Vertebrae Segmentation in 3D CT Images Based on a Variational Framework

Kerstin Hammernik, Thomas Ebner, Darko Stern, Martin Urschler and Thomas Pock

Abstract Automatic segmentation of 3D vertebrae is a challenging task in medical imaging. In this paper, we introduce a total variation (TV) based framework that incorporates an a priori model, i.e., a vertebral mean shape, image intensity and edge information. The algorithm was evaluated using leave-one-out cross validation on a data set containing ten computed tomography scans and ground truth segmentations provided for the CSI MICCAI 2014 spine and vertebrae segmentation challenge. We achieve promising results in terms of the Dice Similarity Coefficient (DSC) of 0.93 ± 0.04 averaged over the whole data set.

1 Introduction

Due to reduced physical activity and modern office jobs that require prolonged sitting during work hours, pathological conditions affecting the spine have become a growing problem of modern society. As most spinal pathologies are related to vertebrae conditions, the development of methods for accurate and objective vertebrae segmentation in medical images represents an important and challenging research area. While manual segmentation of vertebrae is tedious and too time consuming to be used in clinical practice, automatic segmentation may provide means for a fast and

K. Hammernik (✉) · T. Ebner · D. Stern · T. Pock
Institute for Computer Graphics and Vision, BioTechMed, Graz University of Technology, Graz, Austria
e-mail: hammernik@icg.tugraz.at

D. Stern
e-mail: stern@icg.tugraz.at

T. Pock
e-mail: pock@icg.tugraz.at

M. Urschler
Ludwig Boltzmann Institute for Clinical Forensic Imaging, Graz, Austria
e-mail: martin.urschler@cfi.lbg.ac.at

objective analysis of vertebral condition. A current state of the art method for detecting, identifying and segmenting vertebrae in computed tomography (CT) images is proposed by Klinder et al. [5]. The method is based on a complex and computationally demanding alignment of statistical shape models to the vertebrae in the image. Using the deformable surface model and training an edge detector to bone structure, Ma et al. [6] segment and identify the thoracic vertebrae in CT images. In the work of Kadoury et al. [4], the global shape representation of individual vertebrae in the image is captured with a non-linear low-dimensional manifold of its mesh representation, while local vertebral appearance is captured from neighborhoods in the manifold once the overall representation converges during the segmentation process. Ibragimov et al. [3] used transportation theory to build their landmark-based shape representations of vertebrae and game theory to align the model to a specific vertebra in 3D CT images. In this paper, we propose a method for vertebrae segmentation in 3D CT images based on a convex variational framework. In contrast to the previously proposed methods that use sophisticated vertebral models, our segmentation method incorporates only a mean shape model of vertebrae initialized in the center of the vertebral body.

2 Methods

Our vertebrae segmentation algorithm is based on two representations of a priori information, a mean shape model and a bone probability map obtained from intensity information of the input vertebra image. The main steps of our algorithm are illustrated in Fig. 1. Firstly, an intensity based prior map of the bone is estimated by comparing the intensity values to trained bone and soft tissue histograms. This resembles our learned bone prior. The vertebral mean shape is then registered to the
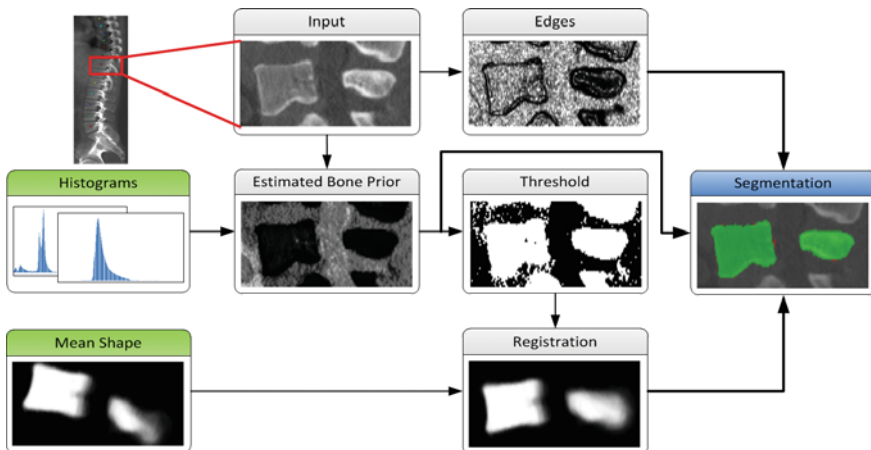
Fig. 1 Overview of our proposed algorithm. *Green boxes* represent a priori information obtained from training images. *Bold arrows* indicate parts that are included in the variational segmentation algorithm (color in online)

thresholded bone prior map to obtain the orientation of the individual vertebrae. This information is used to formulate a total variation (TV) based active contour segmentation problem, which combines the registered mean shape and the bone prior, and additionally incorporates edge information.

