

Lumbar Spine Segmentation Using a Statistical Multi-Vertebrae Anatomical Shape+Pose Model

Abtin Rasoulia, *Student Member, IEEE*, Robert Rohling, *Senior Member, IEEE*, and Purang Abolmaesumi*, *Senior Member, IEEE*

Abstract—Segmentation of the spinal column from computed tomography (CT) images is a preprocessing step for a range of image-guided interventions. One intervention that would benefit from accurate segmentation is spinal needle injection. Previous spinal segmentation techniques have primarily focused on identification and separate segmentation of each vertebra. Recently, statistical multi-object shape models have been introduced to extract common statistical characteristics between several anatomies. These models can be used for segmentation purposes because they are robust, accurate, and computationally tractable. In this paper, we develop a statistical multi-vertebrae shape+pose model and propose a novel registration-based technique to segment the CT images of spine. The multi-vertebrae statistical model captures the variations in shape and pose simultaneously, which reduces the number of registration parameters. We validate our technique in terms of accuracy and robustness of multi-vertebrae segmentation of CT images acquired from lumbar vertebrae of 32 subjects. The mean error of the proposed technique is below 2 mm, which is sufficient for many spinal needle injection procedures, such as facet joint injections.

Index Terms—Computed tomography (CT), multi-vertebrae anatomical model, registration, segmentation, spinal intervention, statistical shape+pose model.

I. INTRODUCTION

S PINAL needle injection is widely used for anesthetic, analgesic, and diagnostic purposes [1]. Facet joint injection for the relief of chronic lower back pain, is an example of such procedures performed in large numbers of hospitals and radiology clinics. Facet joint injection requires careful placement of the injection needle, both to ensure effective therapy delivery and to avoid damaging the nerves of the spinal cord. The current gold standard for the procedure is to use fluoroscopy [or computed tomography (CT)] for guiding the injection. However, fluoroscopy needle guidance for anesthesia delivery has

significant drawbacks, particularly the significant dose of ionizing radiation and the need for a specialized pain management clinic with access to fluoroscopy equipment. To alleviate these issues, several new image-guided techniques have been recently proposed [2]–[5]. In these techniques, a segmented anatomical model from CT is used either separately or in conjunction with other image modalities such as ultrasound, for planning and guiding the intervention. However, the key step for smooth integration of such techniques within a clinical workflow is a robust segmentation of CT images to create the 3-D anatomical model of spine. Unfortunately, despite the high contrast of bony structures in CT images, the segmentation task remains challenging due to the presence of unclear boundaries, the complex structure of vertebrae, and substantial inter-subject variability in performing and assessing segmentation [6].

Several semi-automatic and automatic methods have been proposed for vertebral column segmentation from CT images [6]–[12]. Most methods consist of two steps: 1) identification of vertebrae, and 2) separate segmentation of each vertebra. The identification may be performed manually by using seeds placed within the vertebral body [8] or by drawing contours [7]. In automatic techniques, the spine curvature is initially extracted and then vertebrae are detected by finding the inter-vertebral disk [10], [13], [14], searching for some unique characteristics of vertebrae shape [6], [9], or prediction of the vertebrae location using image features extracted from the entire image [15]. Techniques have been developed that unify these two steps by employing a local articular spine model [16], [17]. Notably, Kadoury *et al.* cluster the training subjects (vertebral column extracted manually from the CT images) into sub-groups using manifold learning, and construct linear point distribution models for each sub-group. An unseen target is then segmented by pairwise selection of the closest data from the training set based on the measures defined on the manifold. Although the results are promising, this approach requires a large enough training data set to cover the range of feasible postures. Several techniques are also proposed for the segmentation of a single vertebra. A typical approach is to use a template (or a mean shape) for initialization and warp it until it fits the edge of the vertebra [7], [8], [10]. Others use active appearance models (AMMs) to account for the inhomogeneity of the vertebrae intensity in CT images and also to incorporate the appearance statistics of the vertebrae in the segmentation task [17], [18]. Another approach addresses this problem by fitting a 3-D superquadric model to the vertebral body [19].

Separate segmentation of the individual vertebrae has several disadvantages. A clear boundary may not exist in regions

Manuscript received April 22, 2013; revised May 31, 2013; accepted June 07, 2013. Date of publication June 12, 2013; date of current version October 02, 2013. This work was supported in part by the Natural Sciences and Engineering Research Council (NSERC) and in part by the Canadian Institutes of Health Research (CIHR). Asterisk indicates corresponding author.

A. Rasoulia is with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, V6T 1Z4 Canada (e-mail: abtinr@ece.ubc.ca).

R. Rohling is with the Department of Electrical and Computer Engineering and Department of Mechanical Engineering, University of British Columbia, Vancouver, BC, V6T 1Z4 Canada (e-mail: rohling@ece.ubc.ca).

*P. Abolmaesumi is with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, V6T 1Z4 Canada (e-mail: purang@ece.ubc.ca).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMI.2013.2268424

between two vertebrae such as regions near the intervertebral disk and facet joints, which may lead to missegmentation of these areas or an overlap between the segmentation of neighboring vertebrae. Moreover, some anatomically useful information such as the common shape variations among vertebrae, is discarded. Although the former problem has been tackled by Klinder *et al.* [6] by simultaneous segmentation of all vertebrae and applying a penalty to overlapped areas, the latter has not yet been addressed.

To alleviate these issues, we propose to segment a section of the vertebral column, containing the target facet joint, using a statistical multi-object shape+pose model. Such models have recently been introduced to integrate variations of multiple anatomies (also referred to as complex) to allow simultaneous registration of them to an image [20]–[25]. The underlying assumption is the existence of a strong correlation between shapes and poses of different anatomies in the same patient, making these models capable of generating all possible variations of shapes and poses of a complex. Using these models, other advantages arise. In particular, by embedding the statistics of the relative position and orientation in the model, the pre-processing step for finding the spine curvature and detecting each vertebra can be avoided. The orientation of each vertebra is also constrained by its neighbors, which leads to more accurate segmentation with lower computational cost. An early application of multi-object statistical shape models is performed by Duta and Sonka [21]. In their technique, all segmented surface points are concatenated to form a single object. Next, the principal component analysis (PCA) is applied to find the main variations of the shapes. Wang *et al.* [22] uses the same approach to model multi-organs in mouse micro-CT images except that the shape correlations between different organs are described using a conditional Gaussian model. Tsai *et al.* [23] also follow the same concept integrated with a level-set function. They initially compute the mean shape by averaging all distance functions generated for all objects in the training set. The difference between samples and the mean shape are all stacked in a single matrix and then PCA is applied to extract the main modes of variation. One of the problems with such concatenation is the high-dimensionality of the problem, which may lead to inaccurate statistical analysis. To resolve this problem, Lu *et al.* [25] utilize M-reps, a medial representation of the objects, to capture shape features at various scale levels. In each scale, the modes of variation are described by a series of deformations, which are residues left from the coarser level. Another approach uses a hierarchical formulation of multi-object structures using the wavelet transform [24].

In all mentioned methods, the pose statistics are either neglected or jointly represented with shape statistics. However, this representation may lead to two major issues. First, the pose statistics are not necessarily correlated with the shape statistics, since they may depend on external factors such as the position and orientation of the patient during data acquisition. Second, the shape deformations are assumed to lie on a Euclidean space. However, the poses are represented by similarity transformations, i.e., rigid+scale transformations. These transformations form a Lie group, which is a Riemannian manifold where anal-

ysis performed in Euclidean space is not applicable [26]. Bossa and Olmos [20] addressed this problem by separating the pose and shape information and presenting statistical tools to extract variations among a set of objects. They used this technique to describe principal variation of brain structures such as the sub-cortical nuclei and lateral ventricles.

In this paper, we start with the approach proposed by Bossa and Olmos [20] and apply two major modifications. First, the original technique performs two levels of statistics. Initially, variations of shapes and poses are separately extracted, and then they are correlated to present a joint statistical model. This presentation is not optimal for the case of vertebrae, since the relative position of vertebrae depends highly on patient posture during data acquisition. Given that observation, we represent shape and pose variations separately in our model. Second, Bossa and Olmos's technique extracts shape variations for each anatomy separately whereas we perform the analysis on all vertebrae combined. The primary assumption for such analysis is that the shapes of different vertebrae in the same patient are strongly correlated.

We then propose a novel iterative expectation maximization (EM) registration technique to align the model to CT images of spine. In this technique, the alignment is considered as a probability density estimation problem. The points on the surface of the model are assumed to be centroids of a Gaussian mixture model (GMM) and the edge points extracted from the CT image are the observations. The model is then deformed using the modes of variation to maximize the likelihood of the GMM generating the observations. In order to account for variations in the segmentation target that are not within the pose and shape statistics of the model, we incorporate an additional nonrigid transformation step in the segmentation pipeline. Such transformation is represented by displacement vectors defined on each surface point of the model.

