35th Annual International Conference of the IEEE EMBS
Osaka, Japan, 3 - 7 July, 2013

Fast and Robust 3D Vertebra Segmentation using Statistical Shape Models

Hengameh Mirzaalian[1,*] , Michael Wels[2,*], Tobias Heimann[3], B. Michael Kelm[3], Michael Suehling[3]

Abstract— **We propose a top-down fully automatic 3D vertebra segmentation algorithm using global shape-related as well as local appearance-related prior information. The former is brought into the system by a global statistical shape model built from annotated training data, i.e., annotated CT volumes. The latter is handled by a machine learning-based component, i.e., a boundary detector, providing a strong discriminative model for vertebra surface appearance by making use of local context-encoding features. This boundary detector, which is essentially a probabilistic boosting-tree classifier, is also learnt from annotated training data. Contextual information is taken into account by representing vertebra surface candidate voxels with high-dimensional vectors of 3D steerable features derived from the observed volume intensities. Our system does not only consider the body of the individual vertebrae but also the spinal processes. Before segmentation, the image parts depicting individual vertebrae are spatially normalized with respect to their bounding box information in terms of translation, orientation, and scale leading to more accurate results. We evaluate segmentation accuracy on 7 CT volumes each depicting 22 vertebrae. The results indicate a symmetric point-to-mesh surface error of $1.37 \pm 0.37\ mm$, which matches the current state-of-the-art.**

I. Introduction

Accurate vertebra segmentation in computed tomography (CT) images is important for numerous medical applications, e.g., diagnosis of osteolytic or osteoblastic cancer metastases within the spinal column [10], diagnosis of lung nodules [12], and detection of osteoporosis [1]. Accurate knowledge of the shape of the individual vertebrae is also important for spinal biopsies, implants, or the insertion of pedicle screws in spinal surgery [11]. However, manually delineating and annotating vertebrae is a subjective, tedious, and error prone task. Preparing an automatic vertebra segmentation system would greatly improve the process, thereby easing the workload on radiologists while also removing operator variability.

Automatic vertebra segmentation in CT images is a challenging task due to the presence of image artifacts, contrast variations, presence of neighboring structures, and shape variation [11]. Recently, a considerable amount of work has been done toward preparing automatic systems for detection and segmentation of vertebrae. In this work, we mainly focus on the segmentation task (rather than detection). In the following, we start by preparing a brief review on the state-of-the-art regarding the segmentation step. In Sec. III, we explain in detail how we extract statistical shape information and train the boundary detector using annotated data (Sec. IV). In Sec. V, we prepare qualitative and quantitative results of our method and compare them with the results achieved by a standard graph cut (GC)-based vertebra segmentation approach and a GC-based approach considering a shape prior [1].

II. Previous work

We classify the existing vertebra segmentation approaches to two main groups: i) the ones, which do not consider shape prior information, and ii) the ones, which do.

Regarding the first group, we can point to the following works: Ghosh et al. [5] extract the vertebra border as high gradient edges. Peng et al. [14] apply the Canny edge detector on 2D slices for vertebra segmentation. Aslan et al. [2] utilize a level set algorithm for vertebra segmentation. However, these methods [5], [14], [2] do not make use of shape prior knowledge. Therefore, they are vulnerable to leakage and thus lead to less accurate segmentation results.

Considering the second group, there exist several vertebra segmentation methods which make use of shape prior information. Aslan et al. [1] consider shape prior information in a graph cut-based framework. Ma et al. [12] propose a template-based segmentation method. However, these methods only rely on mean shape information and do not benefit from the principal modes of variation. Herring et al. [8] compute a coarse segmentation by simple thresholding and then register it to a pre-computed vertebra shape model. However, their method requires a manual initialization; similar to the works in [13] and [18]. Klinder et al. [11] propose a model-based segmentation approach using a region-based appearance model, which includes variance information.

III. Method

The general pipeline of our method is shown in Fig. 1. The input to our system is a 3D CT image of the spinal cord accompanied with vertebra-bounding box information. These bounding boxes, which are represented by their center, orientation, and scale, can be estimated by applying a vertebra body detection as proposed by Kelm et al. [10]. Combining our method with this method would lead to a fully automatic vertebra detection and segmentation system,

[1]Medical Image Analysis Lab, Simon Fraser University, Vancouver, Canada.
[2]Siemens AG, Healthcare, Imaging & Therapy Systems, Computed Tomography & Radiation Oncology, Research & Development, Forchheim, Germany.
[3]Siemens AG, Corporate Technology, Research & Technology Center, Imaging & Computer Vision, Erlangen, Germany.
*H. Mirzaalian and M. Wels contributed to this work when they were with Siemens AG, Corporate Technology, Research & Technology Center, Imaging & Computer Vision, Erlangen, Germany.