2.1 Mean Shape Model

The vertebral mean shape model f_s is calculated separately for three groups of vertebrae to account for variation in shape along the spine: T01–T06, T07–T12, L01–L05. Ground truth segmentations of vertebrae are registered to an arbitrary reference vertebra using an intensity-based registration with a similarity transformation and normalized cross correlation as similarity measure. The vertebral mean shape model is obtained by averaging the registered binary images of the ground truth segmented vertebrae. This step leads to a voxelwise probability for being part of the mean shape. To meet the requirements of the TV optimization framework [10], the obtained values in the probability image are inverted such that negative values represent the mean shape vertebral region and values close to one the non-vertebral region.

2.2 Bone Prior Map

The bone prior map $f_b(x) = log\left(\frac{p_{bg}(x)}{p_{fg}(x)}\right)$ is calculated as the log likelihood ratio between the probability that a voxel x belongs to the bone distribution p_{fg} and the probability that it belongs to the soft tissue distribution p_{bg}. The bone and soft tissue distributions are obtained from the training data set by estimating normalized mean foreground and background histograms of the intensity values using the ground truth segmentations. A coarse segmentation of the bone in the input image is achieved by thresholding the inverted bone map. We select a threshold value of -0.5 to ensure that trabecular bones are included in the segmented bone region, since their image intensities might be close to soft tissue.

2.3 Total Variation Segmentation

To obtain the segmented vertebra u, the following non-smooth energy functional $E_{seg}(u)$ is minimized using the first order primal-dual algorithm from [1]:

$$\min_{u\in[0,1]} E_{seg}(u) = \min_{u\in[0,1]} TV_{g,\,aniso} + \lambda_1 \int_{\Omega} u f_s \, dx + \lambda_2 \int_{\Omega} u f_b \, dx \qquad (1)$$

where Ω denotes the image domain. The trade-off between the vertebral mean shape model, bone prior map and image edge influence is regularized by the parameters λ_1 and λ_2. The term $TV_{g,\,aniso}(u)$ is the anisotropic g-weighted TV norm [7], using a structure tensor $D^{\frac{1}{2}}(x)$ as proposed by [9], incorporating both edge magnitude and

edge direction to be able to segment elongated structures:

$$\text{TV}_{g,\,aniso} = \int_{\Omega} |D^{\frac{1}{2}}(x)\nabla u|\,dx = \int_{\Omega} \sqrt{\nabla u^T D(x)\nabla u}\,dx \qquad (2)$$

$$D^{\frac{1}{2}}(x) = g(x)nn^T + n_0 n_0{}^T + n_1 n_1{}^T. \qquad (3)$$

Here, $n = \frac{\nabla I}{\|\nabla I\|}$ is the normalized image gradient, n_0 denotes an arbitrary vector in the tangent plane defined by n, and n_1 is the cross product between n and n_0. The edge function $g(x)$ is defined as

$$g(x) = e^{-\alpha\|\nabla I(x)\|^\beta}, \ \alpha, \beta \in \mathbb{R}^+. \qquad (4)$$

During minimization of the energy E_{seg}, the segmentation u tends to be foreground, if $f_b, f_s < 0$ and background, if $f_b, f_s > 0$. If f_b, f_s equal zero, the pure TV energy is minimized, thus seeking for a segmentation surface with minimal surface area. The final segmentation is achieved by thresholding the segmentation u between 0 and 1.

3 Experimental Setup

We evaluated our method on the volumetric CT data sets provided for the CSI spine and vertebrae segmentation challenge [11]. The data consists of ten training images and the corresponding ground truth segmentations. The performance of our algorithm was evaluated by a leave-one-out cross validation, i.e., we report average performance over ten experiments.

We implemented the registration as well as the segmentation algorithm on the GPU to exploit hardware parallelization of our algorithms using Nvidia CUDA. For edge detection, we chose the parameters as $\alpha = 20$ and $\beta = 0.55$. The regularization parameters are set to $\lambda_1 = 0.04$ and $\lambda_2 = 0.005$ for all experiments. We achieved the binary segmentations by thresholding the segmentation u with a value of 0.2.

4 Results

For quantitive evaluation, we used the Dice Similarity Coefficient (DSC) to evaluate our segmentation algorithm. We achieved an average DSC of 0.93 ± 0.04 over all vertebrae from the leave-one-out experiment. Our algorithm performs well on lumbar vertebrae (0.96 ± 0.02) and lower thoracic vertebrae T07–T12 (0.95 ± 0.02). The DSC for thoracic vertebrae T01–T06 is 0.89 ± 0.05, which can be explained by the influence of ribs and small intervertebral discs that are connected to the vertebrae. The algorithm did not perform well on case 6 in terms of registration errors due to confusing trabecular bone intensities with soft tissue. All estimated DSC are depicted in Table 1 and 2. A qualitative result of a correctly segmented fifth lumbar vertebra is illustrated in Fig. 2. In contrast to that, Fig. 3 shows an example for the sixth thoracic vertebra where the segmentation is influenced by connected ribs.

