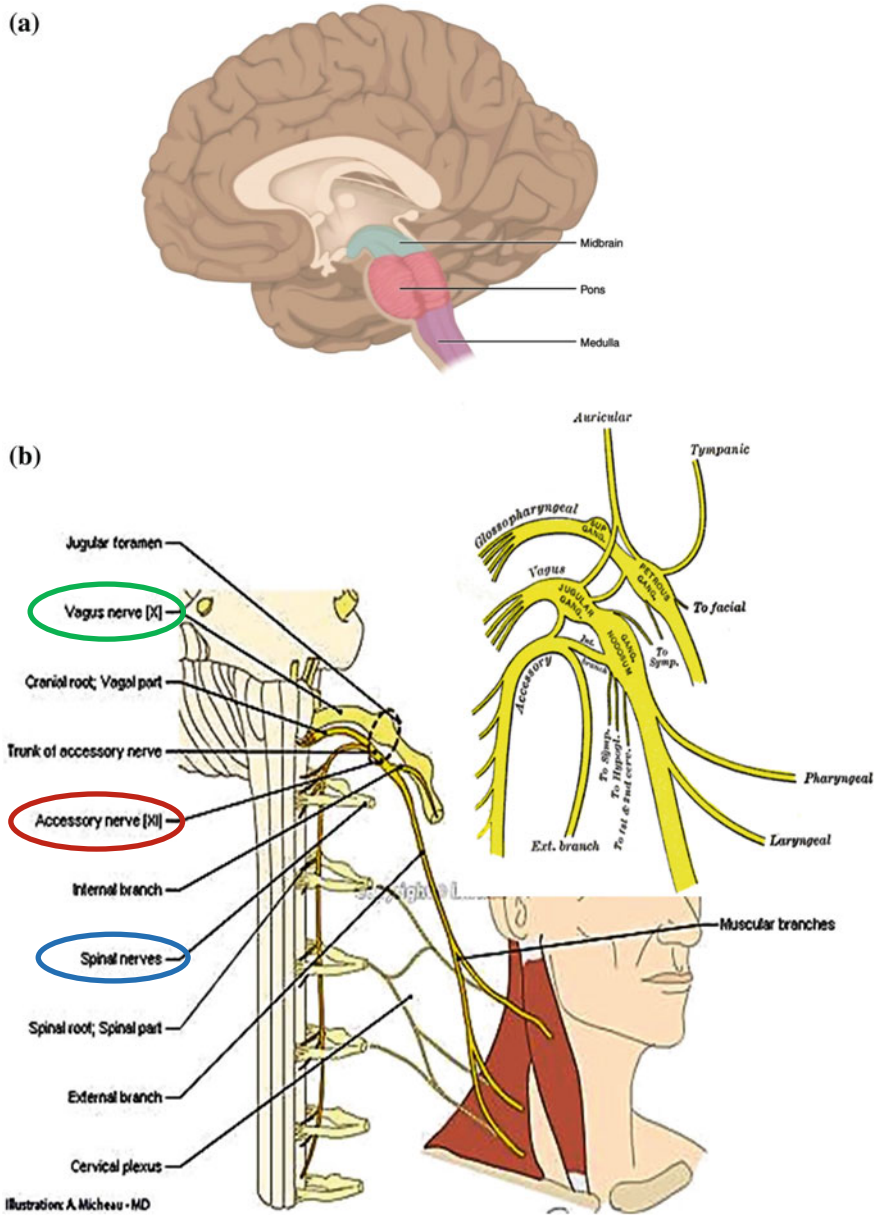

Fig. 1 Relevance of brainstem to spinal anatomy. **a** Structure of brainstem, with medulla oblongata adjacent and structurally similar to spinal cord [3]. **b** Proximity of cranial nerves X (*green*) and XI (*red*) to spinal nerves (*blue*) [17] inset: structure of nerves IX, X and XI [29]

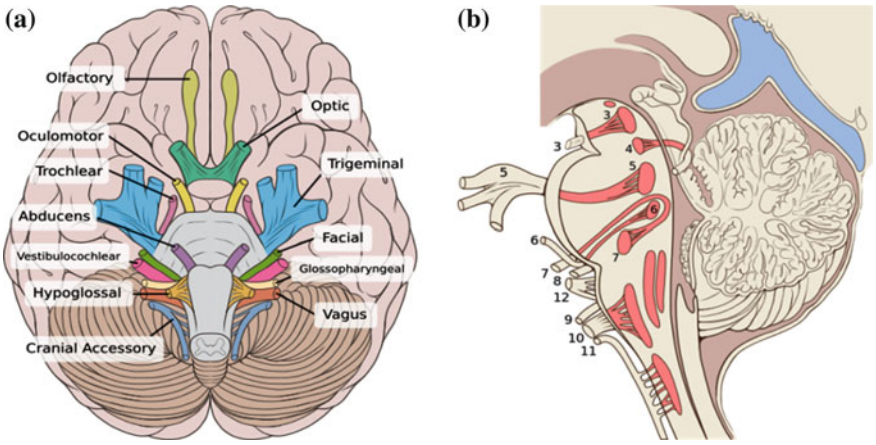

Fig. 2 Relevance to intracranial neuroanatomy: cranial nerves, with nerves III (oculomotor) to XII (hypoglossal) having nuclei within, and attachment points on, the brainstem; **a** inferior view; **b** sagittal view with cranial nerve nuclei and tracts inside the brainstem *shaded red* (reproduced from [9])

software package [11, 12]. To our knowledge, *none* of these digital atlases is dedicated to the brainstem, which relates to the paucity of *printed* atlases that center on the brainstem, with the exception of the Duvernoy [20]. The only competing option is the *Mai online atlas* [19], but its brainstem is low-resolution and not as descriptive as the Duvernoy, while the printed version of the Mai atlas [18] does not include the brainstem.

The main clinical motivation for this minimally supervised segmentation tool of the brainstem is two-fold, to enable the production of patient-specific models of the cranial and spinal nerves in the torso, such as depicted in Fig. 1, and to produce models of the neuroanatomy of the skull base, including the intracranial portion of the cranial nerves depicted in Fig. 2. These models will fulfill an unmet requirement of descriptive neuroanatomy for surgery planning and simulation for cervical spine procedures as well as skull base procedures such as radiotherapy of brainstem gliomas and resection of pituitary adenomas and acoustic neuromas.

In particular, we plan on *using probabilistic knowledge of the position of the tracts, or attachment points, of the cranial nerve in the brainstem to enable model-based identification of cranial nerves, in a manner analogous to the path of a clothesline being constrained by the position of the two poles at each end of it.* Each attachment point on the brainstem equates with the *inner pole* of the clothesline, while the position of the corresponding foramen will provide the other, *outer pole*. Shape statistics of these nerves will provide information on the tortuosity of the clothesline.

The clinical importance of detailed, patient-specific modeling of cranial and cervical spinal nerves is multi-faceted. First, *several complications have arisen in the past due to iatrogenic damage in the skull base and spine by surgery and radiotherapy*, which could have been prevented by more descriptive surgery planning as well as

surgery simulation that would have penalized gestures deleterious to these structures. Second, *topological variability in the form of anastomoses, or unusual connections, can occur between cervical spinal nerves and cranial nerves, and patient-specific descriptive modeling of the anatomy for planning and simulation must cope with this variability* [27].

Iatrogenic complications in neuro- and head-and-neck surgery, alluded to in the previous paragraph, still have an impact on patient outcome in a large number of cases. We argue strongly that improved surgery planning for experts and surgery simulation for residents would have probably improved patient outcome, at least in terms of morbidity and mortality statistics. Iatrogenic complications in neuro- and head-and-neck surgery include the following cases.

- The *spinal accessory nerve* (*SAN*), which is depicted in Fig. 1b, has been described as susceptible to injury in head-and-neck surgery by several authors [5, 31]. Estimates of SAN injury incidence in diagnostic lymph node biopsies of the posterior triangle of the neck are 3–8 %.
- In brain surgery procedures, a *number of complications have been documented in relation to iatrogenic damage or compression of cranial nerves.*
  - The *susceptibility to injury of the optical and oculomotor nerves in pituitary surgery* was reported in [7]. Amongst 939 neuro-surgeons surveyed, 179 reported post-operative visual loss in one or more patients.
  - The *vulnerability of cranial nerves was noted in the case of an acoustic neuroma patient* [23]. Her bone-embedded tumor was resected with an ultrasonic surgical aspirator at 80 % power. The patient experienced right-sided palsies of the 5th, 6th, 7th, and 12th cranial nerves.
  - As described in [4], *oropharyngeal dysphagia, or swallowing disorder,* can arise due to neurological dysfunction related to surgical complications of *posterior fossa and skull base surgery.*

As suggested above, *in order to improve patient outcome in neuro- and head-and-neck surgery, surgical planning and simulation must model cranial and cervical nerves explicitly and on a patient-specific basis, which must exploit both modeling that integrates shape priors,* i.e. *a probabilistic digital atlas, as well as tractographic imaging of the nerves.* MRI-based diffusion tensor imaging has been developed for cranial, spinal and peripheral nerves [16, 24, 25], although tractographic reconstruction of these images is still in its infancy, in contrast with tractography of the brain itself [2].

Moreover, to enable minimally supervised, highly descriptive, patient-specific identification of cranial and cervical nerves, we believe that a digital atlas for their most prevalent topology must first be developed, as a foundation for modeling variations such as those due to anastemoses [27]. One of the objectives of this brainstem atlas is to serve as cornerstone for the development of a cranial and cervical nerve atlas, which in turn will be used to stabilize the tractographic reconstruction of these nerves from diffusion imaging data, irrespective of their patient-specific topology.

Finally, treatment of brainstem tumors is still viewed as challenging, where surgical treatment is usually not an option, and chemotherapy is of limited utility. However, radiotherapy in the form of Gamma Knife radiosurgery (GKRS) has emerged as a promising treatment [22]. We suggest that *atlas-based functional parcellation of brainstem could have the same potentiating effect on brainstem radiosurgery as the whole-brain functional atlas, further potentiated by fMRI, has had on the sparing of functionally important, eloquent areas in neurosurgery of the brain these past few years.* As a result, it is not only the elaboration of cranial nerves that is clinically important in this project, but also the parcellation of the brainstem itself into functional regions.

1.3 Technical Motivation

This paper describes on-going work on the transposition to 3D digital format of a unique printed atlas of the brainstem [20]. This atlas, unlike the Mai atlas of the whole brain for example, does not offer intensity or color cues for regions that we can exploit to facilitate the segmentation. It only features closed contours with numbers imbedded in these regions, which are mostly without color. Moreover, the images also feature open contours that can penetrate some of these regions; a number of these contours coincide with embedded cranial nerves within the brainstem, while others represent symbols of medullary pyramids. At a minimum, these open contours must be purged from the closed contours to enable the labeling of the latter; however in many cases, we must model the embedded cranial nerves to make explicit the corresponding nucleus as well as the attachment point outside the brainstem.

Patient-specific anatomical modeling for neurosurgery planning and simulation entails medical image segmentation, which ideally is minimally supervised and describes the relationship between intensities in medical images, typically MRI or CT data, to anatomically or functionally relevant tissues or structures. Minimally supervised medical image segmentation generally falls under three categories: voxel, boundary and atlas-based, with some overlap being possible. In general, an atlas-based approach enables a functional or anatomical representation of tissues that is far more descriptive than is feasible with the other two methodologies.

Voxel-based approaches involve algorithms based on local properties at each point or voxel. Voxel-based methods tend to be limited to identifying a small number of tissues that they can differentiate, and to be vulnerable to bleeding effects in low contrast, given their paucity of model-based safeguards. The boundary-based approach exploits an active model featuring a trade-off between an internal representation of the tissue boundary, which typically favors a smooth result as well as probabilistic shape priors, and external influences based on intensity gradients or region homogeneity.

An atlas-based approach differs from the other two in that it favors a process inherently based on non-rigid registration, while providing richly descriptive functional information, namely a digital map of tissues or functional structures. This functional

detail of an atlas approach can outstrip the other two, often making it the most clinically relevant method of all, particularly for applications that target functionally complex anatomy such as neurosurgery [21]. *Despite such advances, there is an insufficiency of descriptive digital atlases, particularly of the brainstem, while the Duvernoy printed atlas can be leveraged to address this need.*

Meanwhile, with the advent of active multi-surface models [15], it is now feasible to reconcile the rich functional and anatomical description of digital atlases with the advantages of a boundary-based approach: model-based continuity, probabilistic shape priors [28], a robust multi-resolution non-rigid registration framework, and static collision handling [15] of active surface models. The latter type of integration, i.e. an active multi-surface functional atlas, which we are also developing for deep-brain stimulation applications [1], represents the long-term technical objective of this brainstem modeling project.

2 Purpose

As described above, there is a dire clinical need for a digital atlas of the brainstem, from the standpoint of surgery planning and simulation of the spine and brain. In general, neuroanatomical modeling for the planning of neurosurgery, which includes radiosurgery, as well as for neurosurgery simulation, does not depict the brainstem descriptively, given the traditional slice-by-slice, interactive segmentation approach that relies on manual voxel-based thresholding, which tends to be limited in the detail that is achievable. As a result, existing methods produce anatomical models that are extremely terse about the functional components of the brainstem.

This paper presents preliminary results of the 2D processing stage of the brainstem digital atlas development, which at first glance integrates a number of techniques that appear unrelated, some of which may be incremental improvements on published techniques. However, the clinical impetus for this digital atlas, described in the previous section, is irrefutable; moreover, the printed atlas that we are using imposes limits on our design options. Our newer 2D methods as well as the 3D methodology that will build on the resulting labeled images and explicit models of embedded nerves, leading up to a descriptive 3D atlas, will feature significant improvements, including a representation based on active multi-surface models [15] and 3D active contour models. The final stage that produces a probabilistic brainstem atlas that will also require significant innovations.

3 Methods

Our proximal objective is to semi-automatically process a series of input scanned images from a textbook [20], in order to generate continuously labeled digital images that are suitable for stacking and resampling, to ultimately produce a digital volumetric atlas, featuring functional regions and embedded nerve tracts, suitable for 3D registration with a given patient's brainstem appearing in MRI data.

3.1 Region Labeling by Paint-by-Numbers Level Sets

The printed atlas images typically have labeled regions where the labels are numbers centered within them. A rectangle that covers an entire label while being inside the boundaries of the corresponding region initializes the contour of the region. This rectangular contour is used to initialize an outwardly moving level sets contour model, whose objective is to propagate a user-supplied label, ostensibly the same number, to boundaries of that region.

In order to reduce irregularities within regions, Anisotropic diffusion method was utilized. Anisotropic diffusion tries to preserve the edges as opposed to isotropic diffusion methods which blurs entire image indiscriminately. The Anisotropic function used was modified curvature diffusion equation which is:

$$f_t = |\nabla f| \nabla \cdot c(|\nabla f|) \frac{\nabla f}{|\nabla f|}$$

where is the image and is the conductance function (Whitaker and Xue).

Both region-based and gradient-based level set methods were tested. The Chan-Vese (Chan TF) region-based model would typically bleed across the boundary of the region, which prompted us to adopt a gradient-based method proposed by Caselles [6], was found to be more reliable for most regions. The level set using Geodesic Active Contours evolves according to the following equation:

$$C_t = g(I)(c + k)\,\vec{\mathcal{N}} + (\nabla g \cdot \vec{\mathcal{N}})\vec{\mathcal{N}}$$

where g(I) is the stopping function, k is the curvature, and $\vec{\mathcal{N}}$ is curve normal.

The initial level set function was generated using the Fast Marching method [26], in order to produce a computationally efficient label propagation process. The Fast Marching method also requires an initialization, coinciding with the rectangular initialization contour at time t = 0 from which the level set model evolves. Our labeling program sets the user-provided rectangular contour as a seed. After the zero level set stabilizes at the region boundary, its propagation domain is colored according to the user-supplied label. As a result, the pixels outside the level set are white whereas the pixels inside the final level set result have the user-supplied value.

Finally, the atlases make the assumption of a vertical axis of symmetry. In our labeling interface, the user supplies a few points that determine this axis. These points determine a segment by regression, and in the event of a small difference rotation with respected to the vertical axis, this rotation is corrected to ensure that the axis of symmetry coincides with the y-axis. A mirror image is then generated by flipping the rotated image about the y-axis. To generate the final image, the mirrored and rotated images are arranged side-by-side such that the mirror image is on the negative side of the x-axis and the rotated image is on the positive side of the x-axis.