978-1-4577-0216-7/13/$26.00 ©2013 IEEE

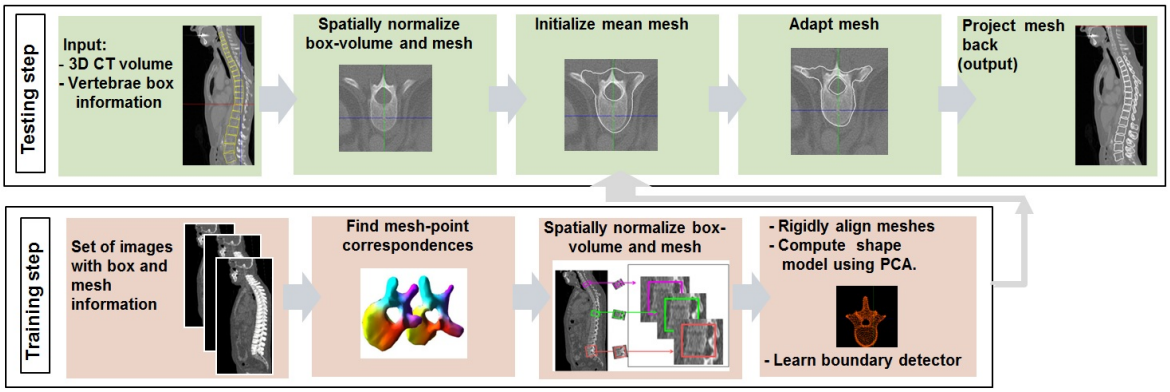


Fig. 1. The block diagram of our proposed 3D vertebra segmentation method.

which does not require any user interaction. Our method combines statistical shape modeling (SSM) to capture global vertebra shape information and machine learning (ML) to capture local appearance-related prior information. We break down our method into two main steps: the training step and the testing step. In the training step, we compute the SSM and the boundary detector model. In the testing step, we make use of the trained models resulting from the previous step to segment vertebrae in an unseen image accompanied with its vertebra bounding box information.

A. Training Step

The training step of our framework consists of four main steps: i) finding the mesh point correspondences, ii) normalizing the meshes and volumes, iii) extracting the SSM, and iv) learning the boundary detector. Note that we do the learning step on cervical (V_1 to V_5), thoracic (V_6 to V_{17}), and lumbar (V_{18} to V_{22}) parts separately, where V_i represents the vertebra number. In Fig. 2(b) the vertebrae inside the cervical, thoracic, and lumbar parts are represented by pink, green, and orange, respectively.

1) Finding Mesh Point Correspondences: Since extracting a SSM requires a set of training shapes with well-defined correspondences [7], we apply a spectral-based algorithm to compute these correspondences between the vertebra meshes [9]. In the respective block in Fig. 1, corresponding points between a pair of meshes are represented with the same color. Note that in our implementation of finding mesh-point correspondences, we use vertebrae V_3, V_{12}, and V_{20} of one patient in the training set as the reference meshes for cervical, thoracic, and lumbar parts and register all other meshes of each group to them.

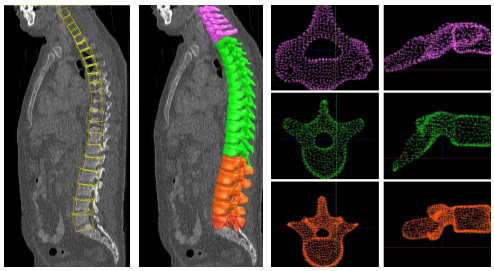
2) Normalizing Meshes and Volumes: The next step as depicted in Fig. 1 is performing spatial normalization on the vertebra volumes and meshes. Regions within the bounding boxes are spatially normalized to image volumes with equal size, resolution, and orientation. The spatial normalization step is important in our machine learning-based approach. Extracting 3D steerable features from these normalized volumes simplifies learning due to more stable appearance patterns of the vertebra edges. A similar normalization step has been proposed by Wels et al. [16] to extract local features from vertebral bodies for spinal bone lesion detection. We apply the same normalization step (normalizing w.r.t. the box information) to the meshes. The normalized meshes are used for extracting the SSM, as explained in the following.