Table 1 Mean values and standard deviations for each vertebra resulting from the leave-one-out cross validation experiment

| DSC	T01	T02	T03	T04	T05	T06
Mean ± std	0.87 ± 0.03	0.92 ± 0.04	0.93 ± 0.03	0.91 ± 0.04	0.87 ± 0.06	0.84 ± 0.08
DSC	T07	T08	T09	T10	T11	T12
Mean ± std	0.94 ± 0.02	0.95 ± 0.02	0.95 ± 0.02	0.96 ± 0.02	0.96 ± 0.00	0.95 ± 0.03
DSC	L01	L02	L03	L04	L05	**All**
Mean ± std	0.96 ± 0.01	0.96 ± 0.01	0.96 ± 0.01	0.95 ± 0.03	0.94 ± 0.04	**0.93 ± 0.04**

We achieved good performance for T07–T12 as well as L01–L05. Our segmentations of T01–T06 have a lower value due to connected ribs and small intervertebral discs that are classified as bone in our pre-segmentation

Table 2 Evaluation in terms of mean values and standard deviations for all vertebrae in the individual data sets

DSC	case1	case2	case3	case4	case5
Mean ± std	0.93 ± 0.04	0.94 ± 0.05	0.92 ± 0.05	0.94 ± 0.03	0.94 ± 0.04
DSC	case6	case7	case8	case9	case10
Mean ± std	0.81 ± 0.18	0.94 ± 0.03	0.93 ± 0.03	0.95 ± 0.03	0.94 ± 0.04
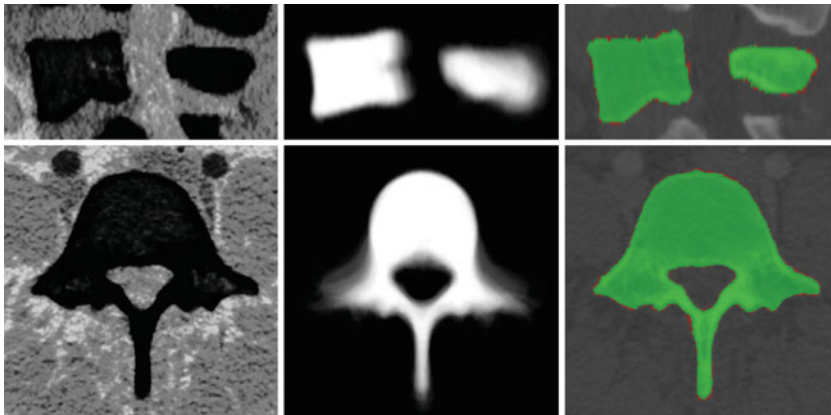
Fig. 2 The main steps of our proposed segmentation algorithm are bone prior estimation, mean shape registration and the final segmentation. The *top row* shows mid-sagittal cross sections and the *bottom row* axial cross sections of the lumbar vertebra L05. *Dark* regions in the bone prior (*left image*) are likely to be bone. The registered mean shape (*middle image*) and the bone prior are used as additional information for the segmentation algorithm. The final segmentation is depicted in the *right image* where *green* regions are correctly segmented and *red* regions differ from the ground truth result. The DSC for this example is 0.97 (color in online)
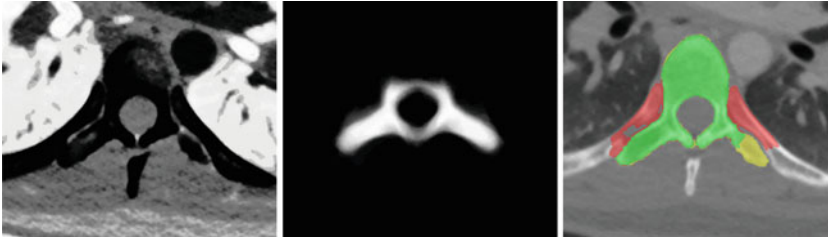
Fig. 3 This example illustrates axial cross sections of T06 where we achieved a DSC of 0.84. The bone prior shown in the *left image* classifies ribs as bone tissue. Therefore, the registered mean shape (*middle image*) is wrongly aligned, as it is attracted by the ribs. The final segmentation (*right image*) contains wrongly segmented ribs illustrated in *red*, while the yellow parts of the vertebra are missing. The *green* area depicts the correctly segmented region (color in online)