We perform experiments on data acquired from L1 to L5 vertebrae of 32 subjects. The results show that the proposed method can reliably segment the lumbar vertebrae and produce segmentations with mean error of below 2 mm, sufficient for facet joint injection procedures [27].

The rest of this paper is organized as follows. In Section II, we provide an overview of the technique used for construction of the multi-vertebrae model. Then, we present our proposed technique for registration of the model to CT images. The evaluation method of the technique is then presented. Results are given in Section III. Section IV discusses and concludes the paper.

II. METHODS AND MATERIALS

A. Methods

1) *Construction of the Statistical Multi-Object Shape+Pose Model:* A brief review on Lie groups and similarity transformations is provided in the Appendix. To construct the model, we adapt a method introduced by Bossa and Olmos [20]. We use superscript “*s*” and “*p*” to differentiate between shape and pose related variables, respectively. Table I provides an overview of the variables used in this section. In case of lumbar vertebrae,

TABLE I
VARIABLES USED IN THE TEXT TOGETHER WITH THEIR DESCRIPTION

Variable	Description
K	number of training data sets
L	number of objects
N_l	Number of surface points in the l th object in the training set
$\mathbf{S}_{k,l}$	$(N_l \times 3)$ matrix, the boundary of the l th object of the k th training set
μ_l^p	Mean pose for l th object
μ_l^s	$(N_l \times 3)$ matrix, the mean shape of the l th object
$\mathbf{T}_{k,l}$	(4×4) matrix, similarity transformation from the l th object in the mean shape to the corresponding object in k th training set
C	Number of shape and pose modes used in the registration
$\mathbf{v}_{c,l}^p$	(4×4) matrix, c th pose principal geodesic for l th object
$\mathbf{v}_{c,l}^s$	$(N_l \times 3)$ matrix, c th shape principal geodesic for l th object
\mathbf{Z}	$(M \times 3)$ matrix, the target which is a point set

K denotes the number of training sets ($K = 32$) and L denotes the number lumbar vertebrae ($L = 5$).

The method assumes the presence of dense correspondences among a training set that contains K samples of a complex with L objects. In this work, the correspondences are established across the boundary of objects, $\mathbf{S}_{k,l}$ which is a point set, using a group-wise GMM-based registration technique [28]. Partial Generalized Procrustes analysis [29], in which scaling is not performed (i.e., the size of the objects is preserved), is applied to the entire complex of all training sets to bring them into the same reference frame. Full generalized Procrustes analysis is also performed on each object separately to find the mean shape and its transformation to each instance. Let $\mathbf{T}_{k,l}$ be the similarity transformation from the l th object of the mean shape to the corresponding object of the k th instance. For each object, the mean transformation, μ_l^p , and the residuals in the tangent space, $\mathbf{U}_{k,l}^p = \log(\mu_l^{p-1} \mathbf{T}_{k,l})$, are found. For each $\mathbf{U}_{k,l}^p$, a 7×1 vector, $\mathbf{u}_{k,l}^p$, is defined according to basis (see (21) in the Appendix). The residual vector for each complex is then reconstructed by concatenating the residual transformations, $\mathbf{u}_k^p = (\mathbf{u}_{k,1}^p, \dots, \mathbf{u}_{k,L}^p)^\top$ and then PCA is applied. The results are C principal components, $\mathbf{v}_c^p = (\mathbf{v}_{c,1}^p, \dots, \mathbf{v}_{c,L}^p)^\top$, and their corresponding eigenvalues, λ_c^p . Corresponding principal geodesics (PG) can be computed by $\mu_l^p \exp(\mathbf{v}_{c,l}^p)$. The formulation of the PCA is shown in (1). Each $\mathbf{u}_{k,l}^p$ is a 7×1 vector representing the similarity transformation in Lie algebra. Concatenation of these vectors results in a matrix of size $K \times 7L$, which includes all the similarity transformations of the training set. The PCA decomposition results in two matrices, the score matrix and the eigenvectors. Matrix $\mathbf{W}_{K \times C}$ is the score matrix and is neglected afterward. The eigenvectors are listed in rows of a matrix of size $C \times 7L$ and can be separated for each object, $\mathbf{v}_{c,l}^p$

$$\begin{bmatrix} \mathbf{u}_{1,1}^p & \dots & \mathbf{u}_{1,L}^p \\ \mathbf{u}_{2,1}^p & \dots & \mathbf{u}_{2,L}^p \\ \vdots & \ddots & \vdots \\ \mathbf{u}_{K,1}^p & \dots & \mathbf{u}_{K,L}^p \end{bmatrix} = \mathbf{W}_{K \times C} \times \begin{bmatrix} \mathbf{v}_{1,1}^p & \dots & \mathbf{v}_{1,L}^p \\ \mathbf{v}_{2,1}^p & \dots & \mathbf{v}_{2,L}^p \\ \vdots & \ddots & \vdots \\ \mathbf{v}_{C,1}^p & \dots & \mathbf{v}_{C,L}^p \end{bmatrix} \quad (1)$$

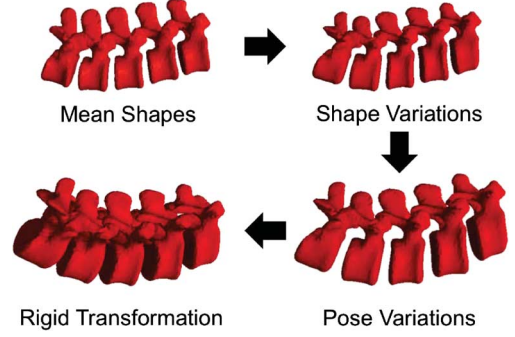


Fig. 1. Steps for instantiation of the model. Shape variations are applied to the mean shape followed by the pose variations. Rigid transformation is applied to the entire complex at the end.

To obtain the statistics of the shapes, one can assume that the distribution can be modeled as a multivariate Gaussian, the parameters of which are obtained using linear principal component analysis (PCA). However, in other cases, a nonlinear model such as Kernel PCA should be used [30]. Recently, a method was proposed by Bossa and Olmos which outperforms conventional PCA in many cases so is used here. This approach assumes that surfaces behave locally as a Riemannian manifold and principal geodesic analysis is performed to extract the modes of variations [31]. To find the shape variations, the instances are transformed into the tangent space at the mean shape, μ_l^s

$$\mathbf{u}_{k,l}^s = \log_{\mu_l^s}(\mathbf{S}_{k,l}). \quad (2)$$

Note that the logarithm mapping of a point set, x , to the tangent space of another point set, y , is defined as [20]

$$\log_y(x) = \text{vec}^{-1} \left(2 \arcsin(1/2 \|\hat{x} - \hat{y}\|) \frac{\hat{x} - \hat{y}(\hat{y}^\top \hat{x})}{\|\hat{x} - \hat{y}(\hat{y}^\top \hat{x})\|} \right). \quad (3)$$

The exponential mapping is defined as

$$\exp_y(x) = \text{vec}^{-1} \left(\cos \left(\frac{\|\hat{x}\|}{\|\hat{y}\|} \right) \hat{y} + \sin \left(\frac{\|\hat{x}\|}{\|\hat{y}\|} \right) \frac{\|\hat{y}\|}{\|\hat{x}\|} \hat{x} \right) \quad (4)$$

where $\hat{x} = \text{vec}(x)$, $\hat{y} = \text{vec}(y)$, and the operator, vec , represents a matrix vectorization.

Contrary to the original algorithm where shape statistical analysis is performed on each object separately, we concatenate all objects in the tangent space and apply PCA to them. Our assumption is that the anatomical shapes obtained from each patient are highly correlated with each other. This assumption is reasonable for vertebrae of the same person but may not be valid for other anatomical complexes. The results are principal components in the tangent space, $\mathbf{v}_c^s = (\mathbf{v}_{c,1}^s, \dots, \mathbf{v}_{c,L}^s)^\top$ (with corresponding eigenvalues, λ_c^s), which produce PGs using so called exponential mapping, $\exp_{\mu_l^s}(\mathbf{v}_c^s)$.

