Available online at www.sciencedirect.com
SciVerse ScienceDirect
International Journal of Spine Surgery 6 (2012) 167–173
INTERNATIONAL SOCIETY for the ADVANCEMENT of SPINE SURGERY
www.sasjournal.com

Variability of manual lumbar spine segmentation

Daniel J. Cook, MS [a], David A. Gladowski, BS [a], Heather E. Acuff [a], Matthew S. Yeager, BS [a], Boyle C. Cheng, PhD [a,b,*]

[a] Department of Neurosurgery, Allegheny General Hospital, Pittsburgh, PA
[b] Department of Neurosurgery, Drexel University College of Medicine, Pittsburgh, PA

Abstract

Background: The application of kinematic data acquired during biomechanical testing to specimen-specific, three-dimensional models of the spine has emerged as a useful tool in spine biomechanics research. However, the development of these models is subject to segmentation error because of complex morphology and pathologic changes of the spine. This error has not been previously characterized.
Methods: Eight cadaveric lumbar spines were prepared and underwent computed tomography (CT) scanning. After disarticulation and soft-tissue removal, 5 individual vertebrae from these specimens were scanned a second time. The CT images of the full lumbar specimens were segmented twice each by 2 operators, and the images of the individual vertebrae with soft tissue removed were segmented as well. The solid models derived from these differing segmentation sessions were registered, and the distribution of distances between nearest neighboring points was calculated to evaluate the accuracy and precision of the segmentation technique.
Results: Manual segmentation yielded root-mean-square errors below 0.39 mm for accuracy, 0.33 mm for intrauser precision, and 0.35 mm for interuser precision. Furthermore, the 95th percentile of all distances was below 0.75 mm for all analyses of accuracy and precision.
Conclusions: These findings indicate that such models are highly accurate and that a high level of intrauser and interuser precision can be achieved. The magnitude of the error presented here should inform the design and interpretation of future studies using manual segmentation techniques to derive models of the lumbar spine.

Keywords: Lumbar spine; Manual segmentation; Precision; Accuracy; Solid models

Modern medical imaging technology, such as computed tomography (CT) and magnetic resonance imaging, has made it possible to explore anatomic features in three dimensions (3D). Furthermore, advances in image processing have led to the development of specimen-specific models of certain anatomic features. These models are obtained by defining the portion of the image corresponding to the feature of interest, such as the brain, a tumor, or a single vertebra. This process is known as segmentation.

These models are becoming more prevalent in biomechanics research. They have been used to develop specimen-specific finite element models, and they have been used in conjunction with kinematic data acquired during biomechanical testing to investigate joint behavior. This technique provides a means of obtaining additional information with regard to the behavior of specific joint features under various loading conditions. The application of kinematic data to rigid body models has been used to investigate the region of contact at the facet joints, carpal bone interaction, and cartilage contact kinematics in the knee.[1–3] The fidelity of the analysis obtained from these techniques is dependent on the accuracy of the kinematic data acquired, the accuracy of registration between the reference frames of the motion capture system and the image data, and the accuracy of the solid models developed from the image data. Segmentation of the spine has proven to be a useful tool for several applications in medicine and medical research and requires varying levels of accuracy depending on which application is used. However, given the natural variation in morphology of spinal anatomy and limitations in imaging technology, segmentation of the human spine presents several challenges. Variation in the density of cortical bone across the surface of the vertebra can result in ambiguity in the boundary between bone and soft tissue in some regions, particularly in the spinous and transverse processes. The thickness of the articular cartilage of the facet is known to vary across

* Corresponding author: Boyle C. Cheng, PhD, 420 E North Ave, Ste 302, Pittsburgh, PA 15212; Tel: 412-359-4020; Fax: 412-359-8464.
E-mail address: boylecheng@yahoo.com

2211-4599
http://dx.doi.org/10.1016/j.ijsp.2012.04.002

its surface and between different levels of the spine. The mean thickness of this layer has been shown to vary between 0.49 and 0.61 mm across the cervical spine.[4] The proximity of adjacent facet surfaces presents the greatest challenge, given the fact that the maximum resolution of medical CT scanners ranges between 0.6 and 1.0 mm in the axial direction depending on the device. Anatomic variation is exacerbated by the presence of pathology. Narrowed or hypertrophic facet joints, bony growths, poor bone density, partial or complete fusion between levels, and degenerated discs contribute to the difficulty of accurately segmenting the lumbar spine. These challenges eliminate the possibility of defining complete models of the vertebral surface for a vast majority of spines based on a simple intensity threshold. Therefore segmentation of the human spine has required either the intelligence of a knowledgeable operator or that of a sophisticated algorithm.

