Spine Segmentation Using Articulated Shape Models

Tobias Klinder[1,2], Robin Wolz[2], Cristian Lorenz[2], Astrid Franz[2], and
Jörn Ostermann[1]

[1] Institut für Informationsverarbeitung, Leibniz University of Hannover, Germany,
`klinder@tnt.uni-hannover.de`,
[2] Philips Research Europe - Hamburg, Sector Medical Imaging Systems, Germany [*]

Abstract. Including prior shape in the form of anatomical models is
a well-known approach for improving segmentation results in medical
images. Currently, most approaches are focused on the modeling and
segmentation of individual objects. In case of object constellations, a si-
multaneous segmentation of the ensemble that uses not only prior knowl-
edge of individual shapes but also additional information about spatial
relations between the objects is often beneficial. In this paper, we present
a two-scale framework for the modeling and segmentation of the spine
as an example for object constellations. The global spine shape is ex-
pressed as a consecution of local vertebra coordinate systems while indi-
vidual vertebrae are modeled as triangulated surface meshes. Adaptation
is performed by attracting the model to image features but restricting the
attraction to a former learned shape. With the developed approach, we
obtained a segmentation accuracy of 1.0 mm in average for ten thoracic
CT images improving former results.

1 Introduction

Segmentation is still one of the main challenging problems in medical image
analysis. In order to improve the segmentation results, prior knowledge is usually
included in the form of anatomical models. For that purpose, complex geometri-
cal models of various individual anatomical objects have been built supporting
several medical applications, e.g., [1, 2].

In case of object constellations, simultaneous multi-object segmentation is
often beneficial compared to the separate segmentation of individual objects.
Model-based segmentation of single objects typically leads to misadaptations in
cases of no clear object boundary, similar structures in close vicinity, or patholo-
gies. The result is thereby sensitive to model initialization. Moreover, the capture
range of the segmentation is often very limited. However, by modeling not only
the shape of individual objects but also the spatial relations to other objects,
significant improvements can be achieved.

So far, little work focused on multi-object modeling and segmentation. Recently, de Bruijne et al. [3] presented a method for modeling relations between shape constellations using conditional probabilities between 2D-contours. A different way for multi-object modeling has been proposed by Boisvert et al. [4] where object relations are modeled as statistics on rigid transformations between the objects. However, both approaches have not been included to a model adaptation framework. One common way to account for multiple objects throughout the segmentation is to introduce a coupling term in the respective formulations as proposed, e.g., for coupled active contours [5] or level sets [6]. While overlapping between neighboring objects is prevented, spatial relations are not modeled.

In this paper, we present a two-scale framework for modeling object constellations with the individual parts interacting on each other throughout the entire adaptation. Although, we work exemplarily on spine segmentation in CT images, the method can be adapted to other object constellations due to its general formulation. Despite showing high contrast in CT, a separate segmentation of individual vertebrae often leads to unsatisfying results in images with low resolution or in case of pathologies. Without including knowledge about the constellation, vertebra shape models have to be positioned very close to their corresponding image structures to prevent adaptation to neighboring vertebrae having similar shape and intensity.

In Sect. 2, the principal idea of our approach is introduced. Basically, our modeling scheme is divided into two parts. A global spine model captures the constellations of objects while local vertebra models provide shape information of individual vertebrae. Patient individualization is achieved by subsequently adapting both parts with detailed formulations given in Sect. 2.1 and 2.2, respectively. Finally, spine segmentation is performed for thoracic CT images using the developed approach with results presented in Sect. 3.

2 Methods

In order to obtain a fast and robust framework, we follow a two-scale modeling scheme. On the one hand, a global model captures the object constellation by expressing the spine shape as a consecution of local vertebra coordinate systems (VCS) which have been earlier defined in [7]. On the other hand, local vertebra models provide shape information in the form of triangulated surface meshes of each individual object. For model adaptation, both parts are to be applied subsequently. By adapting the global model to the image, corresponding positions of individual objects are roughly to be found. Afterwards, local vertebra models are adapted providing the exact vertebra contour.

Adaptation for global and local model is based on the same idea. By using a physical metaphor an external energy (E_{ext}) drives the model towards image features while an internal energy (E_{int}) restricts attraction to a learned shape

$$E = E_{\text{ext}} + \alpha E_{\text{int}} \,. \tag{1}$$

The parameter α controls the trade-off between both energy terms. The final position is in each case found by applying Eq. 1 iteratively.