2) *Model Instantiation*: The shape+pose model inherits two kinds of separate statistics: positions and shapes. We will show later in Section III that these two are not correlated for the case of vertebrae. Therefore, we analyze the statistics of shape and pose separately. New complexes are generated by applying different weights to PGs and combining them. Fig. 1 shows the steps for the model instantiation. Initially, the shape variations

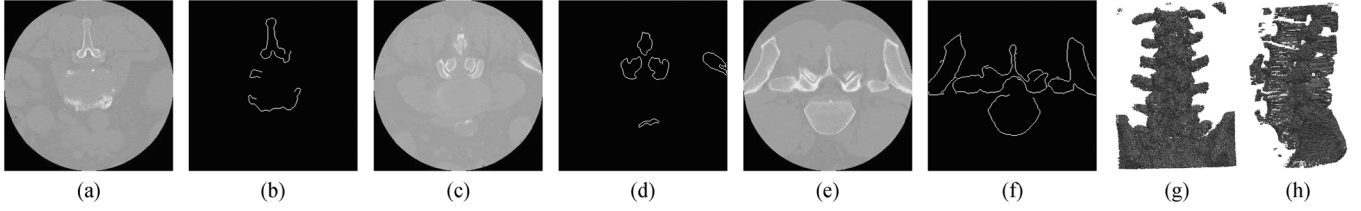


Fig. 2. (a)–(f) Example of results obtained with the Canny edge detector on sample CT images of spine. (g) and (h) Volume rendering of a typical result from the Canny edge detector.

are applied to the mean shapes of vertebrae, followed by pose variations and a rigid registration step, respectively. Assuming that θ_c^s is the weight applied to the c th shape PG and θ_c^p is applied to c th pose PG. A new instance of the l th object is generated as follows:

$$s_l = \Phi(\theta^s, \theta^p) = \Phi_l^p(\Phi_l^s(\theta^s); \theta^p). \quad (5)$$

Note that $\Phi_l^p(\cdot; \theta^p)$ denotes a similarity transformation built by combination of the pose PGs with corresponding weights

$$\Phi_l^p(\cdot; \theta^p) = \mu_l^p \prod_{c=1}^{\hat{C}} \exp(\theta_c^p \mathbf{v}_{c,l}^p). \quad (6)$$

PGs are sorted according to their corresponding eigenvalues. To limit the space of unknown parameters for the registration and to speed up the registration process, we only use a subset of PGs and ignore those with negligible variance, i.e., $\hat{C} \leq C$. Additionally, the shape instantiation, Φ_l^s , can be formulated as follows:

$$\Phi_l^s(\theta^s) = \exp_{\mu_l^s} \left(\sum_{c=1}^{\hat{C}} \theta_c^s \mathbf{v}_{c,l}^s \right). \quad (7)$$

3) Registration to CT Images of the Spine: A preprocessing step is performed to extract edges from the CT images of the lumbar spine. To this end, CT images are smoothed using a Gaussian filter with a standard deviation of 1 mm. Next, CT images are thresholded with a value of 100 Hounsfield units and a Canny edge detector [32] is applied to the 2-D transverse planes. The result is a point set \mathbf{Z} (consisting of points \mathbf{z}_m) representing a rough segmentation of the spinal column. The 2-D Canny detector could be replaced with a 3-D Canny detector, but given the larger inter-slice distances compared to intra-slice pixel spacing in the clinical CT images makes the use of 3-D Canny is likely inconsequential. Examples of such segmentations are shown in Fig. 2. The model is then initialized by a single point selected by the user within the middle vertebra. Since all of our CT images are captured from patients in a supine position, the orientation of the spinal columns are known *a priori* (i.e., along the z-axis). We did not observe any segmentation error regarding the scale variation in our data set; however, this might become an issue when the method is applied to cases that are not included in the training set, e.g., children.

To register the model to the extracted edges, we build upon an EM framework presented in our earlier work [28]. The target is

assumed to be an observation generated by a GMM, which is defined by the shape+pose model points as its centroids. The EM algorithm is used to maximize the likelihood of the observation by finding appropriate parameters for the modes of variations of the model and the extra nonrigid transformation.

Assuming \mathbf{t}_n^l to be the n th point of the model belonging to the l th object. The expectation step is performed by computing the probability of how likely a point in the model generates another point in the target

$$P(\mathbf{t}_n^l | \mathbf{z}_m) = \frac{\exp\left(-\frac{1}{2} \left\| \frac{\mathbf{z}_m - (\Phi(\mathbf{t}_n^l; \theta^s, \theta^p) + \nu_n^l)}{\sigma} \right\|^2\right)}{\sum_{l=1}^L \sum_{j=1}^N \exp\left(-\frac{1}{2} \left\| \frac{\mathbf{z}_m - (\Phi(\mathbf{t}_j^l; \theta^s, \theta^p) + \nu_j^l)}{\sigma} \right\|^2\right)}. \quad (8)$$

Here, $\Phi(\mathbf{t}_n^l; \theta^s, \theta^p)$ is the transformation of point \mathbf{t}_n^l given the parameters of the shape and the pose variations.

The displacement vector, ν_n^l , defined for each point on the model, is an additional nonrigid transformation added to the model to compensate for variations that are not within the statistics represented by the model. The standard deviation of all the Gaussian components, σ , is chosen initially to be large. Its value decreases in each iteration. In the maximization step, the following objective function is minimized:

$$Q = \sum_{l=1}^L \sum_{m,n=1}^{M,N_l} P(\mathbf{t}_n^l | \mathbf{z}_m) \left\| \mathbf{z}_m - (\Phi(\mathbf{t}_n^l; \theta^s, \theta^p) + \nu_n^l) \right\|^2 + \gamma^s \theta^{s\top} \Gamma^s \theta^s + \gamma^p \theta^{p\top} \Gamma^p \theta^p + \alpha \|\mathbf{L}\nu\|^2. \quad (9)$$

The three latter terms are for Tikhonov regularization over the shape, pose, and nonrigid transformation, and are controlled by the trade-off variables, γ^s , γ^p , and α , respectively. Matrices Γ^s and Γ^p are diagonal with elements $1/\lambda_c^s$ and $1/\lambda_c^p$, the corresponding eigenvalues of the shape and the pose PGs, respectively. The regularization over the nonrigid transformation is the same as that suggested by Myronenko *et al.* [33], which is proposed for constraining nonrigid registration of two point sets. Briefly, the operator \mathbf{L} intensifies the energy at high frequencies. In our implementation, it is represented by a Gaussian kernel with standard deviation, β , set initially by the user. Although the value of β depends on the size of the shapes, we set it to 5 mm for all the vertebrae.

The cost function is minimized alternately with respect to the model coefficients and the nonrigid transformation. The

model coefficients are found using the Quasi-Newton method. Ignoring the nonrigid transformation, we expand (9)

$$Q = \sum_{l=1}^L \sum_{m,n=1}^{M,N_l} P(\mathbf{t}_n^l | \mathbf{z}_m) \times \left(\mathbf{z}_m^\top \mathbf{z}_m - 2 \mathbf{z}_m^\top \Phi(\mathbf{t}_n^l; \theta^s, \theta^p) + \Phi(\mathbf{t}_n^l; \theta^s, \theta^p)^\top \Phi(\mathbf{t}_n^l; \theta^s, \theta^p) \right) + \gamma^s \theta^s^\top \Gamma^s \theta^s + \gamma^p \theta^p^\top \Gamma^p \theta^p. \quad (10)$$

The derivative with respect to shape parameter, θ_c^s , is

$$\frac{\partial Q}{\partial \theta_c^s} = \sum_{l=1}^L \sum_{m,n=1}^{M,N_l} P(\mathbf{t}_n^l | \mathbf{z}_m) \left[\left(\Phi(\mathbf{t}_n^l)^\top - \mathbf{z}_m^\top \right) \frac{\partial \Phi(\mathbf{t}_n^l)}{\partial \theta_c^s} \right] + \gamma^s \Gamma_{c,:}^s \theta^s \quad (11)$$

where $\Gamma_{c,:}^s$ is the c th row of the matrix Γ^s . The same can be achieved for the pose parameters

$$\frac{\partial Q}{\partial \theta_k^p} = \sum_{l=1}^L \sum_{m,n=1}^{M,N_l} P(\mathbf{t}_n^l | \mathbf{z}_m) \left[\left(\Phi(\mathbf{t}_n^l)^\top - \mathbf{z}_m^\top \right) \frac{\partial \Phi(\mathbf{t}_n^l)}{\partial \theta_k^p} \right] + \gamma^p \Gamma_{k,:}^p \theta^p. \quad (12)$$

The partial derivatives can be found separately for shape and pose parameters. Following (6), we have

$$\frac{\partial \Phi^s(\cdot; \theta^p)}{\partial \theta_c^p} = \mu_t^p \left(\prod_{i=1}^{c-1} \exp(\theta_c^p v_{c,i}^p) \right) \theta_c^p v_{c,l}^p \left(\prod_{i=c}^C \exp(\theta_c^p v_{c,i}^p) \right). \quad (13)$$

Following (4) and (7), the partial derivative with respect to the shape parameters can be written as

$$\frac{\partial \Phi_l^s(\theta^s)}{\partial \theta_c^s} = \cos \left(\frac{\|h\|}{\|\mu\|} \right) \frac{v_{c,l} h}{\|h\|^2} h + \sin \left(\frac{\|h\|}{\|\mu\|} \right) \left[\frac{v_{c,l} h \cdot \|\mu\|}{\|h\|^3} h + \frac{\|\mu\|}{\|h\|} v_{c,l} - \frac{v_{c,l} h}{\|h\| \cdot \|\mu\|} \mu \right] \quad (14)$$

where $h = \sum_{c=1}^C \theta_c^s v_{c,l}^s$.