3.2 Nerve Identification by Minimal Path and 1-Simplex Contour Modeling

The scanned atlases may contain curvilinear features, typically coinciding with embedded nerves, which can intersect regions. These must be dealt with in two ways. First, these features need to be purged from closed regions to keep the level set evolution in these regions from being prematurely interrupted. Second, many of these curves represent embedded nerves that must be modeled explicitly, which we have begun implementing with 1-simplex active contour models, adapted from 2-simplex active surface modeling developed by Delingette [10].

To segment the nerves for exclusion from regions, the user supplies a start point, an end point, and optionally midway points, which are used to determine a *Minimal Path* (*MP*) [8] that extends from the start point to the end point while crossing the mid points. In order to extract such a path, the Fast Marching minimal path extraction method was used. This method requires a speed function of range between 0 and 1. Ideally, the speed would be close to 0 outside a line and close to 1 inside the line. *A nerve-centered MP is achieved by generating the vesselness of the image* [13], *and restricting the minimal path to high-vesselness pixels.*

For modeling the nerves explicitly, which will be needed for 3D segmentation, we are currently also implementing a 1-simplex model, which is the contour version of Delingette's 2-simplex surface model [10]. An *n*-simplex is a mesh where every vertex is linked by an edge to $n + 1$ neighboring vertices: a 2-simplex implies 3 neighbors, which produces a surface mesh geometrically dual to a triangulated surface mesh, while a 1-simplex produces a contour model where each vertex is linked to 2-neighbors. This implementation is still underway, with preliminary results depicted in the next section. *Our objective for this 1-simplex model is to produce explicit contours in 2D, to which we will assign z-coordinates according to corresponding image slices, and refine these 1-simplex imbedded nerve models with high-resolution histological data.*

4 Results

Figure 3 shows the 2D slices of the printed brainstem atlas and corresponding segmented images. The scanned images were cropped and rotated to correct the orientation, as well as flipped about the y-axis to exploit symmetry. Figure 4 illustrates the elaboration of minimal path and simplex-based contour modeling of the embedded nerves within the brainstem.

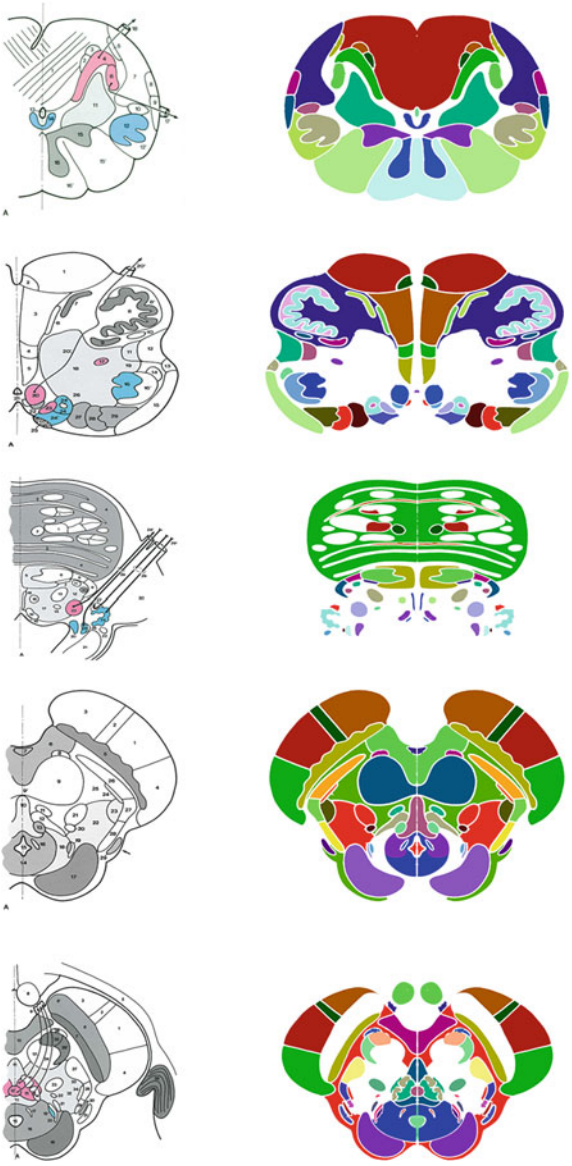

Fig. 3 Paint-by-number level sets results—*left* five input scanned images (after rotation correction), and *right* output of labeling and symmetry application

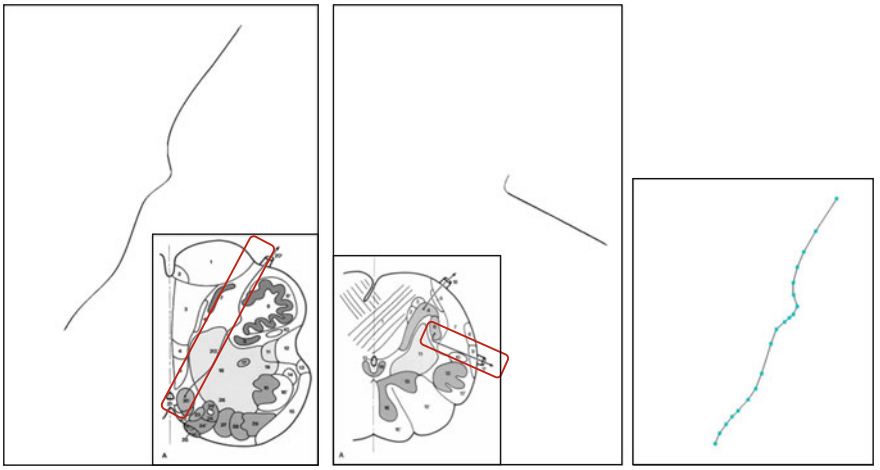

Fig. 4 Modeling of cranial nerves: **a**, **b** Minimal Path-based models of cranial nerves in two scanned images; **c** 1-simplex contour mesh of first MP in (**a**)

5 Discussion

Segmenting a scanned brainstem atlas represents a challenge and doing so without any user supervision is infeasible. Therefore, a limited amount of user assistance is a reasonable trade-off, even if some the methods appear unrelated. Broadly speaking, level set contour models lead to a usable set of user-driven methods for transposing printed labeled regions to digitally labeled regions, in a manner that evokes paint-by-numbers in closed regions. Similarly, the Minimal path and 1-simplex both can be exploited to identify a curvilinear structure such as a nerve within a printed atlas.

Our justification for adding the 1-simplex to the Minimal Path is that an explicit representation of such a structure has advantages over an implicit one, such as the relative ease with which one can generalize a 2D contour, with knowledge of the z-coordinate, into a fully 3D contour, and immunity to bleeding effects. Also, the simplex model has used with explicit static collision detection [15], which will be useful for modeling several nerves and blood vessels proximate to each other.

Finally, we recognize that the validation of this methodology will pose a challenge in itself. After all, how one quantifies success in transposing a printed atlas to a digital atlas is still an open question, despite a number of research efforts in this area.

6 Conclusion

This paper presented a method based on level set contour models and the Minimal Path for converting a printed atlas into densely labeled slices of a digital volumetric atlas, with a specific application to the brainstem. We have begun work on a 1-simplex-based explicit contour model as well. Our 2D algorithms are undergoing further refinements, following which these slices will be interpolated and resampled in 3D. One needed refinement is growing the regions until they abut one another.

References

1. Audette, M.A. et al.: Towards digital atlas and MRA-guided, intraoperative MRI-registered robotic deep brain therapy. IEEE EMBC, San Diego (2012)
2. Basser, P.: In vivo fiber tractography using DT-MRI data. Magn. Resonance Med. **44**, 625–632 (2000)
3. Brainstem Wiki: n.d. http://en.wikipedia.org/wiki/Brainstem
4. Buchholz, D.W.: Oropharyngeal dysphagia due to iatrogenic neurological dysfunction. Dysphagia **10**(4), 248–254 (1995)
5. Cappiello, J., et al.: The spinal accessory nerve in head and neck surgery. Curr. Opin. Otolaryngol. Head Neck Surg. **15**(2), 107–111 (2007)
6. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic active contours. Int. J. Comp. Vis. **22**, 61–79 (1997)
7. Ciric, I., et al.: Complications of transsphenoidal surgery: results of a national survey, review of the literature, and personal experience. Neurosurgery **40**(2), 225–236 (1997). discussion 236–7
8. Cohen, L.D., Kimmel, R.: Global minimum for active contour models: a minimal path approach. Int. J. Comput. Vis. **24**(1), 57–78 (1997)
9. Cranial nerves Wiki. n.d. http://en.wikipedia.org/wiki/Cranial_nerves
10. Delingette, H.: General object reconstruction based on simplex meshes. Int. J. Comput. Vis. **32**(2), 111–146 (1999)
11. Desikan, R.S., et al.: An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage **31**(3), 968–980 (2006). Epub 2006 Mar 10
12. Fischl, B., et al.: Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron **33**, 341–355 (2002)
13. Frangi, R.F., et al.: Multiscale vessel enhancement filtering. Med. Image Comput. Comput.-Assist. Interv. **1496**, 130–137 (1998)
14. FSL Atlases: Templates and atlases included with FSL. n.d. http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases
15. Gilles, B.: Musculoskeletal MRI segmentation using multi-resolution simplex meshes with medial representations. Med. Image Anal. **14**(3), 291–302 (2010)
16. Hiltunen, J., et al.: Diffusion tensor imaging and tractography of distal peripheral nerves at 3 T. Clin. Neurophysiol. **116**(10), 2315–2323 (2005)
17. IMAIOS: IMAIOS anatomy of cranial nerves. n.d. http://www.imaios.com/en/e-Anatomy/Head-and-Neck/Cranial-nerves-diagrams
18. Mai, J.K., Paxinos, G., Voss, T.: Atlas of the Human Brain, 3rd edn. Academic Press, San Diego (2007)
19. Mai, J.K., et al.: Atlas of the Human Brain—Brainstem. n.d. http://www.thehumanbrain.info/head_brain/brainstem_flash.php

20. Naidich, T.P., Duvernoy, H.M., et al.: Duvernoy's Atlas of the Human Brain Stem and Cerebellum: High-Field MRI, Surface Anatomy, Internal Structure, Vascularization and 3 D Sectional Anatomy. Springer (2009)
21. Nowinski, W.L., et al.: Atlas-based system for functional neurosurgery. In: Proceedings of SPIE Medical Imaging. SPIE, Newport Beach, pp. 92–103 (1997)
22. Peterson, H.E., et al.: Gamma knife treatment of brainstem metastases. Int. J. Mol. Sci. **15**(6), 9748–9761 (2014). doi:10.3390/ijms15069748
23. Ridderheim, P.A.: Indirect injury to cranial nerves after surgery with Cavitron ultra-sonic surgical aspirator (CUSA). Acta Neurochir. **89**(1–2), 84–86 (1987)
24. Roundy, N., et al.: Preoperative identification of the facial nerve in patients with large cerebellopontine angle tumors using high-density diffusion tensor imaging. J. Neurosurg. **116**(4), 697–702 (2012)
25. Schmidt, E.J., et al.: Wide-band steady state free precession with small diffusion gradients for spine imaging: application to superior nerve visualization. ISMRM, p. 448 (2010)
26. Sethian, J.A.: A fast marching level set method for monotonically advancing fronts. vol. 93, pp. 1591–1595 (1996)
27. Shoja, M.: A comprehensive review with potential significance during skull base and neck operations, part II: glossopharyngeal, vagus, accessory, and hypoglos-sal nerves and cervical spinal nerves 1–4. Clin. Anat. **27**(1), 131–144 (2014)
28. Tejos, C., et al.: Simplex mesh diffusion snakes: integrating 2D and 3D deformable models and statistical shape knowledge in a variational framework. Int. J. Comput. Vis. **85**(1), 19–34 (2009)
29. Vagus Nerve Wiki: Wikipedia: Vagus Nerve. n.d. http://en.wikipedia.org/wiki/Vagus_nerve
30. Whitaker, R.T., et al.: Variable-conductance, level-set curvature for image denoising. In: IEEE International Conference on Image Processing, pp. 142–145 (2001)
31. Wiater, J.M., Bigliani, L.U.: Spinal accessory nerve injury. Clin. Orthop. Relat. Res. **368**, 5–16 (1999)

Part VI
Segmentation Challenge

Atlas-Based Segmentation of the Thoracic and Lumbar Vertebrae

Daniel Forsberg

Abstract Segmentation of the vertebrae in the spine is of relevance to many medical applications. To this end, the 2nd MICCAI workshop on Computational Methods and Clinical Applications for Spine Imaging organized a segmentation challenge. This paper briefly presents one of the participating methods along with achieved results. The employed method is based upon atlas-based segmentation, where a number of atlases of the spine are registered to the target data set. The labels of the deformed atlases are combined using label fusion to obtain the final segmentation of the target data set. An average DICE score of 0.94 ± 0.03 was achieved on the training data set.

1 Introduction

The spinal column forms an important support structure in the human body and mainly consists of the vertebral bones. As such, the vertebrae form an important part of the diagnosis, treatment planing and the understanding of various conditions affecting the spine. Thus, an accurate segmentation of the vertebrae is of relevance in several applications. The segmentation of the vertebrae is challenging, mainly due to shape variation and neighboring structures of similar intensity (e.g. other vertebrae, other bones and/or other tissues). To this end, the 2nd MICCAI workshop on Computational Methods and Clinical Applications for Spine Imaging has organized a segmentation challenge. This paper briefly describes one of the contributing methods and the achieved results.

D. Forsberg (✉)
Center for Medical Image Science and Visualization (CMIV), Linköping University, Linköping University Hospital, SE-581 85 Linköping, Sweden
e-mail: daniel.forsberg@sectra.se

Sectra, Teknikringen 20, SE-583 30 Linköping, Sweden

© Springer International Publishing Switzerland 2015

J. Yao et al. (eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, Lecture Notes in Computational Vision and Biomechanics 20, DOI 10.1007/978-3-319-14148-0_18

2 Methods

The method used in this work for vertebra segmentation is inspired and to a large extent based upon the work presented in [2, 3], although some components have been changed and others have been added. This has been done to improve the performance but also since the work in [2, 3] was targeted at scoliotic spines. The most notable differences are the use of multiple gray-level atlases instead of a single binary model in the registration step, and the subsequent use of label fusion. The employed method consists of a preprocessing step, where an approximate position and rotation (pose) of each vertebra in the spines of both the target data set and the atlases are estimated. The preprocessing is followed by a registration step, where each atlas is registered to the target data set. The labels of the registered atlases are merged to a single label volume using label fusion to form the segmentation of the spine vertebrae in the target data set.

2.1 Preprocessing

The preprocessing consists of the following sub-steps:

1. **Spinal canal tracking**—Seed points for the spinal canal are detected using the Hough transform on a thresholded axial image in the middle of the image volume. A growing and moving circle is used to detect the center of the spinal canal, and where the growing and moving circle process is repeated for each image as the spinal canal is tracked in both the cranial and the caudal direction.
2. **Disc detection**—Given that the vertebrae are located anterior to the spinal canal, an intensity profile, running through the vertebrae, is sampled. The distinctive pattern of the intensity profile can be used to detect the position of the discs.
3. **Initial vertebral rotation estimation**—In an image slice located between the detected discs, an initial axial vertebral rotation is estimated based upon minimizing an error measure for assessing the lateral symmetry between two halves of the image.
4. **Vertebra pose estimation**—The two previous steps provide an initial estimate of the position and the orientation of each vertebra. To improve the pose estimate, an error measure is defined with six parameters $[x, y, z]$ and $[\theta_X, \theta_Y, \theta_Z]$, defining the vertebra center-point and the vertebral orientation. The error measure is defined to assess the symmetry across various half-planes. The optimal parameters are found using Powell's method.[1] This step is similar to the method used in [5] for estimating the center-point and the orientation of the vertebrae in the spine.

[1] E. Farhi, The iFit data analysis library, ILL, Computing for Science (2011–2012) http://ifit.mccode.org.

The estimated pose is used for establishing an initial alignment between the vertebrae of the atlases and the vertebrae of the spine to segment.

