

# Simultaneous Model-based Segmentation of Multiple Objects

Astrid Franz<sup>1</sup>, Robin Wolz<sup>1</sup>, Tobias Klinder<sup>1,2</sup>, Cristian Lorenz<sup>1</sup>, Hans Barschdorf<sup>1</sup>, Thomas Blaffert<sup>1</sup>, Sebastian P. M. Dries<sup>1</sup>, Steffen Renisch<sup>1</sup>

<sup>1</sup>Philips Research Europe - Hamburg, Sector Medical Imaging Systems

<sup>2</sup>Leibniz University Hannover, Institute of Information Processing

astrid.franz@philips.com

**Abstract.** Deformable models are used for the segmentation of objects in 3D images by adapting flexible meshes to image structures. The simultaneous segmentation of multiple objects often causes problems like violation of spatial relationships. Here we present two methods for including prior knowledge into the segmentation process: The first deals with objects sliding with respect to each other, and the second considers a pre-defined minimal distance between neighboring objects. Using this prior knowledge improved segmentation results can be reached.

## 1 Introduction

Segmentation is still one of the main challenging problems in medical image analysis. In order to improve the segmentation, prior knowledge in form of anatomical models is frequently used. For this reason, deformable models of several organs have been earlier developed and used for segmenting the corresponding objects in medical images [1, 2, 3]. In model-based segmentation, flexible meshes, normally triangle meshes, are adapted to the corresponding image structures. Usually this adaptation is carried out for every object separately by optimizing a weighted sum of two competing energies: an external energy  $E_{\text{ext}}$  driving the mesh triangles towards image features and an internal energy  $E_{\text{int}}$  preserving the form of the model [4].

The segmentation is often complicated by various factors, like, e.g., no clear boundaries of the object to be segmented, pathologies or intensity similarity of adjacent structures, thus leading to wrong adaptation results. In most cases the segmentation can be improved when not only using prior knowledge of the object's shape but also information about spatial relations to neighboring structures. First attempts to include object relations are described in [5, 6, 7] and references therein. For instance, additional edges connecting neighboring meshes are introduced, or non-overlapping constraints are used.

Here we propose a method for simultaneous segmentation of objects close to each other by introducing additional edges similar to [5]. For these edges, a connection energy is defined in different ways incorporating prior knowledge of the corresponding anatomical region. Hence sliding organs or a required minimal distance of two organs can be represented.

## 2 Materials and Methods

Let  $M_i$  ( $i = 1, \dots, n$ ) denote the pre-positioned meshes for  $n$  objects which are to be segmented. Then for any two objects being near to each other, a connecting object  $C_i$  is defined by a set of additional edges connecting points of the two corresponding meshes  $M_{a(i)}$  and  $M_{b(i)}$ . For each connecting object, a transformation class  $T_i$  is chosen which describes the allowed deformations, or a required minimal distance  $d_i$  is chosen. Then iteratively the following steps are performed:

1. For each mesh  $M_i$ , a feature search is carried out, and the corresponding external energy  $E_{\text{ext},i}$  is computed as in [4].
2. For each mesh  $M_i$ , a similarity transform, i.e. a rotation matrix  $r_i$  and a scaling factor  $s_i$  is computed, which best aligns the model mesh with the actual mesh configuration. The remaining deformation results in an internal energy  $E_{\text{int},i} = \sum_{(j,k) \in \text{edges}(M_i)} (\hat{x}_j - \hat{x}_k - s_i r_i (x_j - x_k))^2$  [4], where the  $\hat{x}_j$  denote the coordinate vectors of the corresponding mesh vertices in the pre-positioned model and the  $x_j$  denote the coordinate vectors of the actual mesh configuration.
3. For each connecting object  $C_i$  (connecting the meshes  $M_{a(i)}$  and  $M_{b(i)}$ )
  - (a) a transformation  $t_i \in T_i$  is determined best aligning the pre-positioned model connection object with the actual connection object, and the connection energy  $E_{\text{con(a)},i} = \sum_{(j,k) \in \text{edges}(C_i)} (\hat{x}_j - \hat{x}_k - t_i(x_j - x_k))^2$  is computed, or
  - (b) two parallel planes are fitted through the vertices, one plane through the vertices belonging to  $M_{a(i)}$  and the other plane through the vertices belonging to  $M_{b(i)}$ . Let  $\nu_i$  be the normal vector to these planes, going from  $M_{a(i)}$  to  $M_{b(i)}$ . For every edge  $(j, k)$  in  $C_i$  ( $j \in M_{a(i)}$ ,  $k \in M_{b(i)}$ ), the scalar product  $\langle x_k - x_j, \nu_i \rangle$  is computed, which results in the connection energy  $E_{\text{con(b)},i} = \sum_{(j,k) \in \text{edges}(C_i)} (\langle x_k - x_j, \nu_i \rangle - \max(\langle \hat{x}_k - \hat{x}_j, \hat{\nu}_i \rangle, d_i))^2$ .
4. The overall energy  $E = \sum_{i=1}^n \alpha_i E_{\text{ext},i} + \sum_{i=1}^n \beta_i E_{\text{int},i} + \sum_{i=1}^m \gamma_i E_{\text{con},i}$ , with weighting factors  $\alpha_i$ ,  $\beta_i$  and  $\gamma_i$ , is minimized, resulting in new coordinate vectors for each mesh.