2.1 Global Model Adaptation

Model-based adaptation is known to be dependent on careful initialization. Once adaptation is misled, it can hardly recover. Thus, a global adaptation framework moves the individual objects closer to the respective image structure. By expressing the spine as a consecution of K rigid transformations between $K+1$ VCSs, a flexible representation of the object constellation is found that provides a large to small scale approach. Note that the representation as a consecution of rigid transformations has to be expressed relative to a reference VCS.

For the definition of the respective energies, an appropriate distance measure between two rigid transformations has to be found. For that purpose, the representation of the rotation as a rotation vector is introduced. The rotation vector is defined as the product of the axis of rotation expressed as a unit vector $\mathbf{n}$ and an angle of rotation θ. With the rotation vector representation, we use the left-invariant distance definition between two rigid transformations in [4]:

$$d(T_1, T_2) = N_\omega(T_2^{-1} \circ T_1) \quad \text{with} \quad N_\omega(T)^2 = N_\omega(\{\mathbf{r}, \mathbf{t}\})^2 = \|\mathbf{r}\|^2 + \|\omega\mathbf{t}\|^2 \quad (2)$$

where ω is used to weight the relative effect of rotation and translation, $\mathbf{r}$ is the rotation vector, and $\mathbf{t}$ the translation vector. As proposed in [4], the weight factor ω is set to 0.05. With this distance function, the respective energies as explained in detail below are defined as quadratic differences between a current and an 'optimal' position.

External Energy The key idea of the global external energy is to drive each object towards its corresponding image structure. Thus, we define the feature function

$$F_i(\mathbf{x}_i) = -\mathbf{n}_i{}^T \nabla I(\mathbf{x}_i) \frac{g_{\max}(g_{\max} + \|\nabla I(\mathbf{x}_i)\|)}{g_{\max}^2 + \|\nabla I(\mathbf{x}_i)\|^2} \quad (3)$$

that is carried out for each mesh surface independently judging the position in the image. The feature function is indeed evaluated for each triangle of the surface mesh at the position of the triangle barycentre $\mathbf{x}_i$. The image gradient $\nabla I(\mathbf{x}_i)$ is projected onto the face normal $\mathbf{n}_i$ of each triangle and damped by $g_{\max}$. The feature values of all triangles are summed up providing one value per vertebra for the current position.

The search for new positions is performed by testing for each object various discrete locations inside a local neighborhood around a given position. For that purpose, a cartesian grid inside a bounding box around the given position is defined. In order to not only cope with translations but also rotations, the original object is rotated in discrete steps around all axes obtaining L rotation matrices $\mathbf{R}_l$. At each of the N grid positions, the feature function is evaluated for the corresponding surface mesh translated by $\mathbf{t}_n$ as well as its L rotated versions. Due to the frequent presence of local minima, this exhaustive search is preferred to other optimization strategies but will be replaced in future by stochastic optimization methods.

The transformation resulting in the highest feature strength determines the new position of the object:

$$\text{argmax}_{\mathbf{R}_l, \mathbf{t}_n} \sum_{i \in M} F_i(\mathbf{R}_l \mathbf{x}_i + \mathbf{t}_n) \,, \tag{4}$$

where M is the number of triangles per mesh.

After finding the most promising positions for a pair of neighboring objects m and n, we can define the global external energy for the corresponding k-th transformation:

$$E_{\text{ext}_k} = d(\tilde{T}_{\text{VCS}_m}^{-1} \cdot \tilde{T}_{\text{VCS}_n}, T_k)^2 \tag{5}$$

where $\tilde{T}_{VCS_m}$ and $\tilde{T}_{VCS_n}$ are the transformations of the corresponding VCSs at the new positions to the world coordinate system. Thus, $\tilde{T}_{\text{VCS}_m}^{-1} \cdot \tilde{T}_{\text{VCS}_n}$ gives the transformation between the corresponding VCSs. The final transformation T_k between the neighboring objects will be determined by minimizing Eq. 1.

Internal Energy Driving the model towards high image features as performed by the external energy is restricted by the internal energy to prevent false attraction. The internal energy preserves similarity of the ensemble towards an earlier learned constellation model. In this case, our spine model covers the mean spine shape as well as its variability. Since the spine is expressed as a consecution of rigid transformations that do not belong to a vector space but to a Lie Group, conventional statistics can not be applied. Instead, statistical methods applied to Riemannian manifolds are used. The following formulation is based on [4] that recently presented a statistical model of the spine expressed as a consecution of rigid transformations. However, an inclusion to an adaptation framework has so far not been presented.