2.2 Atlas-Based Registration

Given an initial alignment between the vertebrae of the target data set and an atlas, the vertebrae are then registered in groups of five, starting from the caudal end of the spine, i.e. L5-L1, L1-T9, T9-T5, T5-T1. A sub-volume is extracted from both the target data set and the atlas at hand, containing the vertebrae to be registered along with any surrounding vertebrae. After the initial alignment, a non-rigid registration step is applied, minimizing the local phase-difference [4]. For computational performance, an implementation of the non-rigid registration method on graphics processing units was used [1]. The finally computed transformation is used to deform the atlas onto the target data set, and, thus, the labels of the deformed atlas provide a segmentation of the vertebrae in the target data set.

2.3 Label Fusion

The final step is to merge the labels of the different deformed atlases into a single atlas volume. In this case, a straight forward majority voting has been employed for label fusion.

3 Data

The data used for evaluation consists of ten CT data sets acquired during daily clinical routine work in a trauma center, and is provided as training data for the 2nd MICCAI Workshop on Computational Methods and Clinical Applications for Spine Imaging. Information about the image data is found in [6]. Before processing, the image data was resampled to an isotropic resolution of $1 \times 1 \times 1$ mm^3 using linear interpolation, in order to allow the usage of quadrature filters with isotropic resolution in the non-rigid registration step. Since ten data sets are available, each data set was segmented using the remaining nine as atlases.

The ground truth data was compared with the segmentations obtained from the atlas-based segmentation using the DICE coefficient, defined as

$$DICE = \frac{2 * |GT \cap S|}{|GT| + |S|} \tag{1}$$

where GT and S refer to the ground truth and the computed segmentations respectively, and where $|\ldots|$ denotes the volume in voxels.

4 Results

The results from the atlas-based segmentation of the ten CT data sets are provided in Table 1. In the results, it can be noted that lumbar vertebrae and the lower thoracic vertebrae, in general, obtain a DICE coefficient of $\geq$0.95, which has to be considered as very good. Good scores are also achieved for the remaining thoracic vertebrae but where there are some vertebrae that appear to fail for data set #4 and #8. In addition, vertebra T1 obtains, almost consistently, lower scores than the other vertebrae. Example visualizations of the results are provided in Fig. 1. Note the failed segmentation of vertebrae T5-T8 in Fig. 1c, d, which correspond to the segmentation results for data set #8.

Table 1 DICE scores as estimated from the atlas-based segmentation of the ten data sets

Vertebra/data set	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	All
T1	0.93	0.78	0.91	0.78	0.93	0.91	0.65	0.92	0.94	0.87	0.86
T2	0.94	0.94	0.94	0.93	0.94	0.93	0.93	0.93	0.94	0.94	0.94
T3	0.94	0.94	0.94	0.85	0.93	0.94	0.93	0.92	0.95	0.95	0.93
T4	0.92	0.93	0.94	0.65	0.94	0.94	0.93	0.87	0.94	0.95	0.90
T5	0.91	0.93	0.93	0.62	0.92	0.94	0.93	0.72	0.94	0.95	0.88
T6	0.93	0.95	0.93	0.72	0.93	0.95	0.94	0.41	0.94	0.95	0.86
T7	0.94	0.94	0.94	0.86	0.94	0.94	0.94	0.42	0.94	0.95	0.88
T8	0.94	0.95	0.95	0.93	0.94	0.94	0.95	0.63	0.95	0.95	0.91
T9	0.95	0.95	0.96	0.94	0.95	0.95	0.95	0.81	0.95	0.96	0.94
T10	0.95	0.95	0.96	0.95	0.94	0.96	0.95	0.89	0.96	0.96	0.95
T11	0.95	0.95	0.96	0.95	0.95	0.96	0.96	0.94	0.96	0.96	0.95
T12	0.95	0.95	0.96	0.95	0.95	0.96	0.95	0.95	0.96	0.96	0.95
L1	0.95	0.95	0.96	0.96	0.96	0.96	0.96	0.95	0.96	0.96	0.96
L2	0.95	0.95	0.96	0.96	0.96	0.96	0.96	0.95	0.96	0.96	0.96
L3	0.96	0.96	0.96	0.96	0.96	0.96	0.95	0.95	0.96	0.97	0.96
L4	0.95	0.95	0.96	0.96	0.96	0.97	0.96	0.96	0.96	0.97	0.96
L5	0.95	0.94	0.96	0.94	0.96	0.96	0.96	0.95	0.94	0.95	0.95
All	0.95	0.94	0.95	0.91	0.95	0.96	0.95	0.87	0.95	0.96	0.94

Note that the final column contains the means of the respective rows

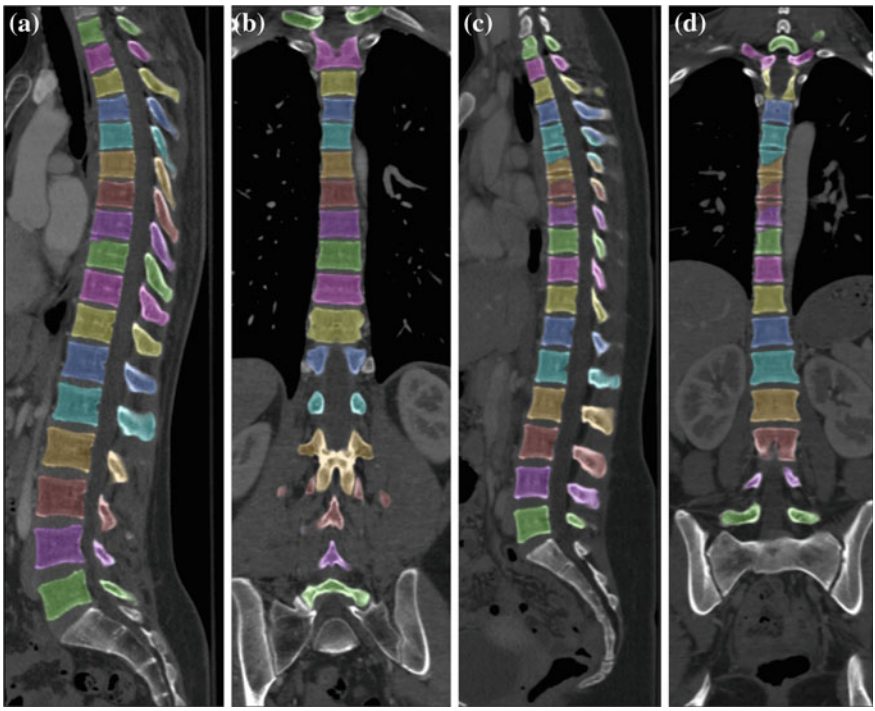

Fig. 1 Example results of the atlas-based registration, where **a** and **b** depict the results for data set #6 and **c** and **d** for data set #8, i.e. the data sets with the best respectively the worst results for the DICE coefficient

5 Discussion

In this work, a method for segmentation of the thoracic and lumbar vertebrae has been described and evaluated in relation to the segmentation challenge hosted by the 2nd MICCAI workshop on Computational Methods and Clinical Applications for Spine Imaging. The described method achieves a an average DICE coefficient of 0.94 ± 0.03 and is based upon atlas-based segmentation.

References

1. Forsberg, D., Eklund, A., Andersson, M., Knutsson, H.: Phase-based non-rigid 3D image registration—from minutes to seconds using CUDA. In: HP-MICCAI/MICCAI-DCI 2011 (2011)
2. Forsberg, D., Lundström, C., Andersson, M., Knutsson, H.: Model-based registration for assessment of spinal deformities in idiopathic scoliosis. Phys. Med. Biol. **59**(2), 311–326 (2014)

3. Forsberg, D., Lundström, C., Andersson, M., Vavruch, L., Tropp, H., Knutsson, H.: Fully automatic measurements of axial vertebral rotation for assessment of spinal deformity in idiopathic scoliosis. Phys. Med. Biol. **58**(6), 1775–1787 (2013)
4. Knutsson, H., Andersson, M.: Morphons: Segmentation using elastic canvas and paint on priors. In: 2005 IEEE International Conference on Image Processing (ICIP), pp. II–1226-9. IEEE (2005). doi:10.1109/ICIP.2005.1530283
5. Vrtovec, T.: Modality-independent determination of vertebral position and rotation in 3D. In: Dohi, T., Sakuma, I., Liao, H. (eds.) Medical Imaging and Augmented Reality. Lecture Notes in Computer Science, vol. 5128, pp. 89–97. Springer, Berlin (2008)
6. Yao, J., Burns, J., Munoz, H., Summers, R.: Detection of vertebral body fractures based on cortical shell unwrapping. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) Medical Image Computing and Computer-Assisted Intervention MICCAI. Lecture Notes in Computer Science, vol. 7512, pp. 509–516. Springer, Heidelberg (2012)

Lumbar and Thoracic Spine Segmentation Using a Statistical Multi-object Shape+Pose Model

A. Seitel, A. Rasoulian, R. Rohling and P. Abolmaesumi

Abstract The vertebral column is of particular importance for many clinical procedures such as anesthesia or anaelgesia. One of the main challenges for diagnostic and interventional tasks at the spine is its robust and accurate segmentation. There exist a number of segmentation approaches that mostly perform segmentation on the individual vertebrae. We present a novel segmentation approach that uses statistical multi-object shape+pose models and evaluate it on a standardized data set. We could achieve a mean dice coefficient of 0.83 for the segmentation. The flexibility of our approach let it become valuable for the specific segmentation challenges in clinical routine.

1 Introduction

Segmentation of the spinal column is an important task for many computer-aided diagnosis and intervention procedures. Despite the high contrast of bony structures in CT volumes, it remains challenging due to the presence of unclear boundaries, the complex structure of vertebrae, and substantial inter-subject variability of the

A. Seitel and A. Rasoulian contributed equally to this work.

A. Seitel (✉) · R.Rohling · A. Rasoulian (✉) · P. Abolmaesumi
Department of Electrical and Computer Engineering, University of British Columbia (UBC), Vancouver BC V6T 1Z4, Canada
e-mail: aseitel@ece.ubc.ca
A. Rasoulian
e-mail: abtinr@ece.ubc.ca
P. Abolmaesumi
e-mail: purang@ece.ubc.ca

R.Rohling
Department of Mechanical Engineering, UBC, Vancouver BC V6T 1Z4, Canada
e-mail: rohling@ece.ubc.ca

© Springer International Publishing Switzerland 2015

J. Yao et al. (eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, Lecture Notes in Computational Vision and Biomechanics 20, DOI 10.1007/978-3-319-14148-0_19

anatomy. Most of the proposed methods for automatic or semi-automatic spine segmentation rely on an initialization step of one or multiple vertebrae followed by a separate segmentation of each vertebra [1–7]. Considering each vertebra separately, however, may result in overlapping segmentations in areas where a clear boundary is missing in the volume data. Although there exist approaches as the one of Klinder et al. [2] that e.g. penalize overlapping areas, to our knowledge there is no method that incorporates common shape variations among the vertebrae of one subject which can be of great benefit for the segmenation quality. We thus propose an approach for segmentation of the spine in CT data which is based on a statistical multi-object model which incorporates both shape and pose information of the vertebral column.

2 Methods

Our segmentation technique is based on a statistical multi-vertebrae shape+pose model which is registered to the bony edges of the spinal column as extracted from the CT volume. The basic principles of this method have previously been presented in [8, 9] and will be summarized in the following paragraphs.

2.1 Model Construction

For construction of the model the idea is to analyze the pose and shape statistics separately as they are not necessarily correlated and are not formulated in the same parameter space. The model training then results in the modes of variations for both shape and pose, represented by v^s and v^p, respectively. Hence, a new instance of the model can be calculated as follows

$$S = \Phi\Big(\sum_{k=1}^{N_s} w_k^s v_k^s, \sum_{l=1}^{N_p} w_l^p v_l^p\Big). \tag{1}$$

where Φ is a similarity transform, N_s and N_p are the number of modes of variations for shape and pose, and w_k^s and w_l^p are the corresponding weights.

Building a single model for the entire vertebral column would require all vertebrae to be present in the training images and the images to be segmented. This limits the choice of volumes for the training data set and restricts the applicability of the segmentation method to such “complete” volumes. To be able to cope with arbitrary number of vertebrae present in the CT images and for segmentation of the whole spinal column, we propose to construct and align small sub-models with limited number of vertebrae. For this purpose, training data is collected for every vertebrae (in this case T1 to L5) and is used to build individual sub-models each containing 3 vertebrae and the ensemble of all models covering the whole spinal column (Fig. 1). The training step then results in many models and their associated modes of variation.

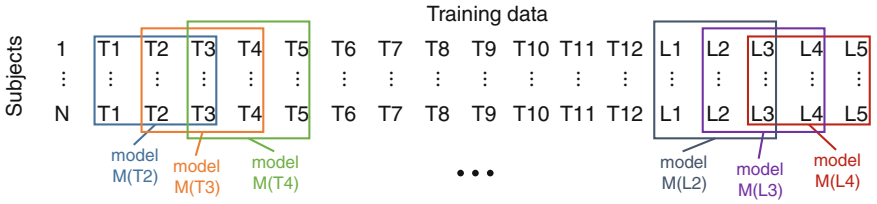

Fig. 1 Construction of the 3-vertebrae statistical shape+pose sub-models. Training data is available for all vertebrae. The individual models are build as detailed in [9]

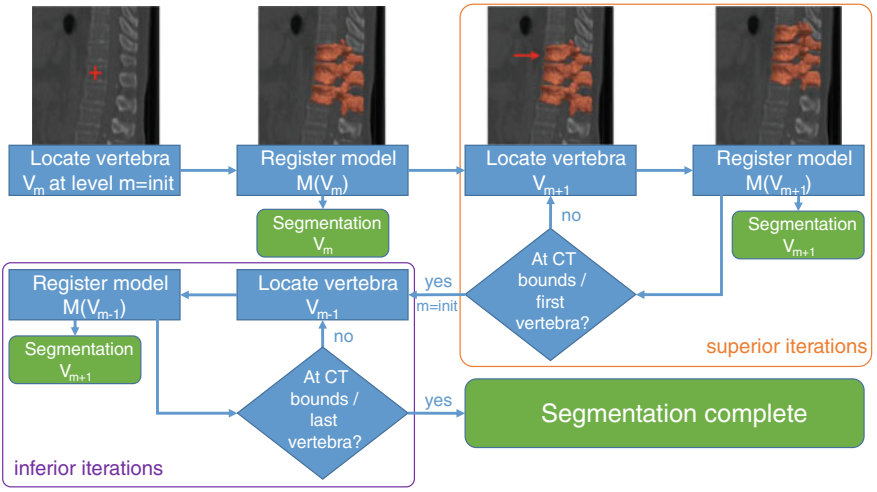

Fig. 2 Workflow of the segmentation approach. Initially, the center of gravity of one vertebral body is selected in the CT volume (*cross*). Next, the corresponding 3-vertebrae model is registered and the *middle* vertebra is segmented. The *last* vertebra (superior iterations) or the *first* vertebra (inferior iterations) of the registered model (*arrow*) is then used to initialize the next model. This process continues until it reaches the extents of the CT volume or the first/last vertebra

2.2 Segmentation

Segmentation using a single statistical multi-object shape+pose model can be formulated as a registration problem where the model is registered to the bone edge point cloud extracted from the CT volume using a canny edge detection preceded by a median filter (kernel radius 1). The transformation parameters as well as the described weights are then optimized using the Expectation Maximization (EM) algorithm such that the resulting model maximizes its likelihood of observing the CT edge point data [9].

The workflow for segmentation of the whole spinal column is depicted in Fig. 2. For initialization, the user has to specify the center of gravity of one specific vertebral body. After registration of the model starting at this initial position, the resulting registered instance is used to initialize the neighboring model either one level superior or one level inferior. This iterative registration is repeated until the new models reach

Table 1 Mean (μ) and standard deviation (σ) of dice coefficient for segmentation of individual vertebrae averaged over $n = 10$ cases

	L5	L4	L3	L2	L1	T12	T11	T10	T9	T8	T7	T6	T5	T4	T3	T2	T1
μ	0.82	0.85	0.85	0.87	0.83	0.82	0.75	0.75	0.76	0.76	0.76	0.75	0.75	0.84	0.83	0.82	0.75
σ	0.10	0.05	0.05	0.02	0.06	0.06	0.06	0.07	0.07	0.05	0.05	0.05	0.06	0.06	0.06	0.06	0.05

the extent of the CT volume or the first/last vertebra covered. The segmentation is then obtained from the registered models.