The nonrigid transformation, (ν_n^l) , is also updated in each iteration. For fixed values of shape and pose coefficients, the optimization of the cost function becomes a nonrigid registration with respect to ν_n^l , which is solved using the closed-form solution proposed by Myronenko *et al.* [33].

B. Materials

1) *Data*: CT images of the lumbar (L1–L5) vertebrae were acquired from 32 subjects. Data are acquired at two institutions, Vancouver General Hospital (VGH) and Kingston General Hospital (KGH). The data from KGH ($n = 25$) was acquired for other research studies where patients' biometrics was not recorded. The spines appear normal, though we observe some cases with mild scoliosis. Subjects at VGH were characterized as follows: four male/three female; age 66.3 ± 5.8

years; and BMI $28.2 \pm 5.0 \text{ kg.m}^{-2}$. No abnormalities were seen. Written informed consent was obtained from all patients. Manual CT segmentations were performed semi-automatically using ITK-SNAP [34]. For each subject, three independent segmentations (performed by three different users) were averaged using majority voting to form the final segmentation, then triangulated using the marching cubes algorithm [35]. The CT resolution for all image sets ranged from $0.6 \text{ mm} \times 0.6 \text{ mm} \times 0.6 \text{ mm}$ to $0.9 \text{ mm} \times 0.9 \text{ mm} \times 3.2 \text{ mm}$ spacing.

2) *Validation Criteria*: Evaluations are based on the distances between the registered model and the triangulated surfaces. The shortest distances between vertices of the two surfaces are computed. The rms, mean, median, 95th percentile, and the maximum (referred to as the Hausdorff distance) of these distances are reported.

The model is evaluated by computing well-known measures of generalization, specificity, and compactness [36]. All measures are provided as a function of c , the number of modes used in model instantiation. Generalization, which measures the capability of the model to be adopted to a new observation, is measured using the leave-one-out method: a model is built using all but one member of the training set and then fitted to the excluded member. The generalization is then measured as an average of the distance error. The specificity measures the ability to represent only valid complexes of the object. To calculate it, the model is randomly instantiated within the range of valid parameters, $[-3\sqrt{\lambda_k^s} + 3\sqrt{\lambda_k^s}]$ and $[-3\sqrt{\lambda_k^p} + 3\sqrt{\lambda_k^p}]$, and is rigidly registered to all members of the training set. Specificity is then measured as the distance error to the closest member. Compactness is the ability to use a minimal set of parameters and is measured using the cumulative variance for the first c modes of the model.

Both accuracy and capture range are computed for the segmentation of the CT images using our proposed technique. The capture range is defined as the maximum possible distance between the center of the mass of the middle vertebra (L3) in the model and in the manual segmentation before the registration, for which the registration error is less than 2 mm in at least 90% of all cases. Since the narrow spaces between facets are between 2 and 4 mm [37], and for facet joint injections, a target accuracy of 5 mm has been reported to be sufficient [38], we expect that an accuracy of 2 mm should be accurate enough for clinical procedures.

III. EXPERIMENTS AND RESULTS

A. Statistical Multi-Vertebrae Shape+Pose Model

The statistical shape+pose model was constructed from the datasets, containing L1–L5 vertebrae of 32 patients, using the approach presented in Section II-A1. Each vertebra had 1500 vertices in the model. Fig. 3 illustrates the changes in the shape of the multi-vertebrae model that result from changing the weights corresponding to the first two pose modes of the variation by $\pm 3\sqrt{\lambda_k^p}$ and first two shape modes of the variation by $\pm 3\sqrt{\lambda_k^s}$.

An experiment was performed to assess the correlation between the shapes of vertebrae. Initially, correspondences be-

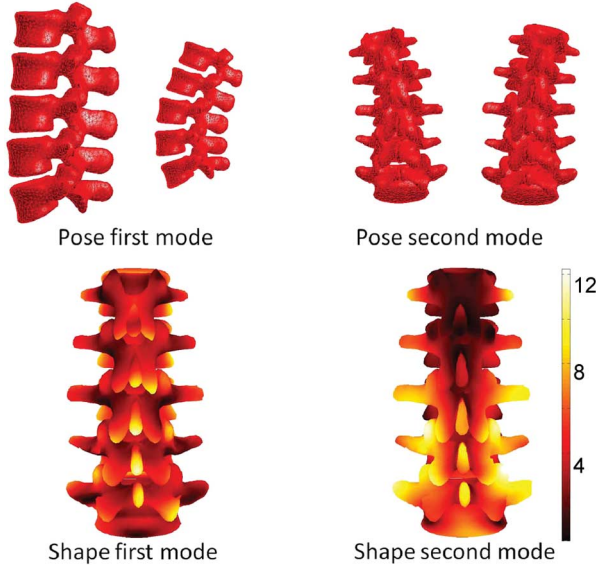


Fig. 3. Graphical representation of the L1–L5 vertebrae variations described by changing the weights corresponding to the first two principal modes of pose and shape variation by $\pm 3\sqrt{\lambda_k^p}$ and $\pm 3\sqrt{\lambda_k^s}$, respectively. The first mode of pose variations represents mainly the scaling and the lateral curvature, and the second mode represents the anterior–posterior curvature of the spine. The first and second modes of variation of shape represent variations in the transverse processes and vertebral body, respectively. For sake of better visualization, the amount of the shape variations is color coded on the mean shape. Unit is in millimeter.

TABLE II
PAIR-WISE CORRELATION BETWEEN SHAPES OF LUMBAR VERTEBRAE

	L2	L3	L4	L5
L1	0.98±0.02	0.98±0.04	0.96±0.08	0.94±0.10
L2		0.98±0.02	0.96±0.07	0.95±0.09
L3			0.98±0.04	0.96±0.07
L4				0.97±0.04

tween each two vertebrae were detected using a nonrigid registration [33]. Next, pair-wise correlation between correspondences were computed and averaged. The average correlations between each two vertebrae are presented in Table II. The correlation values suggest that the shapes of vertebrae closer to each other (e.g., L2 and L3) are more correlated than those far from each other (e.g., L1 and L5).

To show a lack of correlation between the pose and the shape variations, we performed pair-wise correlation between the variation modes score for the training set and calculated the R^2 (square of the Pearson correlation) accordingly. All R^2 values were below 0.4 and only six variations had $R^2 > 0.3$ which showed an insignificant relationship between pose and shape statistics. Given the results, we performed separate analysis on pose and shape for the rest of this paper.

Quantitative results for model accuracy are shown in Fig. 4. Compared to shape, pose statistics are more compact, i.e., the variations of the pose can be represented with fewer modes [compare Fig. 4(a) and (b)]. This is to be expected since the number of pose parameters are significantly less than the number of shape parameters. According to the generalization measure, shown in Fig. 4(c), the model is capable of reconstructing a new observation with a mean distance error

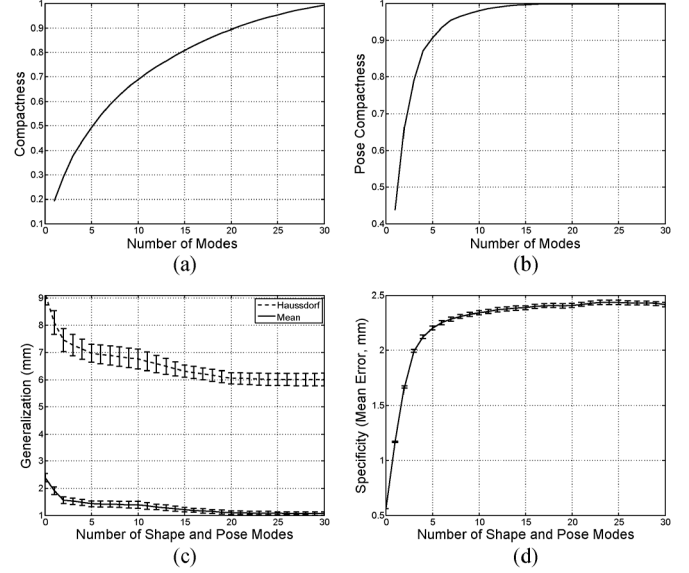


Fig. 4. Generalization (for both Mean Error and Hausdorff), specificity, and compactness ability of the statistical multi-vertebrae shape+pose model. Error bars show the standard deviation divided by square root of the number of trials. (a) Pose compactness. (b) Generalization. (c) Specificity. (d) Specificity.

around 1.4 mm when using the first 10 modes of the variation. Specificity is below 2.5 mm and, as we observed in our results, the error doubles when extreme values of different modes are used, i.e., the boundaries of the (hyper-rectangular) space of allowable shapes and allowable poses.