This iteration is repeated for a given number of steps, or it is stopped when the new vertex coordinates do not differ significantly from the ones before.

The transformation class  $T_i$  in step 3a can for instance be the class of affine transformations (without reflection). Then shearing is allowed which ensures that the connected meshes can slide with respect to each other. If the distance of the two meshes is to be preserved, the scaling factors of the affine transformation can be restricted. Hence the transformation class can be defined in order to match the requirements of the underlying application.

If two meshes are to have a given minimal distance but this distance can not be reached by pre-positioning, step 3b can be applied. In this step, the connecting edges are projected onto the normal vector of the parallel planes describing the adjacent mesh surfaces. The connection energy penalizes a length of these

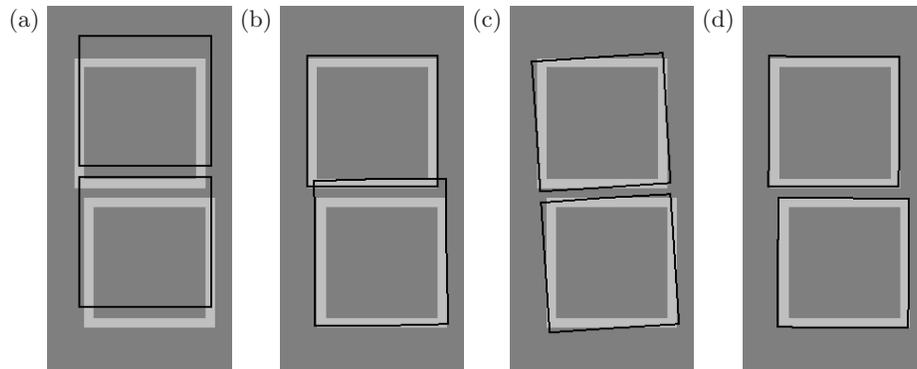
**Table 1.** Mean  $d_{\text{mean}}$ , maximum  $d_{\text{max}}$  and standard deviation  $\sigma$  of the distance of the adapted vertebral meshes to the surface of the ground truth meshes for three CT datasets (all numbers are given in mm)

connection energy	dataset 1			dataset 2			dataset 3		
	$d_{\text{mean}}$	$d_{\text{max}}$	$\sigma$	$d_{\text{mean}}$	$d_{\text{max}}$	$\sigma$	$d_{\text{mean}}$	$d_{\text{max}}$	$\sigma$
without	0.90	4.86	0.71	1.35	5.77	1.06	0.72	3.96	0.65
$E_{\text{con(a)}}$	0.83	4.11	0.62	1.24	5.31	0.93	0.57	3.23	0.46
$E_{\text{con(b)}}$	0.81	4.38	0.61	1.23	5.17	0.90	0.61	3.63	0.51

projections smaller than the given minimal distance. Hence an initial mesh overlap (indicated by a negative length of connecting edges) can be corrected during the mesh adaptation process. If, on the other hand, the minimal distance condition is fulfilled, then the projected length of the connecting edge is compared with the projected length in the model, and a deviation is penalized. Hence in this case the distance of the two connected mesh surfaces is to be preserved.