3 Results

Data from 87 CT volumes containing parts of the lumbar or thoracic spine were used for model construction. Evaluation and parameter optimization was performed on the provided training data. The segmentation results for the training data of the MICCAI challenge (leave-one-out approach) yielded a mean dice coefficient of 0.83 ± 0.04 (averaged over the ten cases) for the complete spine segmentation. The results for the individual vertebrae are shown in Table 1.

4 Discussion

Statistical multi-object models that incorporate both pose and shape statistics are evaluated with respect to their applicability for segmentation of the whole spinal column. We could achieve a mean dice coefficient of the segmentations of 0.83 ± 0.04 which is comparable with other approaches for spine segmentation. The usage of 3-vertebrae sub-models for the segmentation task let our method become flexible in terms of vertebrae covered by the input CT volume. This flexibility comes to price of a possible segmentation overlap at the boundaries of the sub-models especially for the closely spaced thoracic vertebrae. We are currently working on a generic n-vertebrae model that is able to cope with this issue and also allows for automatic model initialization. Further improvement is to be expected by consideration of the CT intensity information e.g. by means of an appearance modeling approach. We thus believe that the segmentation approach can be of great benefit for various interventional and diagnostic applications.

Acknowledgments This work was funded by the CIHR.

References

1. Burnett, S.S.C., et al.: A deformable-model approach to semi-automatic segmentation of ct images demonstrated by application to the spinal canal. Med. Phys. **31**(2), 251–263 (2004)
2. Klinder, T., et al.: Automated model-based vertebra detection, identification, and segmentation in CT images. Med. Image Anal. **13**(3), 471–482 (2009)
3. Ma, J., Lu, L.: 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 (2013)
4. Mastmeyer, A., et al.: 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 (2006)
5. Schmidt, S., et al.: Spine detection and labeling using a parts-based graphical model. Inf. Process Med. Imaging **20**, 122–133 (2007)
6. Shen, H., et al.: Localized priors for the precise segmentation of individual vertebras from ct volume data. In: Metaxas, D., Axel, L., Fichtinger, G., Székely, G. (eds.) MICCAI, LNCS, vol. 5241, pp. 367–375. Springer, Heidelberg (2008)
7. Vrtovec, T., et al.: Automated curved planar reformation of 3d spine images. Phys. Med. Biol. **50**(19), 4527–4540 (2005)
8. Rasoulian, A., et al.: Group-wise registration of point sets for statistical shape models. IEEE Trans. Med. Imaging **31**(11), 2025–2034 (2012)
9. Rasoulian, A., et al.: Lumbar spine segmentation using a statistical multi-vertebrae anatomical shape+pose model. IEEE Trans. Med. Imaging **32**(10), 1890–1900 (2013)

Vertebrae Segmentation in 3D CT Images Based on a Variational Framework

Kerstin Hammernik, Thomas Ebner, Darko Stern, Martin Urschler and Thomas Pock

Abstract Automatic segmentation of 3D vertebrae is a challenging task in medical imaging. In this paper, we introduce a total variation (TV) based framework that incorporates an a priori model, i.e., a vertebral mean shape, image intensity and edge information. The algorithm was evaluated using leave-one-out cross validation on a data set containing ten computed tomography scans and ground truth segmentations provided for the CSI MICCAI 2014 spine and vertebrae segmentation challenge. We achieve promising results in terms of the Dice Similarity Coefficient (DSC) of 0.93 ± 0.04 averaged over the whole data set.

1 Introduction

Due to reduced physical activity and modern office jobs that require prolonged sitting during work hours, pathological conditions affecting the spine have become a growing problem of modern society. As most spinal pathologies are related to vertebrae conditions, the development of methods for accurate and objective vertebrae segmentation in medical images represents an important and challenging research area. While manual segmentation of vertebrae is tedious and too time consuming to be used in clinical practice, automatic segmentation may provide means for a fast and

K. Hammernik (✉) · T. Ebner · D. Stern · T. Pock
Institute for Computer Graphics and Vision, BioTechMed, Graz University of Technology, Graz, Austria
e-mail: hammernik@icg.tugraz.at

D. Stern
e-mail: stern@icg.tugraz.at

T. Pock
e-mail: pock@icg.tugraz.at

M. Urschler
Ludwig Boltzmann Institute for Clinical Forensic Imaging, Graz, Austria
e-mail: martin.urschler@cfi.lbg.ac.at

© Springer International Publishing Switzerland 2015

J. Yao et al. (eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, Lecture Notes in Computational Vision and Biomechanics 20, DOI 10.1007/978-3-319-14148-0_20

objective analysis of vertebral condition. A current state of the art method for detecting, identifying and segmenting vertebrae in computed tomography (CT) images is proposed by Klinder et al. [5]. The method is based on a complex and computationally demanding alignment of statistical shape models to the vertebrae in the image. Using the deformable surface model and training an edge detector to bone structure, Ma et al. [6] segment and identify the thoracic vertebrae in CT images. In the work of Kadoury et al. [4], the global shape representation of individual vertebrae in the image is captured with a non-linear low-dimensional manifold of its mesh representation, while local vertebral appearance is captured from neighborhoods in the manifold once the overall representation converges during the segmentation process. Ibragimov et al. [3] used transportation theory to build their landmark-based shape representations of vertebrae and game theory to align the model to a specific vertebra in 3D CT images. In this paper, we propose a method for vertebrae segmentation in 3D CT images based on a convex variational framework. In contrast to the previously proposed methods that use sophisticated vertebral models, our segmentation method incorporates only a mean shape model of vertebrae initialized in the center of the vertebral body.

2 Methods

Our vertebrae segmentation algorithm is based on two representations of a priori information, a mean shape model and a bone probability map obtained from intensity information of the input vertebra image. The main steps of our algorithm are illustrated in Fig. 1. Firstly, an intensity based prior map of the bone is estimated by comparing the intensity values to trained bone and soft tissue histograms. This resembles our learned bone prior. The vertebral mean shape is then registered to the

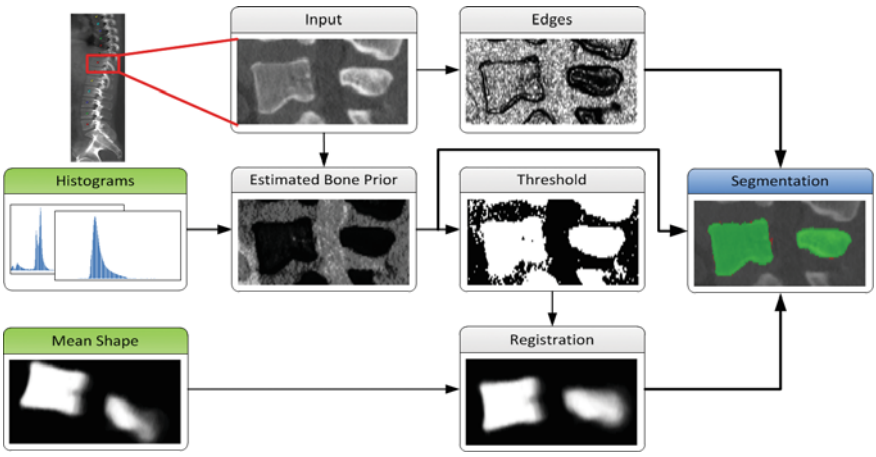

Fig. 1 Overview of our proposed algorithm. *Green boxes* represent a priori information obtained from training images. *Bold arrows* indicate parts that are included in the variational segmentation algorithm (color in online)

thresholded bone prior map to obtain the orientation of the individual vertebrae. This information is used to formulate a total variation (TV) based active contour segmentation problem, which combines the registered mean shape and the bone prior, and additionally incorporates edge information.

2.1 Mean Shape Model

The vertebral mean shape model f_s is calculated separately for three groups of vertebrae to account for variation in shape along the spine: T01–T06, T07–T12, L01–L05. Ground truth segmentations of vertebrae are registered to an arbitrary reference vertebra using an intensity-based registration with a similarity transformation and normalized cross correlation as similarity measure. The vertebral mean shape model is obtained by averaging the registered binary images of the ground truth segmented vertebrae. This step leads to a voxelwise probability for being part of the mean shape. To meet the requirements of the TV optimization framework [10], the obtained values in the probability image are inverted such that negative values represent the mean shape vertebral region and values close to one the non-vertebral region.

2.2 Bone Prior Map

The bone prior map $f_b(x) = log\left(\frac{p_{bg}(x)}{p_{fg}(x)}\right)$ is calculated as the log likelihood ratio between the probability that a voxel x belongs to the bone distribution p_{fg} and the probability that it belongs to the soft tissue distribution p_{bg}. The bone and soft tissue distributions are obtained from the training data set by estimating normalized mean foreground and background histograms of the intensity values using the ground truth segmentations. A coarse segmentation of the bone in the input image is achieved by thresholding the inverted bone map. We select a threshold value of -0.5 to ensure that trabecular bones are included in the segmented bone region, since their image intensities might be close to soft tissue.

2.3 Total Variation Segmentation

To obtain the segmented vertebra u, the following non-smooth energy functional $E_{seg}(u)$ is minimized using the first order primal-dual algorithm from [1]:

$$\min_{u\in[0,1]} E_{seg}(u) = \min_{u\in[0,1]} \mathrm{TV}_{g,\,aniso} + \lambda_1 \int_{\Omega} u f_s \, dx + \lambda_2 \int_{\Omega} u f_b \, dx \qquad (1)$$

where Ω denotes the image domain. The trade-off between the vertebral mean shape model, bone prior map and image edge influence is regularized by the parameters λ_1 and λ_2. The term $\mathrm{TV}_{g,\,aniso}(u)$ is the anisotropic g-weighted TV norm [7], using a structure tensor $D^{\frac{1}{2}}(x)$ as proposed by [9], incorporating both edge magnitude and

edge direction to be able to segment elongated structures:

$$\mathrm{TV}_{g,\,aniso} = \int_{\Omega} |D^{\frac{1}{2}}(x)\nabla u|\,dx = \int_{\Omega} \sqrt{\nabla u^T D(x) \nabla u}\,dx \tag{2}$$

$$D^{\frac{1}{2}}(x) = g(x)nn^T + n_0 {n_0}^T + n_1 {n_1}^T. \tag{3}$$

Here, $n = \frac{\nabla I}{\|\nabla I\|}$ is the normalized image gradient, n_0 denotes an arbitrary vector in the tangent plane defined by n, and n_1 is the cross product between n and n_0. The edge function $g(x)$ is defined as

$$g(x) = e^{-\alpha \|\nabla I(x)\|^{\beta}},\ \alpha, \beta \in \mathbb{R}^{+}. \tag{4}$$

During minimization of the energy E_{seg}, the segmentation u tends to be foreground, if f_b, $f_s < 0$ and background, if f_b, $f_s > 0$. If f_b, f_s equal zero, the pure TV energy is minimized, thus seeking for a segmentation surface with minimal surface area. The final segmentation is achieved by thresholding the segmentation u between 0 and 1.

3 Experimental Setup

We evaluated our method on the volumetric CT data sets provided for the CSI spine and vertebrae segmentation challenge [11]. The data consists of ten training images and the corresponding ground truth segmentations. The performance of our algorithm was evaluated by a leave-one-out cross validation, i.e., we report average performance over ten experiments.

We implemented the registration as well as the segmentation algorithm on the GPU to exploit hardware parallelization of our algorithms using Nvidia CUDA. For edge detection, we chose the parameters as $\alpha = 20$ and $\beta = 0.55$. The regularization parameters are set to $\lambda_1 = 0.04$ and $\lambda_2 = 0.005$ for all experiments. We achieved the binary segmentations by thresholding the segmentation u with a value of 0.2.

4 Results

For quantitive evaluation, we used the Dice Similarity Coefficient (DSC) to evaluate our segmentation algorithm. We achieved an average DSC of 0.93 ± 0.04 over all vertebrae from the leave-one-out experiment. Our algorithm performs well on lumbar vertebrae (0.96 ± 0.02) and lower thoracic vertebrae T07–T12 (0.95 ± 0.02). The DSC for thoracic vertebrae T01–T06 is 0.89 ± 0.05, which can be explained by the influence of ribs and small intervertebral discs that are connected to the vertebrae. The algorithm did not perform well on case 6 in terms of registration errors due to confusing trabecular bone intensities with soft tissue. All estimated DSC are depicted in Table 1 and 2. A qualitative result of a correctly segmented fifth lumbar vertebra is illustrated in Fig. 2. In contrast to that, Fig. 3 shows an example for the sixth thoracic vertebra where the segmentation is influenced by connected ribs.

Table 1 Mean values and standard deviations for each vertebra resulting from the leave-one-out cross validation experiment

DSC	T01	T02	T03	T04	T05	T06
Mean ± std	0.87 ± 0.03	0.92 ± 0.04	0.93 ± 0.03	0.91 ± 0.04	0.87 ± 0.06	0.84 ± 0.08
DSC	T07	T08	T09	T10	T11	T12
Mean ± std	0.94 ± 0.02	0.95 ± 0.02	0.95 ± 0.02	0.96 ± 0.02	0.96 ± 0.00	0.95 ± 0.03
DSC	L01	L02	L03	L04	L05	**All**
Mean ± std	0.96 ± 0.01	0.96 ± 0.01	0.96 ± 0.01	0.95 ± 0.03	0.94 ± 0.04	**0.93 ± 0.04**

We achieved good performance for T07–T12 as well as L01–L05. Our segmentations of T01–T06 have a lower value due to connected ribs and small intervertebral discs that are classified as bone in our pre-segmentation

Table 2 Evaluation in terms of mean values and standard deviations for all vertebrae in the individual data sets

DSC	case1	case2	case3	case4	case5
Mean ± std	0.93 ± 0.04	0.94 ± 0.05	0.92 ± 0.05	0.94 ± 0.03	0.94 ± 0.04
DSC	case6	case7	case8	case9	case10
Mean ± std	0.81 ± 0.18	0.94 ± 0.03	0.93 ± 0.03	0.95 ± 0.03	0.94 ± 0.04

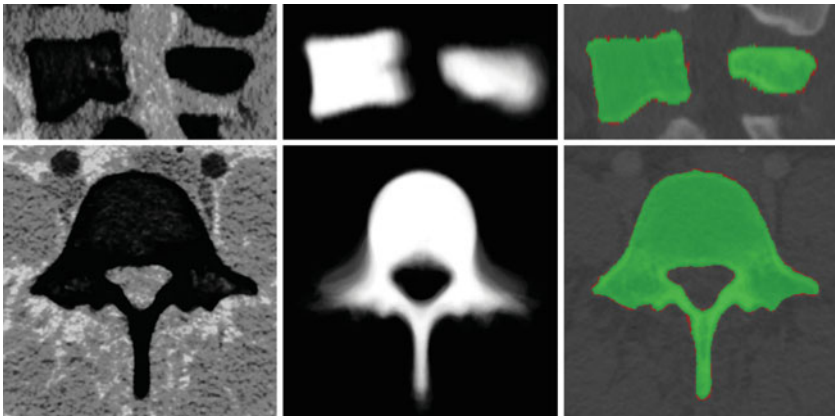

Fig. 2 The main steps of our proposed segmentation algorithm are bone prior estimation, mean shape registration and the final segmentation. The *top row* shows mid-sagittal cross sections and the *bottom row* axial cross sections of the lumbar vertebra L05. *Dark* regions in the bone prior (*left image*) are likely to be bone. The registered mean shape (*middle image*) and the bone prior are used as additional information for the segmentation algorithm. The final segmentation is depicted in the *right image* where *green* regions are correctly segmented and *red* regions differ from the ground truth result. The DSC for this example is 0.97 (color in online)

Fig. 3 This example illustrates axial cross sections of T06 where we achieved a DSC of 0.84. The bone prior shown in the *left image* classifies ribs as bone tissue. Therefore, the registered mean shape (*middle image*) is wrongly aligned, as it is attracted by the ribs. The final segmentation (*right image*) contains wrongly segmented ribs illustrated in *red*, while the yellow parts of the vertebra are missing. The *green* area depicts the correctly segmented region (color in online)

5 Discussion

Our proposed method for vertebrae segmentation based on a variational framework has been successfully applied to ten volumetric CT data sets provided for the CSI spine and vertebrae segmentation challenge. A common problem in vertebrae segmentation is that edges of the vertebrae are not clearly defined. Furthermore, trabecular bone intensities sometimes resemble soft tissue. These limitations could be easily overcome with our approach by adding manual user constraints as proposed in [8]. We see this as a great benefit of using our variational framework in vertebrae segmentation. In contrast to other methods that do not guarantee the convergence to an optimal solution, our proposed TV energy minimization (1) is convex, hence, it yields a globally optimal solution given the successfully registered mean shape. Initialized with a single point in the center of the vertebra, the proposed algorithm can be considered fully automatic since various methods have already been proposed for automatic detection and labeling of the center of the vertebral body [2]. While other methods depend on the sophisticated shape models, whose generation usually requires a great amount of time consuming manual interaction, in our proposed method we used coarse mean shape models built separately for all upper thoracic, lower thoracic and lumbar vertebrae. The vertebral mean shape model is registered to the thresholded bone prior map, which may lead to wrong alignment, if ribs are present or if trabecular bone intensities are low, i.e., close to soft tissue values. The overall result of 0.93 ± 0.04 in terms of the DSC is similar to other published methods. The result in lumbar region (0.96 ± 0.02) is better than 0.95 presented by Kadoury et al. [4] and the result of Ibragimov et al. [3] evaluated only on the lumbar vertebrae (0.93 ± 0.02). In the lower thoracic part of the spine, our result of 0.95 ± 0.02 exceeds the overall result for thoracic vertebrae (i.e. 0.93) presented by Kadoury et al. [4]. However, these methods were evaluated on different data sets, so results are not fully comparable. The lower DSC values for thoracic vertebrae T01–T06 can be explained by the influence of ribs and small intervertebral discs that mislead the segmentation.