We should also note that the proposed technique relies on performing linear analysis on data with large dimensions and few observations. This will reduce the quality of the model in terms of compactness and generalization. A larger training set may resolve this problem.

B. Segmentation of CT Images Using the Multi-Vertebrae Model

CT images are randomly divided into two sets. One set, consisting of five CT images, is used to optimize the trade-off variables of the algorithm. The rest (27 CT images) are used to measure the segmentation accuracy and its capture range. Since the size of the training set is limited, in all experiments explained in the following sections, the model is reconstructed using all manual segmentations excluding the manual segmentation of the target (i.e., using 31 CT images).

1) *Selection of the Trade-Off Variables of the Registration Algorithm:* The appropriate choice of the trade-off variables (γ^p , γ^s , and α) and number of modes, c , were found by fitting the model to a subset of five CT images with different values for these variables. The mean distance error of the fitted model was computed. Fig. 5 shows the distance error versus a range of values of the trade-off variables.

Constraints on the pose, shape, and nonrigid transformation. The distance error increases as the value of γ^s and α decreases, since the model is more likely to be deviated from the mean shape. The same is observed for the large value of these variables, since then the model is forced to be closer to the main shape. However, this was not the case for

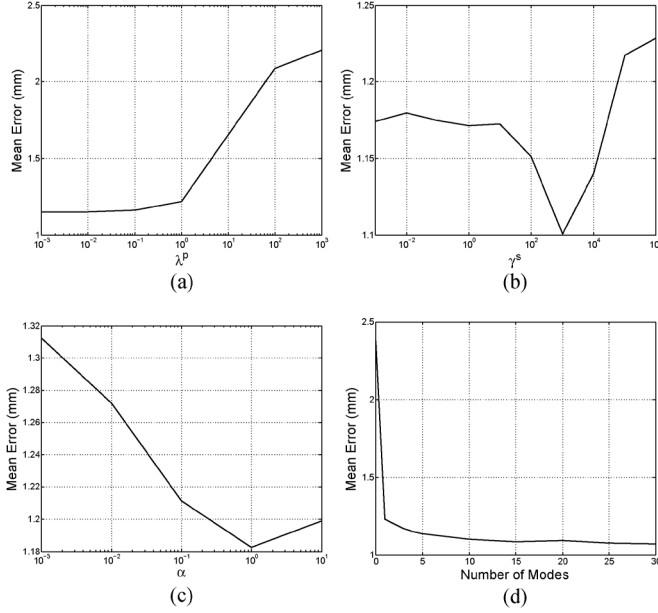


Fig. 5. Mean distance error of the registration for different values of the trade-off variables. Registration results are averaged across all lumbar vertebrae. (a) Pose trade-off. (b) Shape trade-off. (c) Deformation trade-off. (d) Number of modes.

TABLE III
VALUES SELECTED FOR THE TRADE-OFF VARIABLES OF THE REGISTRATION OF THE MULTI-VERTEBRAE MODEL TO THE CT IMAGES

Variable	Description	Value
γ^P	Constrains the coefficients for pose modes	0
γ^S	Constrains the coefficients for shape modes	10^3
α	Constrains the smoothness of the non-rigid transformation	1
c	Number of modes used in the registration	10
M	Number of target points	20,000

the pose trade-off variable, γ^P , and interestingly the registration error remained below 1.2 mm for $\gamma^P < 1$, meaning that this constraint does not necessarily improve the segmentation accuracy.

Number of modes. Confirming the generalization capability of the model [see Fig. 4(c)], the more modes, c , used in the registration, the more accurate the segmentation becomes. However, for $c > 10$ the error remains almost the same (approximately 1.1 mm).

It should be noted that the results in Fig. 4(c) and Fig. 5(d) are not necessary the same since experiments in this part are performed on only a subset (five CT images) of our training set and the registration to CT images has an extra nonrigid transformation.

Table III gives an overview of all variables with a short description and their chosen values for the segmentation. We selected the number of target points to be $M = 20000$. A larger number gives higher accuracy because it provides a more detailed description of the target. For faster but subsequently lower accurate segmentation, one should decrease this number (see Section IV).

2) *Segmentation Accuracy:* Leave-one-out experiments were performed on 27 CT images to assess the accuracy of

the segmentation. The center of mass of the L3 vertebra in the model and in the manually segmented CT image were initially aligned and then the registration was performed. The statistics of the distance error for each vertebra are detailed in Table IV. Additionally, the errors for some of the subregions including the articular process (part of the vertebra that surrounds the facet joint) are calculated and are presented in Table V. The errors are almost equally distributed in all extremities such as the spinous process, transverse process and articular process (< 2 mm). Moreover, the vertebral body shows much smaller error than the other regions. The error reaches a minimum for the middle vertebra (L3). Examples of the segmentation results are shown in Fig. 6.

The additional nonrigid transformation slightly improves the segmentation accuracy (see Fig. 7 as an example). The mean distance error for registration without the nonrigid transformation is 1.65 ± 0.44 mm, compared to a distance error of 1.38 ± 0.56 mm achieved with the nonrigid transformation. The improvement in registration error is statistically significant ($p < 0.05$; using a paired t-test). The small difference between these errors confirms that the additional deformable registration step results in submillimeter changes in segmented boundary achieved with the statical shape+pose model alone.

3) *Capture Range:* Leave-one-out experiments were performed to measure the sensitivity of the algorithm to the initial point selection, i.e., a point around the center of L3, which is marked by the user. Initially, the center of mass of the L3 vertebra in the model and the L3 vertebra in the manually segmented CT image were aligned. Next, a displacement ranging from 0 to 20 mm, in a random direction, was added to the model, followed by registration. Initial displacements were divided into bins with 2 mm width. For each bin, 27 experiments (one experiment per subject) were performed. The mean distance error against the initial displacement is shown in Fig. 8(a). The percentage of the registrations with mean distance error below 1.5 mm and 2 mm are shown in Fig. 8(b). The capture range is ± 8 mm which covers a reasonable area (16 mm) within the vertebrae, given that the inter-spinous distance is around 33 mm for lumbar spine [39]. For images with low resolution, the selection of this point might become challenging since the capture range will be only few voxels. The current unoptimized MATLAB code running on an Intel Xeon X5650 2.67 GHz, requires 15 min. This includes the initial segmentation of the CT images with Canny edge detection and the entire registration pipeline.

IV. DISCUSSION AND CONCLUSION

To the best of our knowledge, this is the first time that a statistical multi-vertebrae model is created and registered to CT images of spine. Our approach enjoys joint representation of all vertebrae by embedding statistics derived from curvature and shape of spinal column into a single model. Such representation reduces the complexity of the segmentation by combining many steps such as detection and identification of the individual vertebrae. Our approach makes two assumptions on the task of spinal column segmentation. First, the shapes of the lumbar vertebrae are strongly correlated and can be learned together. Therefore,

TABLE IV
SEGMENTATION RESULT DEMONSTRATED BY THE DISTANCES BETWEEN THE FITTED MODEL AND THE MANUAL SEGMENTATION. MEAN, RMS AND MAXIMUM (REFERRED TO AS HAUSDORFF) OF THESE DISTANCES ARE PRESENTED

	L1	L2	L3	L4	L5	All
Median	1.01±0.31	0.93±0.25	0.94±0.32	1.04±0.40	1.34±0.65	1.05±0.43
95 Percentile	3.67±1.28	3.23±1.08	3.26±1.25	3.79±1.53	5.04±2.09	3.80±1.61
Haussdorf (Maximum)	9.07±2.27	8.97±3.00	7.87±1.64	8.41±1.95	10.22±2.53	8.91±2.42
Mean	1.33±0.31	1.19±0.33	1.21±0.41	1.37±0.53	1.78±0.80	1.38±0.56
RMS	1.79±0.58	1.59±0.46	1.59±0.46	1.83±0.70	2.41±1.00	1.84±0.74