Following the statistics for Riemannian manifolds, the generalization of the usual mean called the Fréchet mean is defined for a given distance as the element μ that minimizes the sum of the distances with a set of elements $x_0, \ldots, x_N$ of the same manifold $\mathcal{M}$:

$$\mu = \arg\min_{x \in \mathcal{M}} \sum_{i=0}^{N} d(x, x_i)^2 \,. \tag{6}$$

Since the mean is given as a minimization, we use the gradient descent method performed on the summation obtaining

$$\mu_{n+1} = \text{Exp}_{\mu_n} \left(\frac{1}{N} \sum_{i=0}^{N} \text{Log}_{\mu_n}(x_i) \right) \,. \tag{7}$$

The functions Exp and Log are respectively the exponential map and the log map associated with the distance $d(x, y)$. The exponential map projects an element of the tangent plane on the manifold $\mathcal{M}$ and the log map is the inverse function. For the calculation of exponential and log map associated with the defined distance of Eq. 2, we refer to [4].

Finally, the generalized cross covariance Σ_{xy} is the expectation in the tangent plane of the mean using the log map

$$\Sigma_{xy} = \frac{1}{N} \sum_{i=0}^{N} \mathrm{Log}_{\mu_x}(x_i) \mathrm{Log}_{\mu_y}(y_i)^T \ . \tag{8}$$

In our case, the constellation of the spine is expressed as a multivariate vector $\mathbf{C}$ of K individual transformations $\mathbf{C} = [T_1, T_2, T_3, \dots, T_K]^T$ obtaining for the mean and the covariance

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_K \end{pmatrix} \quad \text{and} \quad \Sigma = \begin{pmatrix} \Sigma_{T_1 T_1} & \Sigma_{T_1 T_2} & \dots & \Sigma_{T_1 T_K} \\ \Sigma_{T_2 T_1} & \Sigma_{T_2 T_2} & \dots & \Sigma_{T_2 T_K} \\ \vdots & \vdots & & \vdots \\ \Sigma_{T_K T_1} & \Sigma_{T_K T_2} & \dots & \Sigma_{T_K T_K} \end{pmatrix} \ . \tag{9}$$

With Eq. 9, the formulation for the statistical spine model is given. In order to reduce the dimensionality of our model, principal component analysis is performed on the covariance matrix.

Now, we define the global internal energy for each transformation T_k as

$$E_{\mathrm{int}_k} = d\left(\mathrm{Exp}_{\mu_k} \left(\sum_{i=1}^{t} b_i \mathbf{a}_{i,k} \right), T_k \right)^2 \quad \text{with} \quad \mathbf{b} = \mathbf{A}^T (\mathrm{Log}_{\mu} \mathbf{C}) \tag{10}$$

where the matrix of the individual eigenvectors $\mathbf{a}_i$ is denoted as $\mathbf{A}$ and b_i is the coordinate of the weight vector $\mathbf{b}$ associated with the i-th principal component. The internal energy penalizes differences between the model and the current constellation. The closest constellation to the model is determined by projecting the given constellation $\mathbf{C}$ in the sub-space defined by the principal components.

Optimization After calculating the respective energy terms, the final transformations T_k between the individual objects are determined by minimizing

$$E(T_k) = d(T_{\mathrm{ext}_k}, T_k)^2 + \alpha \cdot d(T_{\mathrm{int}_k}, T_k)^2 \tag{11}$$

for each object separately. The transformations T_{ext_k} and T_{int_k} are obtained from the external and internal energy, respectively. The final transformation $T_{k_{\mathrm{opt}}}$ is found using a Downhill-Simplex optimizer. As the representation of the ensemble as a consecution of rigid transformations requires a reference VCS, the vertebra with the highest feature strength is taken as the reference in each iteration.

2.2 Local Model Adaptation

After positioning the vertebra models using the global model adaptation, a local non-rigid free-form deformation similar to [8] of the individual surface meshes is carried out. Segmentation of all vertebrae is performed simultaneously with the individual shapes interacting on each other to prevent misadaptations.