6 Conclusion

In this work, a fully automatic system for vertebrae segmentation from CT was shown. It builds upon a TV based convex active contour segmentation that incorporates shape and intensity priors learned from training data and combines this prior information with image edges to achieve a minimal surface segmentation. Our results on the data of the MICCAI 2014 CSI challenge are promising and comparable to state of the art methods.

Acknowledgments This work was supported by the Austrian Science Fund (FWF) under the START project BIVISION, No. Y729, by the city of Graz (A16-21628/2013), and a European Community FP7 Marie Curie Intra European Fellowship (331239).

References

1. Chambolle, A., Pock, T.: A first-order Primal-Dual algorithm for convex problems with applications to imaging. J. Math. Imaging Vis. **40**(1), 120–145 (2011)
2. Glocker, B., Zikic, D., Konukoglu, E., Haynor, D.R., Criminisi, A.: Vertebrae localization in pathological spine CT via dense classification from sparse annotations. Medical Image Computing and Computer-Assisted Intervention—MICCAI 2013. Lecture Notes in Computer Science, vol. 8150, pp. 262–270. Springer, Berlin (2013)
3. Ibragimov, B., Likar, B., Pernus, F., Vrtovec, T.: Shape representation for efficient landmark-based segmentation in 3D. IEEE Trans. Med. Imaging **33**(4), 861–874 (2014)
4. Kadoury, S., Labelle, H., Paragios, N.: Automatic inference of articulated spine models in CT images using high-order Markov Random Fields. Med. Image Anal. **15**(4), 426–437 (2011)
5. Klinder, T., Ostermann, J., Ehm, M., Franz, A., Kneser, R., Lorenz, C.: Automated model-based vertebra detection, identification, and segmentation in CT images. Med. Image Anal. **13**(3), 471–482 (2009)
6. Ma, J., Lu, L., Zhan, Y., Zhou, X., Salganicoff, M., Krishnan, A.: Hierarchical segmentation and identification of thoracic vertebra using learning-based edge detection and coarse-to-fine deformable model. Medical Image Computing and Computer-Assisted Intervention—MICCAI 2010. Lecture Notes in Computer Science, vol. 6361, pp. 19–27. Springer, Berlin (2010)
7. Reinbacher, C., Pock, T., Bauer, C., Bischof, H.: Variational Segmentation of Elongated Volumetric Structures. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3177–3184 (2010)
8. Unger, M., Pock, T., Trobin, W., Cremers, D., Bischof, H.: TVSeg—interactive total variation based image segmentation. In: Proceedings of the British Machine Vision Conference (2008)
9. Weickert, J.: Anisotropic Diffusion in Image Processing. Teubner, Stuttgart (1998)
10. Werlberger, M., Pock, T., Unger, M., Bischof, H.: A variational model for interactive shape prior segmentation and real-time tracking. Scale Space and Variational Methods in Computer Vision. Lecture Notes in Computer Science, vol. 5567, pp. 200–211. Springer, Berlin (2009)
11. Yao, J., Burns, J.E., Munoz, H., Summers, R.M.: Detection of vertebral body fractures based on cortical shell unwrapping. Medical Image Computing and Computer-Assisted Intervention—MICCAI 2012. Lecture Notes in Computer Science, vol. 7512, pp. 509–516. Springer, Berlin (2012)

Interpolation-Based Shape-Constrained Deformable Model Approach for Segmentation of Vertebrae from CT Spine Images

Robert Korez, Bulat Ibragimov, Boštjan Likar, Franjo Pernuš and Tomaž Vrtovec

Abstract This paper presents a method for automatic vertebra segmentation. The method consists of two parts: vertebra detection and vertebra segmentation. To detect vertebrae in an unknown CT spine image, an interpolation-based optimization approach is first applied to detect the whole spine, then to detect the location of individual vertebrae, and finally to rigidly align shape models of individual vertebrae to the detected vertebrae. Each optimization is performed using a spline-based interpolation function on an equidistant sparse optimization grid to obtain the optimal combination of translation, scaling and/or rotation parameters. The computational complexity in examining the parameter space is reduced by a dimension-wise algorithm that iteratively takes into account only a subset of parameter space dimensions at the time. The obtained vertebra detection results represent a robust and accurate initialization for the subsequent segmentation of individual vertebrae, which is built upon the existing shape-constrained deformable model approach. The proposed iterative segmentation consists of two steps that are executed in each iteration. To find adequate boundaries that are distinctive for the observed vertebra, the boundary detection step applies an improved robust and accurate boundary detection using Canny edge operator and random forest regression model that incorporates prior knowledge through image intensities and intensity gradients. The mesh deformation step attracts the mesh of

R. Korez (✉) · B. Ibragimov (✉) · B. Likar · F. Pernuš · T. Vrtovec
Faculty of Electrical Engineering, University of Ljubljana,
Tržaska 25, SI-1000 Ljubljana, Slovenia
e-mail: robert.korez@fe.uni-lj.si

B. Ibragimov
e-mail: bulat.ibragimov@fe.uni-lj.si

B. Likar
e-mail: bostjan.likar@fe.uni-lj.si

F. Pernuš
e-mail: franjo.pernus@fe.uni-lj.si

T. Vrtovec
e-mail: tomaz.vrtovec@fe.uni-lj.si

© Springer International Publishing Switzerland 2015

J. Yao et al. (eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, Lecture Notes in Computational Vision and Biomechanics 20, DOI 10.1007/978-3-319-14148-0_21

the vertebra shape model to vertebra boundaries and penalizes the deviations of the mesh from the training repository while preserving shape topology.

1 Methodology

Let set $\mathscr{T}$ contain three-dimensional (3D) images of the thoracolumbar spine, where each image is assigned a series of binary masks representing reference segmentations of each individual thoracolumbar vertebra from level T1 to L5, and let each vertebral level be represented by a 3D face-vertex mesh $\mathscr{M} = \{\mathscr{V}, \mathscr{F}\}$ of $|\mathscr{V}|$ vertices and $|\mathscr{F}|$ faces (i.e. triangles). A chain of mean vertebra shape models represents the mean shape model of the whole thoracolumbar spine used for spine detection, while the mean shape models of individual vertebrae are used for vertebra detection and segmentation in an unknown 3D image I.

1.1 Vertebra Detection

The detection of vertebrae in an unknown 3D image I was performed by a novel optimization scheme based on interpolation theory [3], which consists of three steps: spine detection, vertebra detection and vertebra alignment. To detect the spine in image I, the pose of the mean shape model of the thoracolumbar spine $\mathscr{M}$ is optimized against three translations (i.e. coordinates x, y and z representing sagittal, coronal and axial anatomical directions, respectively), and the resulting global maximum represents the location of the spine in the 3D image, which is further used to initialize the vertebra detection. To detect each vertebra, the pose of the corresponding mean vertebra shape model $\mathscr{M}$ is optimized against three translations, however, in this case all local maxima of the resulting interpolation are extracted, corresponding to locations of the observed and neighboring vertebrae. The correct location of each vertebra is determined by the optimal path that passes through a set of locations, where each location corresponds to a local maximum at a different vertebral level. Finally, a more accurate alignment of the mean vertebra shape model is performed by optimizing the pose of each model against three translations, one scaling (i.e. factor s) and three rotations (i.e. angles φ_x, φ_y and φ_z about coordinate axes x, y and z, respectively). The resulting alignment represents the final vertebra detection result.

1.2 Vertebra Segmentation

After the interpolation-based alignment [3], segmentation of each vertebra in the unknown image I is performed by an improved mesh deformation technique [5] that moves mesh vertices to their optimal locations while preserving the underlying

vertebral shape [4, 6]. In this iterative procedure, the *image object detection* for mesh face centroids that are represented by the centers of mass for mesh faces $\mathscr{F} \in \mathscr{M}$ and *reconfiguration of mesh vertices* $\mathscr{V} \in \mathscr{M}$ are executed in each iteration.

1.2.1 Object Detection

By displacing each mesh face centroid $\mathbf{c}_i; i = 1, 2, \ldots, |\mathscr{F}|$ along its corresponding mesh face normal $\mathbf{n}(\mathbf{c}_i)$, a new candidate mesh face centroid $\mathbf{c}_i^*$ is found in each k-th iteration:

$$\mathbf{c}_i^* = \mathbf{c}_i + \delta\, j_i^*\, \mathbf{n}(\mathbf{c}_i), \tag{1}$$

where δ is the length of the unit displacement, and j_i^* is an element from set $\mathscr{J}$; $j_i^* \in \mathscr{J}$. Set $\mathscr{J}$ represents the search profile along $\mathbf{n}(\mathbf{c}_i)$, called the sampling parcel:

$$\mathscr{J} = \Big\{ -j, -j+1, \ldots, j-1, j \Big\}; \quad j = J - k + 1, \tag{2}$$

which is of size $2J+1$ at initial iteration $k = 1$ and $2(J-K+1)+1$ at final iteration $k = K$. The element j_i^* that defines the location of $\mathbf{c}_i^*$ is determined by detecting vertebra boundaries:

$$j_i^* = \underset{j \in \mathscr{J}}{\arg\max} \Big\{ F\big(\mathbf{c}_i, \mathbf{c}_i + \delta\, j\, n(\mathbf{c}_i)\big) - D\, \delta^2\, j^2 \Big\}. \tag{3}$$

where $\mathbf{c}_i' = \mathbf{c}_i + \delta\, j_i\, \mathbf{n}(\mathbf{c}_i)$ is the candidate location for $\mathbf{c}_i^*$ (Eq. 1), and parameter D controls the tradeoff between the distance from $\mathbf{c}_i$ to $\mathbf{c}_i'$ and the response of the boundary detection operator F:

$$F(\mathbf{c}_i, \mathbf{c}_i') = \frac{g_{\max}\left(g_{\max} + \left\|\mathbf{g}_W(\mathbf{c}_i')\right\|\right)}{g_{\max}^2 + \left\|\mathbf{g}_W(\mathbf{c}_i')\right\|^2} \left\langle \mathbf{n}(\mathbf{c}_i), \mathbf{g}_W(\mathbf{c}_i') \right\rangle, \tag{4}$$

where $\|\cdot\|$ denotes the vector norm, $\langle \cdot, \cdot \rangle$ denotes the dot product, $g_{\max}$ is the estimated mean amplitude of intensity gradients at vertebra boundaries that is used to suppresses the weighted gradients, which may occur if the gradient magnitude at the boundary of the object of interest is considerably smaller than of another object in its neighborhood (e.g. pedicle screws), and $\mathbf{g}_W$ is the image apperance operator at candidate mesh centroid location $\mathbf{c}_i'$:

$$\mathbf{g}_W(\mathbf{c}_i') = \big(1 + \alpha C(\mathbf{c}_i') + (1-\alpha) R(\mathbf{c}_i')\big)\, \mathbf{g}(\mathbf{c}_i'), \tag{5}$$

where $C(\mathbf{c}_i') \in [0, 1]$ is the response to the Canny edge operator, $R(\mathbf{c}_i') \in [-1, 1]$ is a random forest [1] regression model build upon an intensity-based descriptor and α is the weighting parameter.

1.2.2 Mesh Reconfiguration

Once the new candidate mesh face centroids $\mathbf{c}_i^*$ are detected, mesh $\mathscr{M} = \{\mathscr{V}, \mathscr{F}\}$ is reconfigured in each k-th iteration by minimizing the weighted sum E of energy terms:

$$\min_{\mathscr{M}}\{E\} = \min_{\mathscr{M}}\{E_{\text{ext}} + \beta E_{\text{int}}\}, \tag{6}$$

where β is the weighting parameter. The external energy E_{ext} attracts mesh $\mathscr{M}$ to new face centroids $\mathbf{c}_i^*$, $i = 1, 2, \ldots, |\mathscr{F}|$ (Eq. 1), that are located on vertebra boundaries:

$$E_{\text{ext}} = \sum_{i=1}^{|\mathscr{F}|} w_i^* \left\langle \mathbf{c}_i^* - \mathbf{c}_i, \frac{\mathbf{g}_W(\mathbf{c}_i^*)}{\|\mathbf{g}_W(\mathbf{c}_i^*)\|} \right\rangle^2 \tag{7}$$

where $|\mathscr{F}|$ is the number of mesh faces, $\mathbf{g}_W$ is the image appearance operator (Eq. 5), and w_i; $i = 1, 2, \ldots, |\mathscr{F}|$, are weights that are defined according to the obtained j_i^* (Eq. 3) to give a greater influence to more promising centroid locations:

$$w_i^* = \max\left\{0,\, F(\mathbf{c}_i, \mathbf{c}_i^*) - D\,\delta^2\, j_i^{*2}\right\} \tag{8}$$

The internal energy E_{int} restricts the flexibility of mesh $\mathscr{M}$ by penalizing the deviation between deformation vertices $\mathscr{V}$ and mean vertices $\mathscr{V}^m$:

$$E_{\text{int}} = \sum_{i=1}^{|\mathscr{V}|} \sum_{j \in \mathscr{N}_i} \left\| \left(\mathbf{v}_i - \mathbf{v}_j\right) - \left(s R \left(\mathbf{v}_i^m - \mathbf{v}_j^m\right) + \mathbf{t}\right) \right\|^2 \tag{9}$$

where $\mathbf{v}_i$ and $\mathbf{v}_i^m$ are vertices from sets $\mathscr{V}$ and $\mathscr{V}^m$, respectively, $\mathscr{M}^m = \{\mathscr{V}^m, \mathscr{F}^m\}$ represents the mean shape model of the observed lumbar vertebra, and $\mathscr{N}_i$ is the set of vertices neighboring to $\mathbf{v}_i$ (or $\mathbf{v}_i^m$, since the topology is preserved). The scaling factor s, rotation matrix R and translation vector $\mathbf{t}$ that align mesh vertices $\mathbf{v}_i$ to the mean vertices $\mathbf{v}_i^m$ are determined prior to calculation of Eq. (9) by using Procrustes superimposition [2].

2 Results

The performance of the proposed framework was tested on a database[1] [7] of 10 computed tomography (CT) images of the thoracolumbar spine and was evaluated by the mean symmetric absolute surface distance (MSD), symmetric root-mean-

[1] Publicly available through http://spineweb.digitalimaginggroup.ca.