3) Extracting the SSM: After finding correspondences and normalizing the meshes, we apply Generalized Procrustes Analysis (GPA) [6] to align the meshes rigidly. Let us represent aligned meshes by $\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_N$, $N \in \mathbb{N}^+$, where $\mathbf{x}_i$ consists of the spatial coordinates of the surface points of the meshes. Then, the mean shape $\bar{\mathbf{x}}$ and the corresponding covariance matrix S is given by: $\bar{\mathbf{x}} = \sum_{i=1}^{N} \mathbf{x}_i/N, \quad S = \sum_{i=1}^{N} (\mathbf{x}_i - \bar{\mathbf{x}})(\mathbf{x}_i - \bar{\mathbf{x}})^T$. By applying eigendecomposition on S, we can extract principal modes of variation ϕ_m (eigenvectors) and their respective variance λ_m (eigenvalue). Based on the main concept of SSM theory, each shape in the training dataset can be approximated by a linear combination of the first m^{th} modes, i.e. given by: $\mathbf{x} \approx \bar{\mathbf{x}} + \mathbf{Pb}$, where $\mathbf{P} = (\phi_1, .., \phi_m)$ is the matrix of the m selected eigenvectors, and $\mathbf{b} = (b_1, ...b_m)^T$ are the shape parameters [4].

4) Learning the Boundary Detector: To learn the boundary detector, given the normalized meshes and volumes, the image voxels on the mesh surface are interpreted as positive training samples. Then, a set of 3D steerable features is extracted from the points on the surface [19]. The same set of features is extracted from several neighboring sampling points along the normal line of the mesh surface points providing negative training samples [19], [17]. These feature vectors are used to train the boundary detector using a Probabilistic Boosting-Tree Classifier [19], [17].

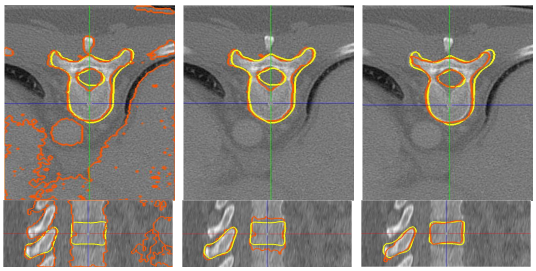
B. Testing Step

Given an unseen image with its vertebra bounding box information, we first spatially normalize the volumes inside the box. On the normalized volumes, an initial estimation of the shape of the vertebra $\mathbf{x}$ is estimated using the computed mean shape $\bar{\mathbf{x}}$, i.e., $\mathbf{x} \leftarrow \bar{\mathbf{x}}$. Then, a set of steerable features is extracted from the mesh points $\mathbf{x}$ and several neighbouring sampling points along the normal line of the mesh-surface



(a) (b) (c)

Fig. 2. (a) Example CT image with the overlaid vertebra bounding boxes (yellow boxes). (b) Overlaid vertebra meshes resulting from our method. The vertebrae inside the cervical, thoracic, and lumbar parts are represented by pink, green, and orange, respectively. Computed mean shapes of vertebrae in: cervical, thoracic, and lumbar groups are shown in (c).



(a) GC (b) GC+SP (c) Ours

Fig. 3. Example detected mesh resulting from: (a) GC [3], (b) GC with shape prior [1], and (c) our method. The groundtruth and segmented meshes are represented by yellow and orange color, respectively. The first and second rows represent axial and sagittal views, respectively.

points. After applying the boundary detector to the extracted feature vectors, $\mathbf{x}$ is updated by a displacement vector $\Delta\mathbf{x} = (\Delta x_1, \Delta y_1, \Delta z_1, ..., \Delta x_n, \Delta y_n, \Delta z_n)^T$. To apply shape constraints on the updated mesh $\mathbf{x} \leftarrow \mathbf{x} + \Delta\mathbf{x}$, it is registered to the SSM model space and projected such that it can be approximated by the mean shape and a linear combination of eigenvectors [4] (see Sec. III-A.3).

As shown in Fig. 1, the final estimation of all the vertebrae in the original image space is made by projecting back the detected meshes in the normalized space to the original image space.