Manual segmentation requires the persistent input of an operator and often uses image filters, intensity thresholding, morphologic filters, and manual "painting" or outline definition. Automated segmentation routines are distinguished from manual ones in that they rely, at least in part, on some image or pattern recognition algorithm.

Several methods have been developed for automated spine segmentation relying on a wide variety, and often a combination, of distinct segmentation frameworks, including thresholding, edge detection, and various manifestations of deformable models coupled with optimization routines. The normalized cuts method of Carballido-Gamio et al[5] segments vertebrae from 2-dimensional magnetic resonance images. The average reported error of this method ranged from 14.44% to 19.34% in vertebral body area from a manual segmentation baseline, depending on input values. de Bruijne et al[6] used a shape particle filtering method on spine radiographs that yielded an average segmentation error of 1.4 mm from manual segmentation by a medical expert and under 2.0 mm in 88 of 91 cases. Kim and Kim[7] developed a fully automatic vertebral segmentation method using deformable 3D fences for CT images, but the method was only evaluated qualitatively. Furthermore, only 80% of the specimens were segmented successfully with this automated routine. A promising class of algorithms that has emerged, model-based algorithms, relies on the inclusion of prior shape information to the segmentation process.[8,9] These prior shape models are usually referred to as active or adaptive in that the location and shape of the model can be modified to achieve optimum correspondence of the model with shape information contained within the image. A physical metaphor of energy is generally used to explain these algorithms with external energy used to describe the attraction of the model to image features and internal energy used to describe the restriction of the adaptation to a known shape. A minimization of the total energy is used to optimize the segmentation process. Klinder et al[8] developed such a technique and applied it to CT images of the thoracic spine. The group reported an average segmentation accuracy of 1.0 mm when compared with segmentation achieved through a similar algorithm using more operator interaction. Using an active shape model to segment the lumbar spine from planar X-rays, Zamora et al[10] reported an average error below 6.4 mm in 50% of cases.

These methods have been developed for fast identification or segmentation of vertebrae in applications such as surgical planning, deformity assessment, and image fusion. However, most model-based investigations of joint articulation have required manual segmentation for some or all of the specimens used in their respective studies.[3,11–15] To our knowledge, no currently available automatic segmentation algorithm has been shown to be successful at segmenting lumbar vertebrae with submillimeter accuracy. These algorithms are subject to variability in the definition of initial conditions and may converge to local minima in some cases.[7] Given that articular cartilage on the facet may be approximately 1.0 mm at its thickest point, these techniques do not appear to be sufficiently accurate for such purposes.[4] Furthermore, as described earlier, many presentations of these automatic algorithms rely on manual segmentation as a standard of accuracy. Because the rate of segmentation is of no relevance compared with accuracy in biomechanical studies of this type, a manual segmentation process was developed for CT images with the aim to produce models with submillimeter accuracy and precision.

The focus of this study is to examine the accuracy, as well as the intrauser and interuser precision, of this technique on a series of human cadaveric lumbar spines.

Methods

Specimen preparation

Eight cadaveric lumbar spine segments from T12 through the sacrum (4 female and 4 male cadavers; mean age, 59.6 years; age range, 51–68 years) were cleaned of muscle, loose connective tissue, and the anterior longitudinal ligament with special care given to preserving the remaining intervertebral ligamentous structures. Each specimen underwent CT scanning with a slice thickness of 0.6 mm in a 64-slice CT scanner (Somatom; Siemens, Munich, Germany).

Vertebral segmentation

Commercially available medical image analysis software (ScanIP; Simpleware, Exeter, England) was used to generate a 3D model of each vertebra. Each specimen was segmented twice by each of 2 operators (operator A and operator B). For ease of comprehension, the following scheme will be used throughout the remainder of this article: segmentation sessions will be abbreviated with operator name (A or B) followed by session number (1 or 2); for example, the second segmentation session by operator B will be abbreviated B2. To ensure that each segmentation session was independent of bias related to memory, the first and