Table 1 Thoracolumbar vertebra segmentation results in terms of mean symmetric absolute surface distance (MSD), symmetric root-mean-square surface distance (RMSSD), maximal symmetric absolute surface distance (MaxSD) and Dice coefficient (DICE), reported as mean ± standard deviation

Vertebral level	MSD (mm)	RMSSD (mm)	MaxSD (mm)	DICE (%)
T1	0.40 ± 0.06	0.71 ± 0.13	5.07 ± 0.94	90.25 ± 1.03
T2	0.33 ± 0.06	0.61 ± 0.13	4.75 ± 1.14	92.07 ± 1.00
T3	0.32 ± 0.05	0.61 ± 0.14	6.36 ± 5.67	92.16 ± 1.09
T4	0.44 ± 0.18	0.82 ± 0.36	5.62 ± 2.60	90.55 ± 2.82
T5	0.51 ± 0.33	1.02 ± 0.78	6.64 ± 6.07	90.27 ± 4.22
T6	0.37 ± 0.13	0.71 ± 0.30	7.35 ± 6.60	92.32 ± 1.27
T7	0.33 ± 0.08	0.60 ± 0.16	5.48 ± 2.25	92.97 ± 1.23
T8	0.31 ± 0.06	0.56 ± 0.11	4.76 ± 1.64	93.52 ± 1.14
T9	0.33 ± 0.07	0.56 ± 0.14	4.17 ± 1.58	93.56 ± 0.93
T10	0.40 ± 0.11	0.78 ± 0.29	6.97 ± 2.73	93.04 ± 1.28
T11	0.33 ± 0.06	0.61 ± 0.15	5.53 ± 1.99	93.97 ± 0.91
T12	0.32 ± 0.08	0.56 ± 0.11	4.72 ± 0.97	94.33 ± 1.03
L1	0.31 ± 0.05	0.56 ± 0.10	5.02 ± 1.74	94.46 ± 0.63
L2	0.31 ± 0.04	0.55 ± 0.10	5.38 ± 2.22	94.53 ± 0.58
L3	0.34 ± 0.07	0.63 ± 0.17	6.45 ± 2.54	94.25 ± 0.68
L4	0.32 ± 0.07	0.60 ± 0.14	6.38 ± 1.42	94.68 ± 0.63
L5	0.34 ± 0.06	0.64 ± 0.15	6.42 ± 2.47	94.14 ± 1.11
All	0.35 ± 0.12	0.65 ± 0.28	5.73 ± 3.15	93.09 ± 2.05

square surface distance (RMSSD), maximal symmetric absolute surface distance (MaxSD) and Dice coefficient (DICE),[2] computed between the resulting 3D meshes and corresponding reference segmentation binary masks. Detailed results for the segmentation of individual vertebrae are presented in Table 1.

References

1. Breima, L.: Random forests. Mach. Learn. **45**(1), 5–32 (2001)
2. Dryden, I.L., Mardia, K.V.: Statistical Shape Analysis. Wiley, New York (1998)
3. Ibragimov, B., Korez, R., Likar, B., Pernuš, F., Vrtovec, T.: Interpolation-based detection of lumbar vertebrae in CT spine images. In: Proceedings of the 2nd MICCAI Workshop on Computational Methods and Clinical Applications for Spine Imaging—MICCAI CSI 2014, pp. 65–75 (2014)

[2] The source code for evaluating segmentation results in terms of MSD, RMSSD, MaxSD and DICE is publicly available through http://mbi.dkfz-heidelberg.de/grand-challenge2007.

4. Klinder, T., Ostermann, J., Ehm, M., Franz, A., Kneser, R., Lorenz, C.: Automated model-based vertebra detection, identification, and segmentation in CT images. Med. Image Anal. **13**(3), 471–482 (2009)
5. Korez, R., Ibragimov, B., Likar, B., Pernuš, F., Vrtovec, T.: An improved shape-constrained deformable model for segmentation of vertebrae from CT lumbar spine images. In: Proceedings of the 2nd MICCAI Workshop on Computational Methods and Clinical Applications for Spine Imaging—MICCAI CSI 2014, pp. 76–85 (2014)
6. Weese, J., Kaus, M., Lorenz, C., Lobregt, S., Truyen, R., Pekar, V.: Shape constrained deformable models for 3D medical image segmentation. In: Information Processing in Medical Imaging—IPMI 2001, vol. LNCS 2082, pp. 380–387 (2001)
7. Yao, J., Burns, J.E., Munoz, H., Summers, R.M.: Detection of vertebral body fractures based on cortical shell unwrapping. In: Medical Image Computing and Computer-Assisted Intervention—MICCAI 2012, vol. LNCS 7512, pp. 509–516 (2012)

3D Vertebra Segmentation by Feature Selection Active Shape Model

Isaac Castro-Mateos, Jose M. Pozo, Aron Lazary and Alejandro Frangi

Abstract In this paper, a former method has been adapted to perform vertebra segmentations for the 2nd Workshop on Computational Methods and Clinical Applications for Spine Imaging (CSI 2014). A statistical Shape Models (SSM) of each lumbar vertebra was created for the segmentation step. From manually placed intervertebral discs centres, the similarity parameters are computed to initialise the vertebra shapes. The segmentation is performed by iteratively deforming a mesh inside the image intensity and then projecting it into the SSM space until convergence. Afterwards, a relaxation step based on B-spline is applied to overcome the SSM rigidity. The deformation of the mesh, within the image intensity, is performed by displacing each landmark along the normal direction of the surface mesh at the landmark position seeking a minimum of a cost function based on a set of trained features. The organisers tested the performance of our method with a dataset of five patients, achieving a global mean Dice Similarity Index (DSI) of 93.4 %. Results were consistent and accurate along the lumbar spine 93.8, 93.9, 93.7, 93.4 and 92.1 %, from L1 to L5.

I. Castro-Mateos (✉) · J.M. Pozo · A. Frangi
Centre for Computational Imaging and Simulation Technologies in Biomedicine (CISTIB), Department of Mechanical Engineering, University of Sheffield, Sheffield, UK
e-mail: isaac.casm@sheffield.ac.uk

J.M. Pozo
e-mail: j.pozo@sheffield.ac.uk

A. Frangi
e-mail: a.frangi@sheffield.ac.uk

A. Lazary
National Center for Spine Disorder (NCSD), Budapest, Hungary
e-mail: lazary.aron@gmail.com

© Springer International Publishing Switzerland 2015

J. Yao et al. (eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, Lecture Notes in Computational Vision and Biomechanics 20, DOI 10.1007/978-3-319-14148-0_22

1 Introduction

Statistical Shape Model (SSM) is a powerful tool to restrict shapes to a certain population and they have been successfully applied to vertebra segmentation [1–3]. The main drawback of such models is the large dataset required for the training set, which must be manually or semi-automatically segmented.

In this paper, we adapted our previous work on Intervertebral disc segmentation [4] to handle vertebrae. The method employs an SSM in a similar manner as in Active Shape Models (ASM) [5]. It also includes a relaxation step at the end of the process to allow segmentations to be out of the SSM space achieving highly accurate results even with a small training dataset.

2 Method

2.1 Initialization

In the training phase, the upper and lower endplate middle landmarks (l_a, l_b) of our template vertebra were selected. These two points allow to compute the centre of the vertebral body (VB) as well as the direction from l_b to l_a ($\hat{\mathbf{u}}$) of each initial vertebra.

In the segmentation step, the intervertebral discs (IVD) centres were manually pinpointed. From these points, the new VB centres were computed as the mean point between two consecutive IVD centres, with the exception of the L1, which was extrapolated from the other points. The scaling factor (s) and the rotation matrix ($\mathbf{R}$) were computed as:

$$s = \alpha \frac{\|l_a - l_b\|}{\|\mathbf{v}\|} \quad \text{and} \quad \mathbf{R} = \mathbf{I} + \mathbf{B} + \frac{1}{1 + \hat{\mathbf{u}} \cdot \hat{\mathbf{v}}} \mathbf{B}^2 \quad \text{with} \quad \mathbf{B} = \hat{\mathbf{u}}\hat{\mathbf{v}}^T - \hat{\mathbf{v}}\hat{\mathbf{u}}^T \tag{1}$$

where ($\mathbf{v}$) is the vector that goes from the IVD below the vertebra to the IVD on the top, $\mathbf{u} \cdot \mathbf{v}$ defines the dot product between $\mathbf{u}$ and $\mathbf{v}$. α was set to 0.9 and it was used to compensate the difference between the endplate landmarks and the distance between two consecutive IVDs.

2.2 Vertebra Segmentation

The algorithm follows our previous work [4], where a set of features and weights (w_i) were trained. For this task, the training database (Sect. 2.3) was used. Vertebrae were divided into two regions (see Fig. 1), the articular processes (AP) and the rest (RV). Three features were trained for each region, for the RV: The oriented derivative from high to low intensity, the intensity itself and the distance from the current position

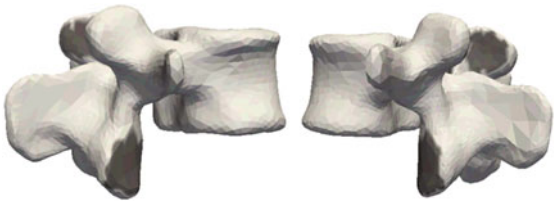

Fig. 1 The two vertebral regions, the *dark grey* areas in both figures are the AP region, where the rest of the vertebra, in beige, is the RV region (Color figure online)

were used. For the AP, the same features were employed but the first one, which was exchanged for a feature more adequate to the space between the articular processes. This feature seeks a sudden decrease of intensity followed by an increase of it. Thus, the cost functions are:

$$\mathrm{E}_{RV} = (w_1 I(\mathbf{v}_i) + w_2(\hat{\mathbf{n}}(\mathbf{s}) \cdot \nabla I(\mathbf{v}_i)) + w_3 EU(\mathbf{v}_i, \mathbf{p}_i)) \quad (2)$$

$$\mathrm{E}_{AP} = (w_1 I(\mathbf{v}_i) + w_2 \mathrm{min}_a(-F_a * I(\mathbf{v}_i)) + w_3 EU(\mathbf{v}_i, \mathbf{p}_i)) \quad (3)$$

where $\mathbf{p}_i$ is the current position of the mesh point i, $\mathbf{v}_i$ is a vector, whose values are the intensity profile positions in the image along the normal direction of the mesh surface at point $\mathbf{p}_i$, $\hat{\mathbf{n}}$ is the normal direction at $\mathbf{p}_i$ from inside to outside of the mesh, I is the image intensity, $EU(a, b)$ represents the euclidean distance from a to b and $F * I(\mathbf{v}_i)$ represents the convolution of a set of filter F and the profile of intensities. F is in the form of [1's, 0's, −1's, 0's, 1's], where the length (a) of the components are variable.

Having acquired the set of features, each mesh landmark is iteratively displaced, along the normal direction of the mesh surface, seeking the minimum of the previous cost function, and the resultant shape is projected into an SSM space. After the convergence of this iterative process a relaxation step is employed. This step starts with the same landmark displacement but the projection is performed over a B-spline space, with higher degrees of freedom.

2.3 Evaluation of the Method

Five SSMs were constructed using a training database of 30 Lumbar Spine CT images with a resolution of $0.608 \times 0.608 \times 0.3\,\mathrm{mm}^3$, collected at the National Center for Spinal Disorders (Budapest, Hungary). Those patients were selected for participating in the European MySpine project.

The testing dataset was composed of 10 CT scans and their ground-truth, both provided by the organisers, of thoracic and lumbar spines with resolution $0.3125 \times 0.3125 \times 1\,\mathrm{mm}^3$.

Furthermore, a second testing set of 5 CT scans was used by the organiser to test the performance of the algorithm. This second dataset was not available and only the results were given.

The segmentations were evaluated using the Dice Similarity Index (DSI), defined as $\frac{2\times(GT\cap SV)}{GT+SV} \times 100\,\%$, between the segmented vertebrae (SV) and the ground truth (GT).

3 Results and Discussion

The result of the Lumbar spine segmentations, of the first testing dataset, are shown in Table 1 and that of the second testing dataset in Table 2. The method shows stable performance and high accuracy along the different vertebrae, even when the training step was performed in a dataset with different characteristics as that of the testing (Sect. 2.3). Visual results (only for the first testing dataset) are presented in Fig. 2. This

Table 1 Mean and 2 standard deviations of Dice Similarity Index between the ground truth and the automatic segmentations of the first testing dataset composed of 10 patients

DSI	Lumbar spine: L1–L5 (%)	Global (%)
Mean	88.9, 88.8, 89.1, 89.2, 88.4	88.8
2 × STD	1.7, 3.2, 3.0, 2.5 1.7	2.5

Table 2 Mean and 2 standard deviations of Dice Similarity Index between the ground truth and the automatic segmentations of the second testing dataset composed of 5 patients

DSI	Lumbar spine: L1–L5 (%)	Global (%)
Mean	93.8, 93.9, 93.7, 93.4, 92.1	93.4
2 × STD	3.6, 0.9, 0.6, 1.7, 6.3	3.3

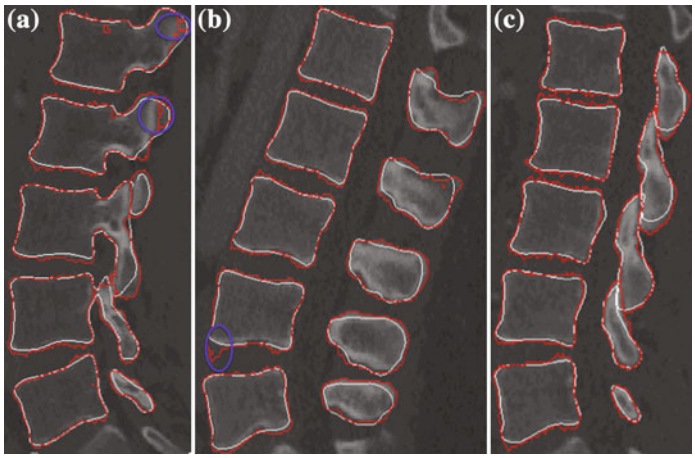

Fig. 2 Automatic segmentation results (*white contours*) and ground truth segmentations (*red contours*) of three patients. The images manifest that although results are accurate, the automatic algorithm under-segments the processes, and the ground truth was over-segmented and contains some artefacts (*blue contours*) (Color figure online)

figure shows that segmentation contours match the border of the vertebra proving the good performance of the method. However, it also reveals an under-segmentation issue in the processes. On the other hand, the ground-truth provided by the organisers seem to be over-segmented, mainly on the vertebral bodies, and it also contains some artefacts.

4 Conclusions

The proposed method employs the centre of the IVDs to compute the initial pose and scaling of each lumbar vertebra, then an improved ASM that employs a trained set of features and weights is used for the actual segmentation. The algorithm was tested only on lumbar vertebrae where SSMs were available. Results show high accuracy when comparing with the ground truth, which contains a high bias due to difficulty of the images.

References

1. Roberts, M.G., Cootes, T.F., Pacheco, E., Oh, T., Adams, J.E.: Segmentation of lumbar vertebrae using part-based graphs and active appearance models. In: Lecture notes in computer science (MICCAI), pp. 1017–1024. Springer (2009)
2. Klinder, T., Ostermann, J., Ehm, M., Franz, A., Kneser, R., Lorenz, C.: Automated model-based vertebra detection, identification, and segmentation in CT images. Med. Image Anal. **13**(3), 471 (2009)
3. Weese, J., Kaus, M., Lorenz, C., Lobregt, S., Truyen, R., Pekar, V.: Shape constrained deformable models for 3D medical image segmentation. Inf. Process. Med. Imaging pp. 380–387 (2001)
4. Castro-Mateos, I., Pozo, J.M., del Rio, L., Eltes, P., Lazary, A., Frangi, A.F.: 3D segmentation of annulus fibrosus and nucleus pulposus from T2-weighted magnetic resonance images. Manuscript accepted for publication in Physics in Medicine and Biology (2014)
5. Cootes, T.F., Taylor, C.J., Graham, J.: Active shape models- their training and applications. Comput. Vis. Image Under **61**(1), 38–59 (1995)

Report of Vertebra Segmentation Challenge in 2014 MICCAI Workshop on Computational Spine Imaging

Jianhua Yao and Shuo Li

Abstract Segmentation is the fundamental step for most spine image analysis tasks. The vertebra segmentation challenge held at the 2014 Computational Spine Imaging Workshop (CSI2014) objectively evaluated the performance of several algorithms segmenting vertebrae in spine CT scans. Five teams participated in the challenge. Ten training data sets and Five test data sets with reference annotation were provided for training and evaluation. Dice coefficient and absolute surface distances were used as the evaluation metrics. The segmentations on both the whole vertebra and its substructures were evaluated. The performances comparisons were assessed in different aspects. The top performers in the challenge achieved Dice coefficient of 0.93 in the upper thoracic, 0.95 in the lower thoracic and 0.96 in the lumbar spine. The strength and weakness of each method are discussed in this paper.