External Energy The local external energy

$$E_{\text{ext}} = \sum_{i \in T} w_i (\mathbf{e}_{\nabla \mathbf{I}}(\mathbf{x}_i + \tilde{\mathbf{c}}_i))^2 \quad \text{with } w_i = \max\{0, F_i^*(\mathbf{x}_i + \tilde{\mathbf{c}}_i) - \delta \|\tilde{\mathbf{c}}_i\|^2\} \quad (12)$$

drives each triangle barycentre $\mathbf{x}_i$ towards a detected potential anatomical surface point $\mathbf{x}_i + \tilde{\mathbf{c}}_i$. The unit vector $\mathbf{e}_{\nabla \mathbf{I}}$ points in direction of the image gradient at the surface point $\mathbf{x}_i + \tilde{\mathbf{c}}_i$.

For each triangle barycentre $\mathbf{x}_i$, surface detection is carried out within a locally defined sampling grid along the triangle surface normal. At each candidate position $\mathbf{c}_k$, the feature function is evaluated and finally the point $\tilde{\mathbf{c}}_i$ is chosen that maximizes the objective function

$$\tilde{\mathbf{c}}_i = \text{argmax}_{\mathbf{c}_k} \left\{ F_i^*(\mathbf{x}_i + \mathbf{c}_k) - \delta \|\mathbf{c}_k\|^2 \right\} . \quad (13)$$

The parameter δ controls the trade-off between feature strength and distance.

As a feature function, we apply the one already defined in Eq. 3 but with an additional factor c penalizing overlapping resulting in $F_i^* = c \cdot F_i$.

If no collision with a neighbored mesh is present, the factor c will be equal to one. In case of a collision, c becomes the smaller the deeper the detected point is inside the neighbored mesh. Note that this formulation does not prohibit collision but makes points that lie inside other meshes less attractive.

An effective implementation of the collision detection is achieved by labeling all meshes and assigning adjacent surface patches of neighboring vertebrae. Since collision detection has to be only carried out for corresponding adjacent surfaces that potentially overlap, the computational complexity is significantly reduced.

Internal Energy As a regularization, the local internal energy

$$E_{\text{int}} = \sum_{j \in V} \sum_{k \in N(j)} ((\hat{\mathbf{v}}_j - \hat{\mathbf{v}}_k) - s\mathbf{R}(\mathbf{v}_j - \mathbf{v}_k))^2 \quad (14)$$

preserves shape similarity of all adapted vertices $\mathbf{v}_j$ to the model vertices $\hat{\mathbf{v}}_j$ with $N(j)$ being the set of neighbors of vertex j. neighboring vertices are those connected by a single triangle edge. The scaling factor s and the rotational matrix $\mathbf{R}$ are determined by a closed-form point-based registration method based on a singular value decomposition prior to calculation of Eq. 14.

Optimization In the final optimization scheme for the local model adaptation, the vertex positions of the triangular surface mesh are indeed the parameters to be varied. As only interdependencies between neighboring vertices exist and the energy terms are of a quadratic form, the conjugate gradient method is used for minimization of the final equation system with a sparsely filled matrix.

3 Results

For our model building, vertebra models were adapted to 18 thoracic CT data sets using [7]. In case of misadaptations, manual corrections were performed.

From the adapted surface meshes, mean shape models of all vertebrae as well as a statistal model of rigid transformations between the VCSs was created.

The new framework was applied to ten test data sets including pathologies like strong calcifications between vertebrae or scoliosis. In each case, model initialization was given by [7]. Reference segmentation was provided by manually positioning and automatically adapating the individual surface models. Misadaptations were again manually corrected. As the main contributions of our framework are the global adaptation and the collision detection during local adaptation, we compared our results to the adaptation using the shape-constrained deformable model method (SCDM) [8]. Automatic segmentation was compared to the reference by calculating for each vertex of the adapted mesh the closest point on the reference surface. One problem that could be circumvented by our framework is the misadaptation of several models to wrong neighboring vertebrae as examplarily shown in Fig. 1. In the ten test cases, this secenario occured for four patients. In the remaining six data sets, SCDM already adapted to the correct image structure. To investigate the effect of collision detection, we compared the segmentation result for these six data sets using SCDM and SCDM with collision detection but no global model. As shown in Tab. 1, the mean error could be slightly reduced from 1.1 to 0.9 mm in average. Comparing the results for all ten cases using SCDM with our two-scale framework, the segmentation results were significanty improved from 2.7 to 1.0 mm in average. However, it has to be noted that the results are slightly biased due to the fact that the creation of the reference segmentation was based on the same algorithm with additional user interaction.
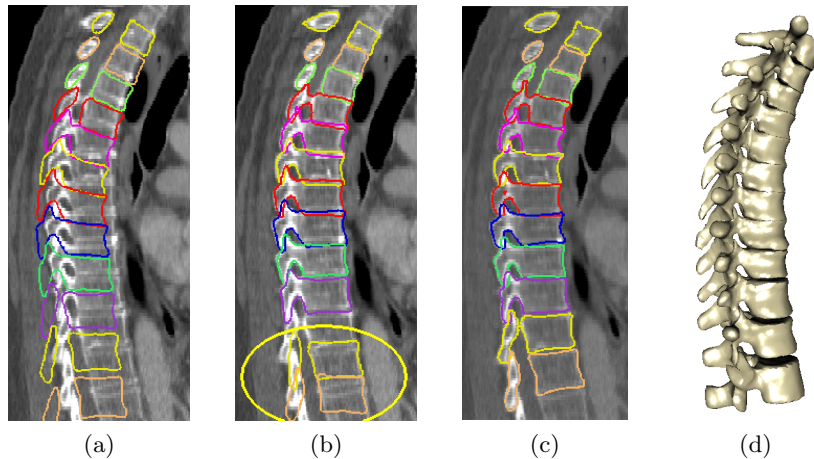