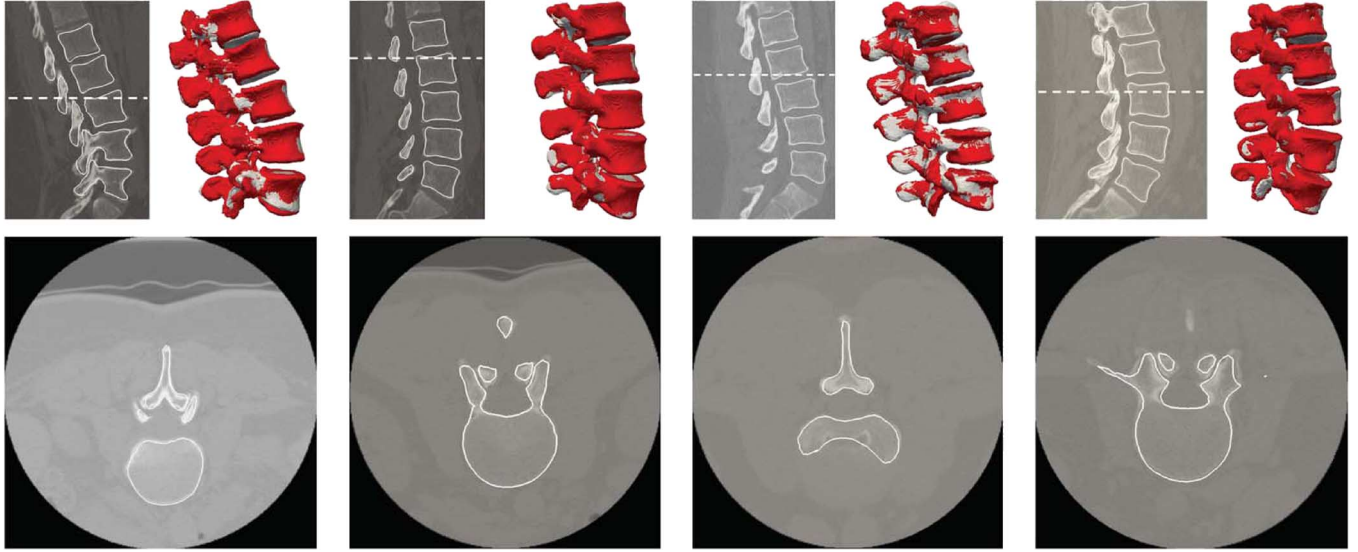


Fig. 6. Registration results and volume overlaps for four chosen volumes from different views. Registered model is highlighted in white. Dashed lines mark where the axial slices are located in the sagittal view. Red surface shows the manual segmentation.

TABLE V
MEAN ERROR (\pm STANDARD DEVIATION) OF THE REGISTRATION IN DIFFERENT REGIONS FOR EACH VERTEBRAE (SP, SPINOUS PROCESS; TP, TRANSVERSE PROCESSES; AP, ARTICULAR PROCESSES; VB, VERTEBRAL BODY)

	SP (mm)	TP (mm)	AP (mm)	VB (mm)
L1	1.96±1.59	2.27±1.22	1.49±0.82	0.69±0.35
L2	1.72±1.36	1.73±0.66	1.28±0.67	0.75±0.36
L3	1.17±0.63	1.37±0.71	1.44±0.87	0.71±0.34
L4	1.85±1.36	2.03±0.91	1.90±1.00	0.73±0.49
L5	2.35±1.17	2.48±1.31	2.51±1.59	0.68±0.30
All	1.81±1.22	1.97±0.96	1.72±0.99	0.71±0.36

since the shapes of all vertebrae are learnt together, no overlapping should be observed if the coefficients of the shape variations stay within acceptable range, i.e., $3\sqrt{\lambda}$. Although we allow for larger coefficients in our algorithm, we did not observe any overlapping of neighboring vertebrae in our segmentation technique. Second, the shape and the relative position/orientation of the vertebrae are not correlated since the latter is more related to the patient posture during data acquisition.

Our registration technique is inspired by an EM approach [28] where soft correspondences, established between two point sets, facilitate a robust alignment. Its superior performance to other surface-to-surface registration has been well explored [28], [33], [40].

Our method successfully segments the entire lumbar vertebrae with a mean point-to-surface error of 1.38 ± 0.56 mm. The surface error around the target facet joint/articular process is 1.72 mm, which is below the clinically accepted error [27].

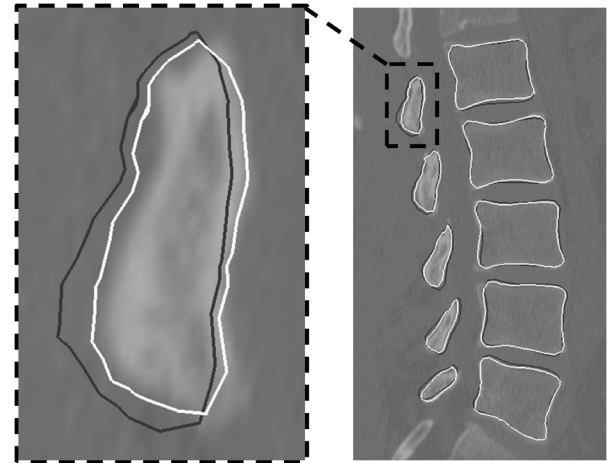


Fig. 7. Registered model with (white color) and without (gray color) the additional nonrigid transformation step. This example shows a location where the errors are especially large in order to best demonstrate the additional deformation step.

Error for the other critical subregions of the spine is also below 2 mm.

Our algorithm requires three regularization parameters to constrain the shape and the pose variations, and the nonrigid transformation smoothness. Although the segmentation accuracy shows a global optimum by changing each parameter, it also exhibits a lack of sensitivity to selection of these parameters. The mean surface error remains below 2.5 mm for a wide range of values (see Fig. 5).

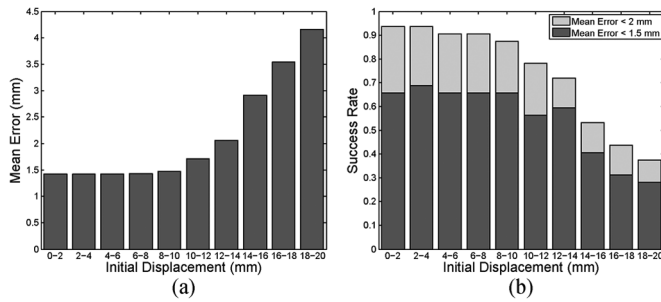


Fig. 8. (a) Mean error and (b) success rate for registration with different initial displacements. Registration results are averaged across all lumbar vertebrae.

The most similar work to what is presented in our paper is performed by Klinder *et al.* [6] with few key differences. Our approach unifies the detection, identification, and segmentation into a single step by using the shape prior and the pose statistics embedded in our model. Additionally, our method requires an initialization that is a single point marked by the user close to the center of the mass of the L3 vertebra. On the other hand, Klinder's method is fully automatic, but requires a few extra steps for detection, identification, and extraction of vertebrae. Moreover, they do not take advantage of the shape statistics, and instead only perform a nonrigid registration of a template of a single vertebra to its isolated sub-volume. It should be noted that our achieved segmentation error is worse (1.38 mm compared to 0.76 mm); however, both algorithms produce segmentation errors that are within the clinically acceptable range for facet joint injections. Additionally, Klinder *et al.* reported a misidentification of 5% whereas we did not observe any in our experiments. Our technique is also similar to a segmentation algorithm recently proposed by Kadoury *et al.* [17] since they represent the relative position of the neighboring vertebrae on a manifold. However, they do not apply any statistical analysis on pose information but find a training subject with the most similar spine curvature to the target. In summary, our segmentation results are better in terms of accuracy (1.38 mm compared to 2.5 mm) and identification of the vertebrae since they reported a misidentification of 12%. We have also evaluated our algorithm on larger dataset (32 patients compared to 20 patients). Generally, the segmentation error achieved with our method is also comparable to the results reported in the literature which have the range of 0.95–1.5 mm [12], [17], [41].