1 Introduction

The spinal column forms the central weight-bearing axis of the human body. Segmentation is the fundamental step for most spinal image analysis tasks. The vertebra segmentation is challenging due to the complex shape variation and inter-object interaction between vertebrae and ribs.

A number of spine segmentation algorithms on computed tomography had been proposed over recent years. Klinder et al. [1] proposed a method that most researchers

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

S. Li
GE Healthcare, Mississauga, ON, Canada
e-mail: shuo.li@ge.com

S. Li
University of Western Ontario, London, ON, Canada

© Springer International Publishing Switzerland (outside the USA) 2015

J. Yao et al. (eds.), *Recent Advances in Computational Methods and Clinical Applications for Spine Imaging*, Lecture Notes in Computational Vision and Biomechanics 20, DOI 10.1007/978-3-319-14148-0_23

referred as the state of the art technique by integrating detection, identification and segmentation of vertebrae in a framework. The method is based on spinal curve extraction and statistical shape models. Ma et al. [2] proposed a method based upon bone-structure edge detectors and coarse-to-fine registration of a deformable surface model for thoracic spine. Lim et al. [3] included the Willmore flow in the level set framework to guide the surface model evolution. Rasoulian et al. [4] incorporated both shape and pose statistics in a multi-vertebrae model for the lumbar spine segmentation. Kim and Kim [5] proposed a deformable fence model to separate lumbar vertebrae and surrounding tissues. Kadoury et al. [6] modeled the individual vertebra in an articulated spine model with a low-dimensional manifold representation and inferred the model using high-order Markov Random Fields. Ibragimov et al. [7] built landmark-based shape representations of vertebrae using transportation theory and aligned the model to a specific vertebra in 3D CT images using game theory. Roberts et al. [8] used part-based models and active shape models to divide the vertebra into several parts and conducted the segmentation collaboratively. Huang et al. [9] combined edge- and region-based level set functions for the vertebra CT image segmentation. Most of the published methods were based on certain type of vertebral models and reported fairly good results. However, these algorithms were mostly evaluated on self-prepared data set and their performances were not independently verified.

It is necessary to establish standardized reference data and validation criteria to objectively compare different segmentation algorithms. Challenge frameworks have been developed in the past few years for several medical image analysis problems. A dedicated website at http://www.grand-challenge.org has the most complete list of organized challenges since 2007, including those focusing on liver, lung, brain and heart etc. By far, there is no grand challenge exclusively for spine image analysis.

The objective to organize the vertebra segmentation challenge has three-folds. First, we intend to construct an annotated reference data set for spine labeling and segmentation; second, we will provide a platform to objectively evaluate the strength and weakness of various spine segmentation algorithms; and third, we want to assess the state of the art segmentation accuracy for vertebra and its substructures.

2 Data Sets

The data sets used in the challenge were acquired at the University of California, Irvine, Medical Center, between March 2013 and May 2013. The data set was manually selected by Dr. Joseph Burns according to the following selection criteria: age between 18 and 35 years; entire spine column scanned; and no obvious spine abnormality or disease. All patients were scanned using a spine CT protocol where a small field of view centered at the spine was reconstructed. The scanning parameters included 0.7–1 mm slice thickness, 120 kVp, soft tissue reconstruction kernel, and intravenous contrast.

There are several characteristics about this data set. It is from a young population with dense bone structures. Complete thoracic and lumbar spines are presented in the image. The spines are healthy. In all, we chose a relatively easy data set for the first challenge in this topic since we want to assess the mechanism of the vertebra segmentation algorithms and encourage participation.

We collected 15 data sets for the challenge, 10 for training and 5 for testing. The training cases were provided before the participants entered the challenge (Jan, 2014). The test cases were provided after they entered the challenge (July, 2014). All data were anonymized and made available in Meta format (MHD/raw). The data sets are released on SpineWeb (http://spineweb.digitalimaginggroup.ca/spineweb/index.php?n=Main.Datasets).

3 Reference Data Generation

The reference segmentation data was generated in two stages. First, the initial segmentation was obtained using a fully automatic algorithm reported in [10]. The algorithm was based on adaptive thresholding, watershed, directed graph search, and connected component analysis. The segmentation was then manually corrected and refined by a medical fellow and a research fellow. A customized software was developed for the manual correction. In the reference data, each vertebra is assigned a unique label and the background (pixels other than vertebra) is assigned label 0. We created reference segmentation for every vertebra for both thoracic (T1–T12) and lumbar (L1–L5) spines. A reference segmentation file is saved in Meta data format for each data set. The reference file has the same resolution as the original CT image file. Figure 1 shows an example of the reference label in the sagittal plane and the 3D surface model generated by the reference segmentation.

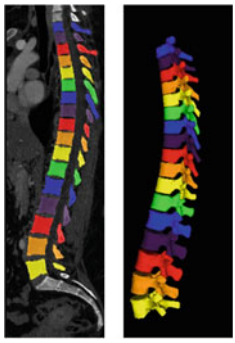

Fig. 1 Example of reference Data. Each vertebra is assigned a unique label (color coded). *Left* Sagittal view, *right* 3D view

4 Participating Algorithms

There are over 60 requests or downloads of the training data from SpineWeb. Five teams entered the challenge held at the Computational Spine Imaging workshop (CSI2014). Among the participants, four teams segmented both thoracic and lumbar spines, and one team only segmented the lumbar spine.

The five participating algorithms are dubbed as Method 1 [11], Method 2 [12], Method 3 [13], Method 4 [14], and Method 5 [15] in this paper according to the order of the submission. The titles of the five methods are listed in the bibliography. The technical comparisons of the five methods are conducted in six aspects: localization, model, bundled model, registration/optimization, feature function and running time. Method 1 and Method 4 comprise steps to automatically localize and initialize the vertebrae. Method 2, 3, and 5 require manual initialization of the model, either at the center of the vertebral body or at the endplates. All methods are based on certain type of shape and intensity models. Method 1 uses multiple atlases directly derived from the reference data. Method 2 uses statistical shape + pose model built from 87 training models. Both Method 3 and Method 4 use mean shape models. Method 4 builds a model for each vertebra level, while Method 3 builds one model for each spine section (one for T1–T6, one for T7–T12, and one for L1–L5). Method 5 computes a statistical shape model from 30 training models for each vertebra level. In Method 1 and Method 2, adjacent vertebrae are bundled together (5 vertebrae in Method 1 and 3 vertebrae in Method 2) in the segmentation. Different registration/optimization frameworks were adopted in the methods. Method 1 first applies a non-rigid registration for each atlas and then performs a label fusion. Method 2 conducts an EM algorithm to optimize the model. Method 3 adopts a total variation framework. Surface mesh deformation and reconfiguration is performed in Method 4. Method 5 employs statistical shape deformation plus b-spline relaxation for the surface optimization. Different feature functions are used in the methods for the optimization. They are mostly based on edge point and intensity model. The running time was reported by the participating teams as follows. Method 1-5 takes approximately 12, 10, 60, 30 and 3 minutes per case respectively. It can only be viewed as a reference since the algorithms were run on different hardware and some have been optimized by GPU.

5 Evaluation

The performance on the training set was evaluated by participants and reported in their submissions to the workshop. The performance of the test set was evaluated by the challenge organizers.

After the test data was released in July, the participants were given 10 days to submit the segmentation results. Each segmented vertebra was assigned a unique label. The results were submitted in Meta format (MHD/raw) with the same resolution of the original CT data.

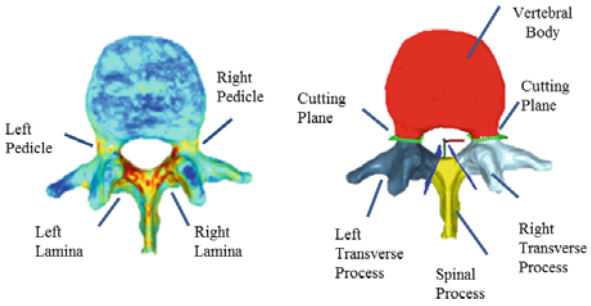

Fig. 2 Partitioning of a vertebra into four substructures. *Left* Density map on vertebra surface, hotter color: higher density. *Right* Partitioning a vertebra into four substructures. The substructures are color-coded with different colors. The cutting planes lie at the border between two substructures

Two metrics were employed for evaluation: dice coefficient (DC) and absolute surface distance (ASD). The evaluation was conducted on each individual vertebra.

By visual inspection, we notice that the segmentation performed differently at different parts of the vertebra. Therefore we evaluate the segmentation performance for both the whole vertebra and its four substructures: vertebra body, left transverse process, right transverse process, and spinal process. We developed an automatic method to partition the vertebra into anatomical substructures [16]. It is based on the anatomical knowledge that pedicles and laminae are the densest parts of the vertebral arch which forms the circle of bones around the spinal canal. We therefore search for the four cutting planes at left pedicle, right pedicle, left lamina and right lamina that go through the cross-sections with highest CT intensity around the spinal canal. Symmetric constraints are also enforced to balance the left and right cutting planes. The vertebra is then partitioned into four substructures (vertebral body, left transverse process, right transverse process and spinal process depending on which side of the cutting planes a pixel lies. Figure 2 shows the partitioning of a vertebra into four substructures.

6 Performance Comparison

The segmentation results were compared both visually and quantitatively. The segmentation result was superimposed on the CT image for visual inspection. Dice coefficient (DC) and absolute surface distance (ASD) were used for quantitative analysis. In this paper, we mainly focus on the results on the test set.

Figure 3 shows the visual comparison of submitted segmentation results for test case 1. All methods achieve visually acceptable segmentation for the thoracolumbar spine column. There is no overly leakage or under-segmentation from the sagittal view. For a closer visual inspection, Fig. 4 shows the visual comparison of the segmentation of the mid-axial slice for three representative vertebrae: T3, T9 and L3. In T3 and T9, all methods successfully separate the vertebra and the ribs. The border of segmented vertebra in Method 1is not smooth, which indicates further refinement

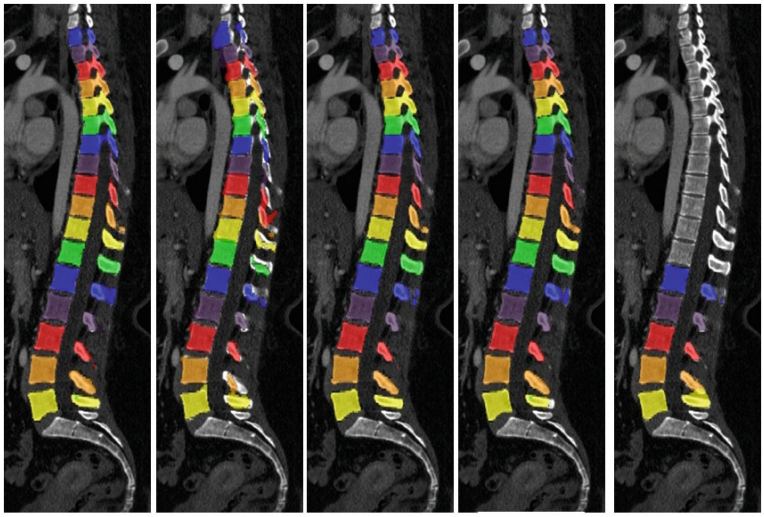

Fig. 3 Visual comparison of segmentation results for test case

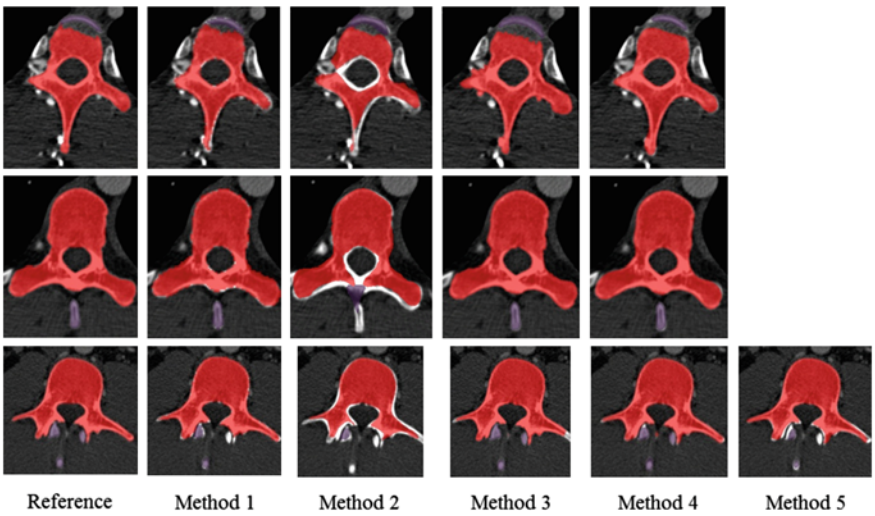

Fig. 4 Visual comparison of segmentation results for specific vertebrae in test case 5. *Row 1* T3 vertebra; *Row 2* T9 vertebra; *Row 3* L3 vertebra Mid-axial slice for each vertebra is shown. The segmentation is superimposed on the CT data

is necessary. The segmentation in Method 2 is off-mark although the location of the vertebra and the overall shape are correct. Another stage of local segmentation should be conducted. Method 3 and 4 both get pretty good segmentation results,

but it is noted that the segmentation of the posterior substructures still have rooms for improvement. The tips of the processes are not completely segmented and some contrast-enhanced vessels are included in the segmentation. Method 5 only segments the lumbar spine and the result is similar to that of Method 1 where the boundary is slightly off.

Figure 5 shows the average dice coefficient for all five methods from T1 to L5. There is a general trend of better performance from upper spine to lower spine as the vertebrae gradually increase the size and density. Both DC and ASD show the same pattern. To further illustrate the pattern, we group the vertebrae into three sections: upper thoracic from T1 to T6, lower thoracic from T7 to T12 and lumbar spine from L1 to L5. The DC goes from 0.867 in the upper thoracic, to 0.909 in the lower thoracic and to 0.933 in the lumbar spine.

Figure 6 summarizes the performance on the sub-structures. It is noted that the DC for the vertebral body segmentation is much higher than that for the posterior

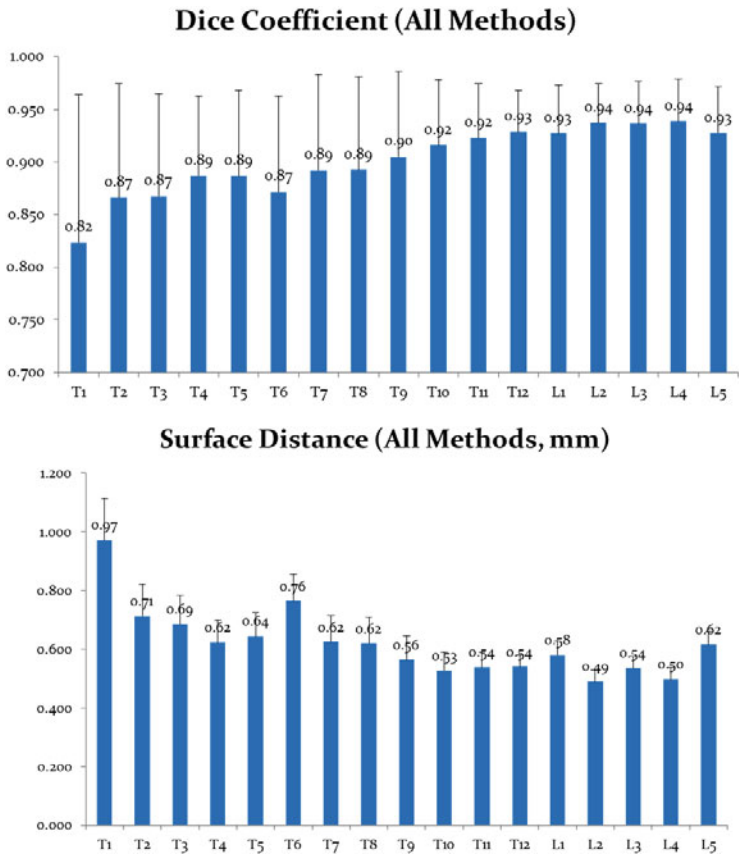

Fig. 5 Mean performance of all methods at each vertebra level

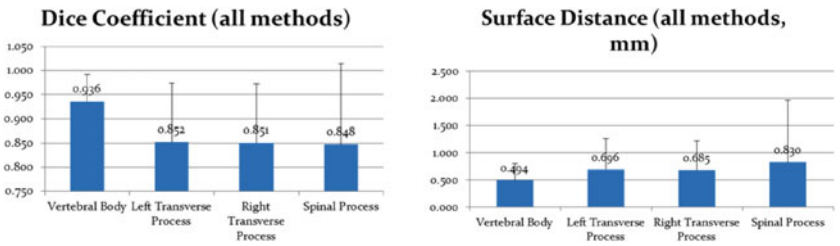

Fig. 6 Mean performance of all methods for substructures

substructures (left transverse process, right transverse process and spinal process). The three processes have comparable performance. This further verifies the visual comparison shown in Fig. 4.