(a)　　　　　　(b)　　　　　　(c)　　　　　　(d)

Fig. 1: Automatically positioned models (a) are adapted using SCDM (b) and two-scale method (c). Due to the consideration of overall constellation, misadaptation of T11 and T12 could be overcome and overlapping between neigbored models could be decreased. Surface rendering of adapted meshes (d).

Table 1: Adaptation result for ten test data sets showing average ± standard deviation (maximum) of segmentation error (values in mm). For each algorithm, the result for one chosen test case is examplarily given in detail.

	Collision Detection		Two-Scale	
	Patient 6	6 Data Sets	Patient 1	10 Data Sets
SCDM	1.3±0.3 (6.9)	1.1±0.2 (5.5)	2.8±1.7 (8.9)	2.7±2.8 (9.0)
new	0.9±0.3 (6.0)	0.9±0.2 (5.4)	0.8±0.2 (5.8)	1.0±0.3 (6.5)

4   Conclusion

In this paper, we presented a two-scale framework for the automatic segmentation of object constellations. The constellation is modeled as a consecution of local coordinate systems while the individual objects' shape is presented as a triangulated surface model. In contrast to former approaches performing only non-rigid deformation for each object separately, we extended the adaptation allowing a simultaneous segmentation of the constellation with individual objects interacting on each other. Compared to former approaches, a significant improvement of the segmentation result has been achieved while at the same time the capture range for the adaption has been increased. Furthermore, we are more robust against pathologies in spine shape with the developed approach due to the inclusion of a statistical spine shape model.

References

1. Lorenz, C., von Berg, J.: A comprehensive shape model of the heart. Medical Image Analysis **10** (2006) 657–670
2. Blaffert, T., Barschdorf, H., von Berg, J., Dries, S., Franz, A., Klinder, T., Lorenz, C., Renisch, S., Wiemker, R.: Lung lobe modeling and segmentation with individualized surface meshes. In: Proc. SPIE Medical Imaging 2008 (to appear). (2008)
3. de Bruijne, M., Lund, M., Tanko, L., Pettersen, P., Nielsen, M.: Quantitative vertebral morphometry using neighbor-conditional shape models. In: Proc. MICCAI. Volume 4190., Springer (2006) 1–8
4. Boisvert, J., Pennec, X., Labelle, H., Cheriet, F., Ayache, N.: Principal spine shape deformation modes using Riemannian geometry and articulated modes. In: Proc. AMDO. (2006) 346–355
5. Zimmer, C., Olivo-Marin, J.: Coupled parametric active contours. In: IEEE Trans. Pattern Anal. Mach. Intell. Volume 27(11). (2005) 1838–1842
6. Rousson, M., Khamene, A., Diallo, M., Celi, J.C., Sauer, F.: Constrained surface evolutions for prostate and bladder segmentation in ct images. In: Proc. CVBIA. Volume 3765 of LNCS., Springer (2005) 251–260
7. Klinder, T., Lorenz, C., von Berg, J., Dries, S., Bülow, T., Ostermann, J.: Automated model-based rib cage segmentation and labeling in CT images. In: Proc. MICCAI. Volume 4792., Springer (2007) pp. 195–203
8. Weese, J., Kaus, M., Lorenz, C., Lobregt, S., Truyen, R., Pekar, V.: Shape constrained deformable models for 3D medical image segmentation. In: Proc. IPMI, Springer (2001) 380–387