5 Discussion

Our proposed method for vertebrae segmentation based on a variational framework has been successfully applied to ten volumetric CT data sets provided for the CSI spine and vertebrae segmentation challenge. A common problem in vertebrae segmentation is that edges of the vertebrae are not clearly defined. Furthermore, trabecular bone intensities sometimes resemble soft tissue. These limitations could be easily overcome with our approach by adding manual user constraints as proposed in [8]. We see this as a great benefit of using our variational framework in vertebrae segmentation. In contrast to other methods that do not guarantee the convergence to an optimal solution, our proposed TV energy minimization (1) is convex, hence, it yields a globally optimal solution given the successfully registered mean shape. Initialized with a single point in the center of the vertebra, the proposed algorithm can be considered fully automatic since various methods have already been proposed for automatic detection and labeling of the center of the vertebral body [2]. While other methods depend on the sophisticated shape models, whose generation usually requires a great amount of time consuming manual interaction, in our proposed method we used coarse mean shape models built separately for all upper thoracic, lower thoracic and lumbar vertebrae. The vertebral mean shape model is registered to the thresholded bone prior map, which may lead to wrong alignment, if ribs are present or if trabecular bone intensities are low, i.e., close to soft tissue values. The overall result of 0.93 ± 0.04 in terms of the DSC is similar to other published methods. The result in lumbar region (0.96 ± 0.02) is better than 0.95 presented by Kadoury et al. [4] and the result of Ibragimov et al. [3] evaluated only on the lumbar vertebrae (0.93 ± 0.02). In the lower thoracic part of the spine, our result of 0.95 ± 0.02 exceeds the overall result for thoracic vertebrae (i.e. 0.93) presented by Kadoury et al. [4]. However, these methods were evaluated on different data sets, so results are not fully comparable. The lower DSC values for thoracic vertebrae T01–T06 can be explained by the influence of ribs and small intervertebral discs that mislead the segmentation.

6 Conclusion

In this work, a fully automatic system for vertebrae segmentation from CT was shown. It builds upon a TV based convex active contour segmentation that incorporates shape and intensity priors learned from training data and combines this prior information with image edges to achieve a minimal surface segmentation. Our results on the data of the MICCAI 2014 CSI challenge are promising and comparable to state of the art methods.

References

1. Chambolle, A., Pock, T.: A first-order Primal-Dual algorithm for convex problems with applications to imaging. J. Math. Imaging Vis. **40**(1), 120–145 (2011)
2. Glocker, B., Zikic, D., Konukoglu, E., Haynor, D.R., Criminisi, A.: Vertebrae localization in pathological spine CT via dense classification from sparse annotations. Medical Image Computing and Computer-Assisted Intervention—MICCAI 2013. Lecture Notes in Computer Science, vol. 8150, pp. 262–270. Springer, Berlin (2013)
3. Ibragimov, B., Likar, B., Pernus, F., Vrtovec, T.: Shape representation for efficient landmark-based segmentation in 3D. IEEE Trans. Med. Imaging **33**(4), 861–874 (2014)
4. Kadoury, S., Labelle, H., Paragios, N.: Automatic inference of articulated spine models in CT images using high-order Markov Random Fields. Med. Image Anal. **15**(4), 426–437 (2011)
5. Klinder, T., Ostermann, J., Ehm, M., Franz, A., Kneser, R., Lorenz, C.: Automated model-based vertebra detection, identification, and segmentation in CT images. Med. Image Anal. **13**(3), 471–482 (2009)
6. Ma, J., Lu, L., Zhan, Y., Zhou, X., Salganicoff, M., Krishnan, A.: Hierarchical segmentation and identification of thoracic vertebra using learning-based edge detection and coarse-to-fine deformable model. Medical Image Computing and Computer-Assisted Intervention—MICCAI 2010. Lecture Notes in Computer Science, vol. 6361, pp. 19–27. Springer, Berlin (2010)
7. Reinbacher, C., Pock, T., Bauer, C., Bischof, H.: Variational Segmentation of Elongated Volumetric Structures. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3177–3184 (2010)
8. Unger, M., Pock, T., Trobin, W., Cremers, D., Bischof, H.: TVSeg—interactive total variation based image segmentation. In: Proceedings of the British Machine Vision Conference (2008)
9. Weickert, J.: Anisotropic Diffusion in Image Processing. Teubner, Stuttgart (1998)
10. Werlberger, M., Pock, T., Unger, M., Bischof, H.: A variational model for interactive shape prior segmentation and real-time tracking. Scale Space and Variational Methods in Computer Vision. Lecture Notes in Computer Science, vol. 5567, pp. 200–211. Springer, Berlin (2009)
11. Yao, J., Burns, J.E., Munoz, H., Summers, R.M.: Detection of vertebral body fractures based on cortical shell unwrapping. Medical Image Computing and Computer-Assisted Intervention—MICCAI 2012. Lecture Notes in Computer Science, vol. 7512, pp. 509–516. Springer, Berlin (2012)