The most computationally expensive component of our algorithm is the calculation of the likelihood function [8]. Therefore, the runtime is proportional to the number of target points. On the other hand, the accuracy decreases when fewer target points are used. Fig. 9 shows the relation between the accuracy, computation time, and the number of target points. The runtime is not directly comparable to the Klinder *et al.* approach, because of differences in coding, implementation, hardware and number of image slices. The detection and identification steps in their algorithm takes around 37 min, and final fine-tuning of the segmentation takes 179 s. Conversely, our method uses a single point to manually identify a vertebra, and our mean error is still clinically acceptable (i.e., 2 mm) for run-times as low as 1 min. Previous work suggested that more than 12 modes are

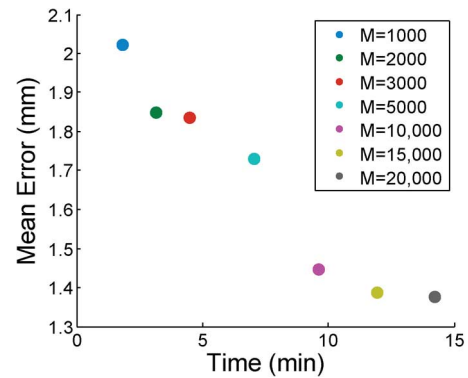


Fig. 9. Time versus accuracy for different number of CT points used in the registration. Time is calculated from an unoptimized MATLAB implementation.

needed to represent 95% of the variations for a single vertebra [42]. Given that, at least 60 modes should be used to represent the shape variations of five vertebrae. For the task of the registration, one must also compute the position and orientation of each vertebra (six parameters) which increases the number of parameters to 90 ($60 + 6 \times 5$). On the contrary, by using a multi-vertebrae shape+pose model, 95% of the shape variations are captured by only 25 shape modes [see Fig. 4(a)], and first 15 pose modes represent the entire possibilities of the relative position and orientations of the vertebrae [see Fig. 4(b)]. In our registration, however, we only used the first 10 modes for both shapes and poses since the additional nonrigid transformation captured the small variations. The additional nonrigid transformation step slightly improves the segmentation accuracy. This involves small variations that are not included in our statistical model. However, segmentation of pathological cases such as fractured/broken bone cannot be recovered with this transformation. Additionally, we believe that enlarging the training set with patients with abnormalities in their spine curvatures (such as scoliosis or kyphosis) will improve the segmentation results for such cases. We should also note that the training set is based on only 32 CT images which may not reflect the entire variability of L1–L5 vertebrae. A larger dataset is expected to reduce errors further.

Future work will involve extension of our point-based model to an *appearance model* [43]. This may improve the results given the large inhomogeneity in the spine intensity in the CT images. We will also focus on the integration of the proposed methodology into the clinical workflow and extension of this model to thoracic and cervical vertebrae. The latter is challenging since the shapes of vertebrae may not be correlated and a larger number of vertebrae are included in the model. We investigate the possibility of the automation of the model initialization and reducing the computational effort by implementing the code on a graphical processing unit. We also plan to exploit the inherent quantitative joint statistics of the vertebral model for detection of spine abnormalities that are exhibited through geometric deviations. In order to further improve accuracy, we plan to integrate alternative frameworks for statistical analysis, such as probabilistic methodologies proposed by Changizi and Hamarneh [44].

APPENDIX

BRIEF REVIEW OF LIE GROUPS AND SIMILARITY TRANSFORMATIONS

A Lie group G is a group and also a differentiable manifold where multiplication and inversion are smooth. Tangent space exists for every member $x \in G$ and is denoted as $T_x G$. The tangent space at the identity element, e , is called Lie algebra, \mathfrak{g} . The exponential map, $\mathfrak{g} : \exp(x) \rightarrow G$, and its inverse logarithm map $G : \log(x) \rightarrow \mathfrak{g}$, are used to map elements in the tangent space into G and vice versa.

The distance between two elements in a Lie group is defined as the length of the geodesic (the shortest path) between two elements and can be found using the log map

$$d(x, y) = \|\log(x^{-1}y)\|. \quad (15)$$

One can define the Fréchet expectation of an empirical mean value of a set of elements, x_1, \dots, x_n

$$\mu = \arg \min_{\mu} \left(\frac{1}{n} \sum_i d(\mu, x_i)^2 \right). \quad (16)$$

The result of the minimization may not exist or not be unique. However, for elements with bounded variance, the Fréchet mean exists and can be computed iteratively as is suggested by Pennec [45]

$$\mu_k = \mu_{k-1} \exp \left(\frac{1}{n} \sum_i \log(\mu_{k-1}^{-1} x_i) \right). \quad (17)$$

Analogous to principal components in the Euclidean space, PG are defined. They can be found using a minimization approach or may be approximated. The approximation, suggested by Fletcher *et al.* [26], is as follows: for a set of elements, x_1, \dots, x_n , the mean, μ , is found. The PCA is then applied to the residuals in the Lie algebra, $\log(\mu^{-1} T_i)$. The results are orthonormal principal components, v_l , which give the PGs by the exponential mapping $\mu \exp(v_l)$.

Similarity transformations include three components: Rotation, R , which is an orthogonal matrix with determinant equal to 1; translation, d ; and scale, s . Such transformations can transform a 3-D point by $T(x) = sRx + d$. It can also be represented in a matrix form

$$T = \begin{bmatrix} sR & d \\ 0 & 1 \end{bmatrix} \quad (18)$$

and be applied to homogeneous coordinates of a 3-D point, $X = (x, y, z, 1)^T$, i.e., $T(X) = TX$. Exponential and logarithm maps are performed using the standard matrix exponential and matrix logarithm; e.g., the matrix exponential is defined by the series

$$\exp(T) = \sum_{k=0}^{\infty} \frac{1}{k!} T^k. \quad (19)$$

The logarithm map of a similarity transformation T results in

$$\log T = \begin{bmatrix} l & -r_z & r_y & x \\ r_z & l & -r_x & y \\ -r_y & r_x & l & z \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (20)$$

where $l = \log(s)$, (r_x, r_y, r_z) defines the rotation axis with angle $\theta = \sqrt{r_x^2 + r_y^2 + r_z^2}$, and (x, y, z) is the translation vector. The transformation can be expressed by a 7×1 vector, $\varphi_{7 \times 1} = (r_x, r_y, r_z, x, y, z, l)^T$, in the tangent space with corresponding bases

$$\begin{aligned} B_1 &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, & B_2 &= \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \\ B_3 &= \begin{bmatrix} 0 & -1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, & B_4 &= \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \\ B_5 &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, & B_6 &= \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \\ B_7 &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \end{aligned} \quad (21)$$

Given above, a similarity transformation can be written as $T = \exp(\sum \phi_k B_k)$.

ACKNOWLEDGMENT

The authors would like to also thank the support provided by the staff at Kingston General Hospital, Kingston, ON, Canada, in acquiring the CT data used in this research.

REFERENCES

- [1] M. Boswell, J. Colson, and W. Spillane *et al.*, "Therapeutic facet joint interventions in chronic spinal pain: A systematic review of effectiveness and complications," *Pain Physician*, vol. 8, no. 1, pp. 101–114, 2005.
- [2] P. Bruners, T. Penzkofer, M. Nagel, R. Elfring, N. Gronloh, T. Schmitz-Rode, R. Günther, and A. Mahnken, "Electromagnetic tracking for CT-guided spine interventions: Phantom, ex-vivo and in-vivo results," *Eur. Radiol.*, vol. 19, no. 4, pp. 990–994, 2009.
- [3] J. Moore, C. Clarke, D. Bainbridge, C. Wedlake, A. Wiles, D. Pace, and T. Peters, "Image guidance for spinal facet injections using tracked ultrasound," *Med. Image Comput. Comput.-Assist. Intervent.*, vol. 5761, pp. 516–523, 2009.
- [4] E. C. S. Chen, P. Mousavi, S. Gill, G. Fichtinger, and P. Abolmaesumi, "Ultrasound guided spine needle insertion," in *SPIE Medical Imaging*, 2010, vol. 7625, pp. 762 538-1–762 538-8.
- [5] C. Yeo, T. Ungi, A. Lasso, R. McGraw, and G. Fichtinger *et al.*, "The effect of augmented reality training on percutaneous needle placement in spinal facet joint injections," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 7, pp. 2031–2037, Jul. 2011.
- [6] T. Klinder, J. Ostermann, M. Ehm, A. Franz, R. Kneser, and C. Lorenz, "Automated model-based vertebra detection, identification, and segmentation in CT images," *Med. Image Anal.*, vol. 13, no. 3, pp. 471–482, 2009.
- [7] S. S. C. Burnett, G. Starkschall, C. W. Stevens, and Z. Liao, "A deformable-model approach to semi-automatic segmentation of CT images demonstrated by application to the spinal canal," *Med. Phys.*, vol. 31, no. 2, pp. 251–263, 2004.