IV. Material

We evaluate our method on CT images of 7 patients. Each image depicts 22 vertebrae. The plain resolution of the images varies between 0.78 and 0.97 mm with a slice thickness between 2.0 to 2.5 mm. Images are of size 512x512 voxels in-plain with 296 to 486 voxels in z direction. Groundtruth meshes were prepared manually.

In total, there were 154 spatially normalized single vertebra volumes available for shape model generation and training boundary detectors. Our system is implemented in C++ and the experiments have been carried on an Intel Core 2Duo CPU (2.6 GHz). Parts of our system were rapidly prototyped using an Integrated Detection Network (IDN) [15].

V. Results

We perform leave-one-patient-out cross validation (CV) on the data collection: given the 6 CT images in each CV-training, we compute 3 shape models for cervical, thoracic, and lumbar vertebrae. Examples of the computed mean shapes are shown in Fig. 2(c). Afterwards, we train boundary detectors on the cervical, thoracic, and lumbar vertebrae. Finally, we apply our segmentation method on the image, which was excluded from the training dataset. Examples of segmented vertebrae resulting from our method are shown in Fig. 2(b).

A qualitative comparison between the segmented meshes resulting from our method compared with the standard graph cut (GC) algorithm [3], and the GC-based algorithm considering a shape prior [1], is shown in Fig. 3. Using a shape prior with GC significantly reduces leaking as compared to plain GC. The more precise shape model employed in our work, however, yields even better results.

For a quantitative evaluation, we measure the error of the segmented meshes compared with the groundtruth meshes in terms of the symmetric point-to-mesh distance, which is a widely used criterion [11]. In Tab. V, we report the errors resulting from the GC-based method considering the shape prior, GC+SP [1], our SSM and ML based approach without and with spatial normalization, SSM+ML and SSM+ML+Norm, respectively. As it can be seen, averaged over all the datasets, the minimum error 1.37 ± 0.37 mm is achieved by SSM+ML+Norm. Comparison between the error resulting from SSM+ML and SSM+ML+Norm highlights the importance of extracting feature vectors from spatially normalized volumes for the learning-based boundary detector. To measure the sensitivity of the method with respect to the box information, we measure segmentation accuracy using perturbed box information in the following range: 5 mm in xyz-translations and xyz scales, and 2° in orientation. The results in the last row of Tab. V indicate that we almost achieve the same accuracy using the perturbed box information.

Note that Klinder et al. [11] report to have 1.12 ± 1.04 mm error evaluating on their own dataset. A numerical comparison between the error resulting by our method compared with [11] indicates that our method has less standard deviation while only having a slightly higher mean error. Further, the running time of the segmentation step over the thoracic part is reported as 3 min in [11], where ours is about 2 min.

VI. Conclusion

We proposed a fully automatic 3D vertebra segmentation algorithm using statistical shape modeling and a machine learning-based boundary detector. The SSMs were extracted from annotated training data and the boundary detector were trained over a set of 3D steerable features, which were extracted from the normalized volumes. Our system considers

	Cervical	Thoracic	Lumbar	All
GC+SP	7.9± 7.9	6.3± 6.4	3.9± 5.8	6.1± 6.7
SSM+ML	2.2±0.7	1.8±0.6	1.6±0.4	1.8±0.6
SSM+ML+Norm	1.4±0.4	1.3±0.3	1.3±0.3	1.3±0.3
SSM+ML+Norm*	1.6±0.6	1.7±0.7	1.3±0.5	1.6±0.7

TABLE I

VERTEBRA SEGMENTATION ERROR IN TERMS OF THE POINT-TO-SURFACE ERROR ON 7 CT IMAGES EACH CONSISTING OF 22 VERTEBRAE (154 VERTEBRA VOLUMES IN TOTAL). THE RESULTS ARE REPORTED FOR THE GC-BASED METHOD CONSIDERING SHAPE PRIOR INFORMATION, GC+SP [1], AND OUR METHOD WITHOUT AND WITH NORMALIZATION, SSM+ML AND SSM+ML+NORM, RESPECTIVELY. NOTE THAT THE MINIMUM ERROR IS ACHIEVED BY SSM+ML+NORM. THE LAST ROW, SSM+ML+NORM*, REPRESENTS THE RESULTS WHEN WE PERTURB THE VERTEBRA BOX INFORMATION ($5\ mm$ IN XYZ-TRANSLATIONS AND XYZ SCALES, AND 2° IN ORIENTATION).

the bodies of the individual vertebrae as well as the spinal processes. The evaluation results on real data indicate that our method is compatible with the state-of-the-art. Applying the boundary detector in [19] and extracting features from the spatially normalized images lead to a more accurate method. Towards preparing a fully automatic vertebra segmentation algorithm, we aim to combine our method with the detection method in [10]. We computed three separate shape models for cervical, thoracic, and lumbar parts. To this end, we apply a spectral-based mesh point matching method [9] to find mesh point correspondences. As our future work, we aim to apply groupwise point registration algorithms [7], which may lead to more accurate correspondences and as a result more accurate SSM information.