Figure 7 compares the performance on individual vertebra for all five methods. Figure 8 summarizes the comparison on the entire spine column. These comparisons

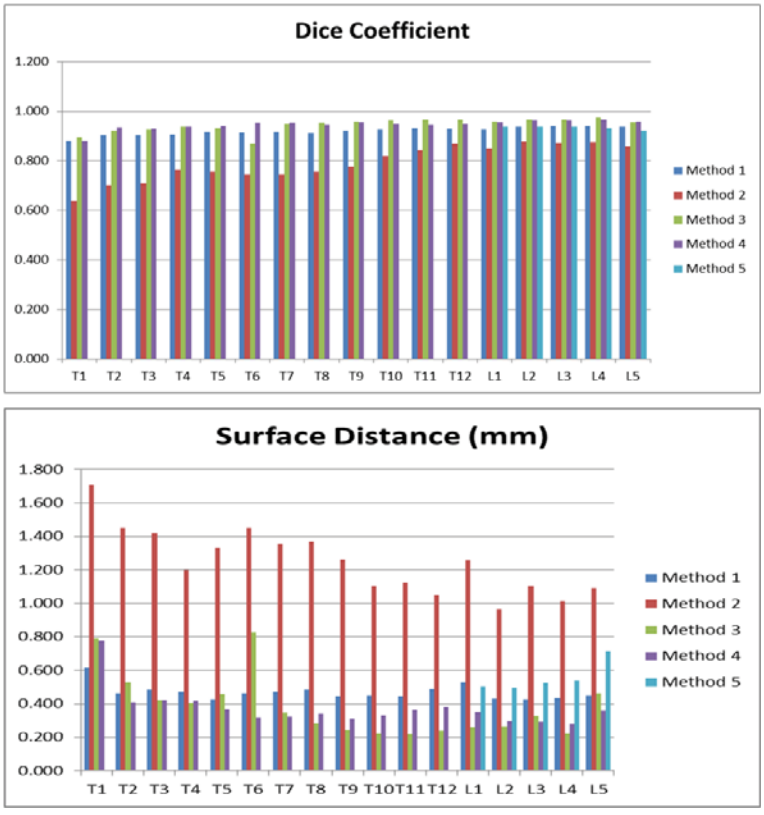

Fig. 7 Performance comparison on individual vertebra

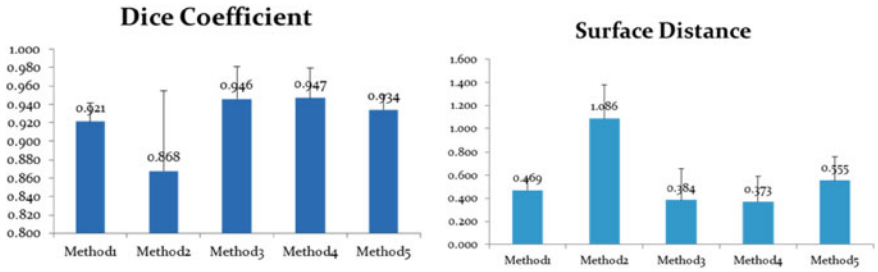

Fig. 8 Performance comparison on entire spinal column

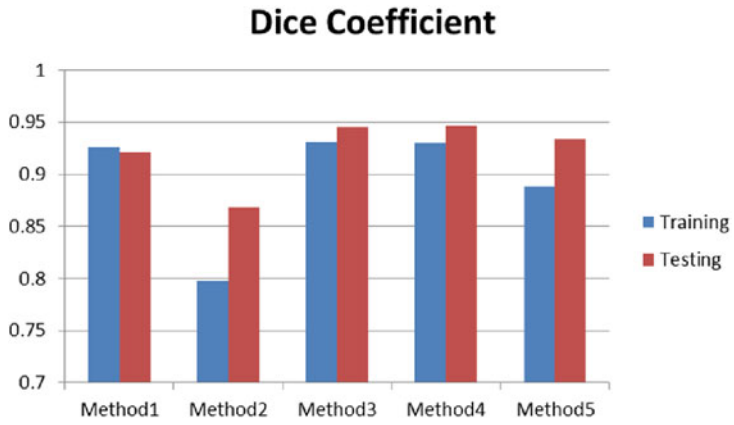

Fig. 9 Comparison between the training and testing cases

show the differences in performance among the methods, in terms of both Dice coefficient and surface distance. The two metrics show very similar difference patterns. Statistical analysis was conducted to evaluate the significance of the performance difference. Only two t-tests did not show statistically significant difference: the comparison between Method 3 and Method 4 ($p = 0.85$), and the one between Method 1 and Method 5 ($p = 0.52$). All other p-values are $<10^{-3}$. Based on this analysis, there are three tiers of performers among the five methods: Method 3 and Method 4 are in tier 1 (DC 0.946 and 0.947 respectively), Method 1 and Method 5 in tier 2 (DC 0.921 and 0.934 respectively), and Method 2 in tier 3 (DC 0.868).

Figure 9 compares the performance between the training and the testing sets. All methods show comparable or better performance in the test set. This comparison shows that the methods are generalizable. The better performance in the test cases may contribute to the quality of the reference data. The reference data for the test set had been corrected and verified by two operators while the training set was verified by only one operator.

7 Discussion

The performance of the participating methods ranges from 0.868 to 0.947 in Dice coefficient and 0.373 to 1.086 mm in absolute surface distance. The best results represent the state-of-the-art performance. However, it also indicates that there are still rooms for improvement. The substructure assessment shows that the vertebral body segmentation is generally excellent (0.94 on average, 0.97 in the best method). However, for other substructures (left and right transverse processes, spinal process), the dice coefficient is 0.85 on average and 0.92 in the best performer. A closer look in Fig. 4 also shows that most segmentation errors occur at the tip of the processes. Further refinement in the posterior substructures is possible and desirable.

The results also show that the segmentation results vary at different vertebra levels. For instance, upper thoracic has DC of 0.87 on average and 0.93 in best, lower thoracic 0.91 on average, 0.95 in best, and lumber 0.93 on average, 0.96 in best. The performance difference is mainly due to two reasons: (1) the size and density at the upper thoracic level is smaller and lower than that at the lumber level; (2) the surrounding structures are more complex at the upper thoracic level, for instance the costovertebral junctions that connecting the ribs and the vertebra. Further investigation is necessary to improve the segmentation at upper spine column.

All participating methods used models computed from some training data to segment the test data. Method 2 and Method 5 used their own training set to build the model, which may contribute to their relative low performance. Method 1 used multiple atlases from labels, Method 2 and Method 5 used statistical shape model and Method 3 and Method 4 used mean shape models. The mean shape models performed the best in this challenge. The possible reason is: the shape of a vertebra is highly variable. A statistical model computed from a small training set may not cover all the variants. More relaxation from the training models may be necessary to accommodate the variations. The mean shape model with large flexibility for deformation may work better especially for healthy and normal vertebra. However, statistical models with stricter domain constraint may be necessary for disease cases where the target shape is far from the mean shape and therefore severe under- or over-segmentation may occur if without shape and domain constraints. In this case, a large training set for a specific disease may be necessary.

Initial location of the model is essential for the accuracy of segmentation results. Three methods in this challenge required manually placed model locations. Automatic vertebra labeling and localization will be important for the methods to be applied in a clinical setting or to a large number of data sets. Manually placed seeds also have the issues of operator subjectivity. Methods based on spinal canal tracking [10, 17] or based on random forest models [18, 19] have shown robust and promising results to automatically locate the vertebrae.

Vertebra models are bundled in Method 1 and Method 2 so that the interaction between adjacent vertebrae can be employed to assist the segmentation. Although the results showed that the bundled models did not seem to improve the segmentation accuracy, this may be due to the data in this challenge comes from healthy

spines where the vertebrae are mostly intact. Therefore, individual vertebra model can handle the segmentation. In the case of disease cases, especially for cases with compression fractures, it would rely on the relatively healthy vertebrae in the neighborhood to assist the segmentation of the damaged vertebra. Thus, the bundled model is expected to be beneficial. Bundled model can also help prevent overlapping and collision between adjacent vertebrae.

Vertebrae at different levels have different shape, size and intensity. Two vertebrae far away in the spinal column such as between upper thoracic vertebrae and lumbar vertebrae show great difference. Therefore, it would be difficult or too complex to characterize all vertebrae in one model. Having noticed this, all participating methods built different models for different vertebra levels or at lease for different vertebral groups (Method 3). Vertebra specific models impose anatomical knowledge in the modeling and would be necessary for a robust segmentation.

Model fitting or image registration is widely accepted as the reliable way to segment complex objects such as a vertebra. Different frameworks for the registration or optimization had been adopted by the participating methods. They all converged to a solution however at different rates and running costs. The running time ranged from a few minutes to more than one hour.

Since vertebral bone has relatively high contrast against its surrounding tissues. Edge and gradient based feature functions were used in all methods. Those feature functions could be sensitive to noise and compromised by surrounding bony tissues (e.g. ribs). Recently machine learning techniques have been explored to classify pixels based on structural information and the contextural features. Furthermore, the image gradients vary at difference places due to different bone density and partial volume effect. The feature function must be adaptive to local features.

Method 3 and Method 4 have comparably the best performance in this challenge. Method 4 includes a component for automatic vertebra detection and localization, which makes it a more complete system. The methods were evaluated on a relatively easy data set. They may perform differently on more difficult data sets. To fully assess the algorithms in a realistic clinical setting, disease cases such as osteoporosis, fracture, degenerative change or bone lesions must be included. Nevertheless, the challenge on this data set revealed the capability of current state-of-the-art vertebra segmentation algorithms. It also identified a few areas that can be improved. The basic mechanism of the vertebra segmentation algorithms presented in this challenge should also work for more general clinical data sets. Through the comparison of various algorithms, the reader should get a sense of the strength and weakness of different vertebra segmentation algorithms and choose the appropriate one for their applications.

Acknowledgments We thank Dr. Joseph Burns in the Department of Radiological Sciences, University of California, Irvine, Medical Center for providing the CT data set. We thank Dr. Sasha Getty and Mr. James Stieger for providing the manual segmentation for the reference data. We also thank Dr. Ronald Summers in the Imaging Biomarkers and Computer-Aided Diagnosis Laboratory, Radiology and Imaging Sciences Department, NIH for providing the resource to carry out the evaluation. We also thank SpineWeb established by Digital Imaging Group of London for hosting the publicly available data set.

References

1. Klinder, T., Ostermann, J., Ehm, M., Franz, A., Kneser, R., Lorenz, C.: Automated model-based vertebra detection, identification, and segmentation in ct images. Med. Image Anal. **13**(3), 471–482 (2009)
2. Ma, J., Lu, L., Zhan, Y., Zhou, X., Salganicoff, M., Krishnan, A.: 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, vol. 6361, pp. 19–27. Springer (2010)
3. Lim, P., Bagci, U., Bai, L.: Introducing willmore flow into level set segmentation of spinal vertebrae. IEEE Trans. Biomed. Eng. **60**(1), 115–122 (2013)
4. Rasoulian, A., Rohling, R., Abolmaesumi, P.: Lumbar spine segmentation using a statistical multi-vertebrae anatomical shape + pose model. IEEE Trans. Med. Imaging **21**(10), 1890–1900 (2013)
5. Kim, Y., Kim, D.: A fully automatic vertebra segmentation method using 3D deformable fences. Compute. Med. Imaging Graph. **33**(5), 343–352 (2009)
6. Kadoury, S., Labelle, H., Paragios, N.: Automatic inference of articulated spine models in CT images using high-order Markov random fields. Med. Image Anal. **15**(4), 426–437 (2011)
7. Ibragimov, B., Korez, R., Likar, B., Pernus, F., Vrtovec, T.: Interpolation-based detection of lumbar vertebrae in CT spine images. In: Yao, J., Glocker, B., Klinder, T., Li, S. (eds.) 2nd MICCAI Workshops on Computational Methods and Clinical Applications for Spine Imaging (CSI2014), pp. 65–75. Springer (2014)
8. Roberts, M., Cootes, T., Adams, J.: Segmentation of lumbar vertebrae via part-based graphs and active appearance models. In: MICCAI, pp. 1017–1024 (2009)
9. Huang, J., Jian, F., Wu, H., Li, H.: An improved level set method for vertebra ct image segmentation. Biomed. Eng. Online **12**(1), 48 (2013)
10. Yao, J., O'Connor, S.D., Summers, R.M.: Automated spinal column extraction and partitioning. In: 3rd IEEE International Symposium on Biomedical Imaging: Nano to Macro, pp. 390–393 (2006)
11. Forsberg, D.: Atlas-based segmentation of the thoracic and lumbar vertebrae. In: Yao, J., Glocker, B., Klinder, T., Li, S. (eds.) 2nd MICCAI Workshops on Computational Methods and Clinical Applications for Spine Imaging (CSI2014), pp. 190–194. Springer (2014)
12. Seitel, A., Rasoulian, A., Rohling, R., Abolmaesumi, P.: Lumbar and thoracic spine segmentation using a statistical multi-object shape + pose model. In: Yao, J., Glocker, B., Klinder, T., Li, S. (eds.) 2nd MICCAI Workshops on Computational Methods and Clinical Applications for Spine Imaging (CSI2014), pp. 195–199. Springer (2014)
13. Hammernik, K., Ebner, T., Stern, D., Urschler, M., Pock, T.: Vertebrae segmentation in 3D CT images based on a variational framework. In: Yao, J., Glocker, B., Klinder, T., Li, S. (eds.) 2nd MICCAI Workshops on Computational Methods and Clinical Applications for Spine Imaging (CSI2014), pp. 200–207. Springer (2014)
14. Korez, R., Ibragimov, B., Likar, B., Pernus, F., Vrtovec, T.: Interpolation-based shape-constrained deformable model approach for segmentation of vertebrae from CT spine images. In: Yao, J., Glocker, B., Klinder, T., Li, S. (eds.) 2nd MICCAI Workshops on Computational Methods and Clinical Applications for Spine Imaging (CSI2014), pp. 208–212. Springer (2014)
15. Castro-Mateos, I., Pozo, J.M., Frangi, A.F.: 3D vertebra segmentation by feature selection active shape model. In: Yao, J., Glocker, B., Klinder, T., Li, S. (eds.) 2nd MICCAI Workshops on Computational Methods and Clinical Applications for Spine Imaging (CSI2014), pp. 213–217. Springer (2014)
16. Yao, J., Burns, J., Getty, S., Stieger, J., Summers, R.: Automated extraction of anatomic landmarks on vertebrae based on anatomic knowledge and geometrical constraints. In: International Symposium on Biomedical Imaging, pp. 397–400 (2014)

17. Forsberg, D.: Atlas-based registration for accurate segmentation of thoracic and lumbar vertebrae in CT data. In: Yao, J., Glocker, B., Klinder, T., Li, S. (eds.) 2nd MICCAI Workshops on Computational Methods and Clinical Applications for Spine Imaging (CSI2014), pp. 43–53. Springer (2014)
18. Glocker, B., Feulner, J., Criminisi, A., Haynor, D.R., Konukoglu, E.: Automatic localization and identification of vertebrae in arbitrary field-of-view ct scans. Med. Image Comput. Comput.-Assist. Interv. **7512**, 590–598 (2012)
19. Ibragimov, B., Likar, B., Pernus, F., Vrtovec, T.: Shape representation for efficient landmark-based segmentation in 3D. IEEE Trans. Med. Imaging **33**(4), 861–874 (2014)