- [8] A. Mastmeyer, K. Engelke, C. Fuchs, and W. A. Kalender, "A hierarchical 3-D segmentation method and the definition of vertebral body coordinate systems for QCT of the lumbar spine," *Med. Image Anal.*, vol. 10, no. 4, pp. 560–577, 2006.
- [9] S. Schmidt, J. Kappes, M. Bergholdt, V. Pekar, S. Dries, D. Bystrov, and C. Schnoerr, "Spine detection and labeling using a parts-based graphical model," *Inf. Process. Med. Imag.*, vol. 4584, pp. 122–133, 2007.
- [10] H. Shen, A. Litvin, and C. Alvino, "Localized priors for the precise segmentation of individual vertebrae from CT volume data," *Med. Image Comput. Comput.-Assist. Intervent.*, vol. 5241, pp. 367–375, 2008.
- [11] T. Vrtovec, B. Likar, and F. Pernus, "Automated curved planar reformation of 3-D spine images," *Phys. Med. Biol.*, vol. 50, pp. 4527–4540, 2005.
- [12] J. Ma, L. Lu, Y. Zhan, X. Zhou, M. Salganicoff, and A. Krishnan, "Hierarchical segmentation and identification of thoracic vertebra using learning-based edge detection and coarse-to-fine deformable model," *Med. Image Comput. Comput.-Assist. Intervent.*, vol. 6361, pp. 19–27, 2010.
- [13] Z. Peng, J. Zhong, W. Wee, and J. Lee, "Automated vertebra detection and segmentation from the whole spine MR images," in *Proc. IEEE Int. Conf. Eng. Med. Biol.*, 2005, pp. 2527–2530.
- [14] B. M. Kelm, S. K. Zhou, M. Suehling, Y. Zheng, M. Wels, and D. Comaniciu, "Detection of 3-D spinal geometry using iterated marginal space learning," in *Medical Computer Vision. Recognition Techniques and Applications in Medical Imaging*. New York: Springer, 2011, pp. 96–105.
- [15] B. Glocker, J. Feulner, A. Criminisi, D. R. Haynor, and E. Konukoglu, "Automatic localization and identification of vertebrae in arbitrary field-of-view CT scans," in *Medical Image Computing and Computer-Assisted Intervention*. New York: Springer, 2012, pp. 590–598.
- [16] Y. Zhan, D. Maneesh, M. Harder, and X. S. Zhou, "Robust MR spine detection using hierarchical learning and local articulated model," in *Medical Image Computing and Computer-Assisted Intervention*. New York: Springer, 2012, pp. 141–148.
- [17] S. Kadoury, H. Labelle, and N. Paragios, "NCE of articulated spine models in CT images using high-order Markov random fields," *Med. Image Anal.*, vol. 15, no. 4, pp. 426–437, 2011.
- [18] M. Roberts, T. Cootes, E. Pacheco, T. Oh, and J. Adams, "Segmentation of lumbar vertebrae using part-based graphs and active appearance models," *Med. Image Comput. Comput.-Assist. Intervent.*, pp. 1017–1024, 2009.
- [19] D. Štern, B. Likar, F. Pernuš, and T. Vrtovec, "Parametric modelling and segmentation of vertebral bodies in 3-D CT and MR spine images," *Phys. Med. Biol.*, vol. 56, no. 23, p. 7505, 2011.
- [20] M. Bossa and S. Olmos, "Multi-object statistical pose+shape models," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2007, pp. 1204–1207.
- [21] N. Duta and M. Sonka, "Segmentation and interpretation of MR brain images. An improved active shape model," *IEEE Trans. Med. Imag.*, vol. 17, no. 6, pp. 1049–1062, Dec. 1998.
- [22] H. Wang, D. Stout, and A. Chatzioannou, "Estimation of mouse organ locations through registration of a statistical mouse atlas with Micro-CT images," *IEEE Trans. Med. Imag.*, vol. 31, no. 1, pp. 88–102, Jan. 2012.
- [23] A. Tsai, W. Wells, C. Tempany, E. Grimson, and A. Willsky, "Coupled multi-shape model and mutual information for medical image segmentation," *Inf. Process. Med. Imag.*, vol. 2732, pp. 185–197, 2003.
- [24] J. Cerrolaza, A. Villanueva, and R. Cabeza, "Hierarchical statistical shape models of multiobject anatomical structures: Application to brain MRI," *IEEE Trans. Med. Imag.*, vol. 31, no. 3, pp. 713–724, Mar. 2012.
- [25] C. Lu, S. M. Pizer, S. Joshi, and J. Jeong, "Statistical multi-object shape models," *Int. J. Comput. Vis.*, vol. 75, no. 3, pp. 387–404, 2007.
- [26] P. Fletcher, C. Lu, and S. Joshi, "Statistics of shape via principal geodesic analysis on lie groups," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, 1, 2003, pp. 95–101.
- [27] S. Gill, P. Abolmaesumi, G. Fichtinger, J. Boisvert, D. Pichora, D. Borschneck, and P. Mousavi, "Biomechanically constrained groupwise ultrasound to CT registration of the lumbar spine," *Med. Image Anal.*, vol. 16, no. 3, pp. 662–674, 2012.
- [28] A. Rasoulouli, R. Rohling, and P. Abolmaesumi, "Group-wise registration of point sets for statistical shape models," *IEEE Trans. Med. Imag.*, vol. 31, no. 11, pp. 2025–2034, Nov. 2012.
- [29] I. Dryden and K. Mardia, *Statistical Shape Analysis*. New York: Wiley, 1998, vol. 4.
- [30] C. J. Twining and C. J. Taylor, "Kernel principal component analysis and the construction of non-linear active shape models," in *Br. Mach. Vis. Conf.*, 2001, vol. 1, pp. 23–32.
- [31] M. Bossa and S. Olmos, "Statistical linear models in Procrustes shape space," in *MICCAI Workshop Math. Foundat. Computat. Anat.: Geomet., Stat. Registrat. Methods Model. Biol. Shape Variabil.*, 2006, pp. 102–111.
- [32] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 8, no. 6, pp. 679–698, Nov. 1986.
- [33] A. Myronenko, X. Song, and M. Carreira-Perpinan, "Non-rigid point set registration: Coherent point drift," *Adv. Neural Inf. Process. Syst.*, pp. 1009–1016, 2007.
- [34] P. A. Yushkevich, J. Pivenc, H. C. Hazlett, R. G. Smith, S. Hob, J. C. Gee, and G. Gerig, "User-guided 3-D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability," *Neuroimage*, vol. 31, no. 3, pp. 1116–1128, 2006.
- [35] W. Lorensen and H. Cline, "Marching cubes: A high resolution 3-D surface construction algorithm," *ACM Siggraph Comput. Graph.*, vol. 21, no. 4, pp. 163–169, 1987.
- [36] M. Styner, K. Rajamani, L. Nolte, G. Zsemlye, G. Szekely, C. Taylor, and R. Davies, "Evaluation of 3-D correspondence methods for model building," *Inf. Process. Med. Imag.*, vol. 2732, pp. 63–75, 2003.
- [37] M. Pathria, D. Sartoris, and D. Resnick, "Osteoarthritis of the facet joints: Accuracy of oblique radiographic assessment," *Radiology*, vol. 164, no. 1, pp. 227–230, 1987.
- [38] M. Greher, L. Kirchmair, B. Enna, P. Kovacs, B. Gustorff, S. Kapral, and B. Moriggl, "Ultrasound-guided lumbar facet nerve block: Accuracy of a new technique confirmed by computed tomography," *Anesthesiology*, vol. 101, no. 5, pp. 1195–1200, 2004.
- [39] H. Rafii-Tari, P. Abolmaesumi, and R. Rohling, "Panorama ultrasound for guiding epidural anesthesia: A feasibility study," *Inf. Process. Comput.-Assist. Intervent.*, vol. 6689, pp. 179–189, 2011.
- [40] H. Chui and A. Rangarajan, "A new algorithm for non-rigid point matching," in *IEEE Conf. Comput. Vis. Pattern Recognit.*, 2000, vol. 2, pp. 44–51.
- [41] S. Ghosh, R. Alomari, V. Chaudhary, and G. Dhillon, "Automatic lumbar vertebra segmentation from clinical CT for wedge compression fracture diagnosis," in *Proc. SPIE Med. Imag.*, 2011, vol. 7963, p. 796303.
- [42] S. Khalaghi, P. Mousavi, R. Gong, S. Gill, J. Boisvert, G. Fichtinger, D. Pichora, D. Borschneck, and P. Abolmaesumi, "Registration of a statistical shape model of the lumbar spine to 3-D ultrasound images," *Med. Image Comput. Comput.-Assist. Intervent.*, vol. 6362, pp. 68–75, 2010.
- [43] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 6, pp. 681–685, Jun. 2001.
- [44] N. Changizi and G. Hamarneh, "Probabilistic multi-shape representation using an isometric log-ratio mapping," *Med. Image Comput. Comput.-Assist. Intervent.*, vol. 6363, pp. 563–570, 2010.
- [45] X. Pennec, "Intrinsic statistics on Riemannian manifolds: Basic tools for geometric measurements," *J. Math. Imag. Vis.*, vol. 25, pp. 127–154, 2006.