REFERENCES

[1] M. Aslan, A. Ali, A. Farag, H. Rara, B. Arnold, and P. Xiang. 3D vertebrae body segmentation using shape based graph cuts. *IEEE ICPR*, pages 3951–3954, 2010.

[2] M. Aslan, A. Farag, B. Arnold, and X. Ping. Segmentation of vertebrae using level sets with expectation maximization algorithm. *IEEE ISBI*, pages 2010–2013, 2011.

[3] Y. Boykov and G. Funka-Lea. Graph cuts and efficient N-D image segmentation. *IJCV*, 70:109–131, 2006.

[4] T. Cootes, C. Taylor, D. Cooper, and J. Graham. Active shape models-their training and application. *Computer Vision and Image Understanding*, pages 38–59, 1995.

[5] S. Ghosh, R. Alomari, V. Chaudhary, and G. Dhillon. Automatic lumbar vertebrae segmentation from clinical CT for wedge compression fracture diagnosis. *SPIE*, pages 1–9, 2011.

[6] J. Gower. Generalized Procrustes Analysis. *Psychometrika*, pages 33–50, 1975.

[7] T. Heimann and H. Meinzer. Statistical shape models for 3D medical image segmentation: A review. *MIA*, 2009.

[8] J. Herring and B. Dawant. Automatic lumbar vertebral identification using surface-based registration. *Computers and Biomedical Research*, pages 74–84, 2001.

[9] V. Jain and H. Zhang. Robust 3D shape correspondence in the spectral domain. *IEEE Shape Modeling International*, pages 118–129, 2006.

[10] M. Kelm, K. Zhou, M. Suehling, Y. Zheng, M. Wels, and D. Comaniciu. Detection of 3D spinal geometry using iterated marginal space learning. *MICCAI: Workshop on Medical Computer Vision*, pages 96–105, 2009.

[11] T. Klinder, J. Ostermann, M. Ehm, A. Franz, R. Kneser, and C. Lorenz. Automated model-based vertebra detection identification and segmentation in CT images. *MIA*, pages 471–482, 2009.

[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. *MICCAI*, pages 19–27, 2010.

[13] A. Mastmeyer, K. Engelke, C. Fuchs, and W. Kalender. A hierarchical 3D segmentation method and the definition of vertebral body coordinate systems for QCT of the lumbar spine. *MIA*, pages 560–577, 2006.

[14] Z. Peng, J. Zhong, W. Wee, and J. Lee. Automated vertebra detection and segmentation from the whole spine MR images. *Engineering in Medicine and Biology*, pages 2527–2530, 2005.

[15] M. Sofka, K. Ralovich, N. Birkbeck, J. Zhang, and K. Zhou. Integrated Detection Network (IDN) for pose and boundary estimation in medical images. *IEEE Biomedical Imaging*, pages 294 – 299, 2011.

[16] M. Wels, M. Kelm, A. Tsymbal, M. Hammon, G. Soza, M. Suehling, A. Cavallaro, and D. Comaniciu. Multi-stage osteolytic spinal bone lesion detection from CT data with internal sensitivity control. *SPIE*, pages 831513–1–8, 2012.

[17] M. Wels, Y. Zheng, G. Carneiro, M. Huber, J. Hornegger, and D. Comaniciu. Fast and robust 3D MRI brain structure segmentation. *MICCAI*, pages 575–839, 2009.

[18] T. Whitmarsh, L. Barquero, S. Gregorio, J. Sierra, L. Humbert, and A. Frangi. Age-related changes in vertebral morphometry by statistical shape analysis. pages 30–39. 2012.

[19] Y. Zheng, A. Barbu, B. Georgescu, M. Scheuering, and D. Comaniciu. Four-chamber heart modeling and automatic segmentation for 3D cardiac CT volumes using marginal space learning and steerable features. *IEEE TMI*, pages 1668–1681, 2008.