### 3 Results

A synthetic test example image (size  $100 \times 200 \times 100 \text{mm}^3$ ) contains two cubes (size  $70^3 \text{mm}^3$ ) with a light outer region like cortical bone and a dark inner region like cancellous bone (Fig. 1). The two cubes are staggered as for instance vertebrae in the spine may be. The two meshes are pre-positioned in such a way that the upper surface of the lower mesh is closer to the cortex-marrow-transition of the upper cube than to the surface of the lower cube, simulating a typical imperfect pre-positioning (Fig. 1a). Hence an adaptation without a connection energy

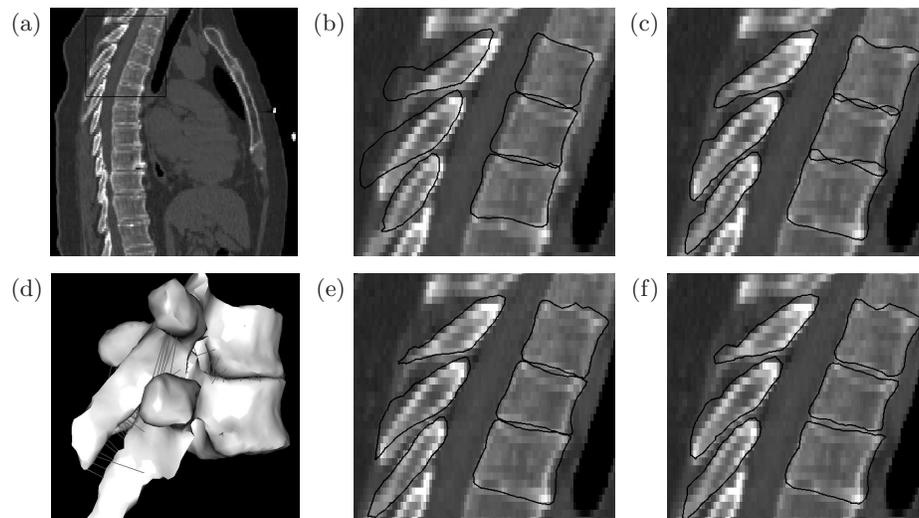


**Fig. 1.** Synthetic example image of two cubes with two pre-positioned meshes (a), adaptation results without connection energy (b) and with connection energies based on rigid transformation (c) and affine transformation (d). Result (d) is also achieved with the connection energy requiring a minimal distance of 7mm for both starting at configuration (a) or starting at configuration (b)

will fail (Fig. 1b). In contrast, a connection energy  $E_{\text{con(a)}}$  based on a rigid or affine transformation can ensure that the upper surface of the lower mesh finds its correct cube surface, since the lower mesh is forced downwards due to the adaptation of the upper mesh. A rigid transformation keeps the relative positioning of the adjacent mesh surfaces (Fig. 1c), whereas an affine transformation allows for shearing (Fig. 1d). The shearing result in Fig. 1d can also be obtained with the connection energy  $E_{\text{con(b)}}$  requiring a minimal distance of 7mm. This connection energy additionally has the potential of correcting initial mesh overlaps: When starting the adaptation process with the mesh configuration shown in Fig. 1b, the result shown in Fig. 1d is achieved as well.

Fig. 3 shows a medical example application. In a computed tomography (CT) image of the spine, the third, fourth and fifth thoracic vertebrae are to be segmented. As the synthetic example illustrated, the transition from cortical to cancellous bone is comparable to the bone boundary, hence a separate adaptation of each vertebra is likely to fail (Fig. 3c). In contrast, the connection energy  $E_{\text{con(a)}}$  based on an affine transformation preserves the spatial relationship of the vertebral meshes and hence results in a reasonable segmentation (Fig. 3e). The connection energy  $E_{\text{con(a)}}$  requiring a minimal distance of 2mm furthermore corrects for the slight initial mesh overlaps (Fig. 3b). It moves the meshes apart during the adaptation process (Fig. 3f).

We validated our methods for three CT datasets of the 12 thoracic vertebrae with a slice thickness between 2 and 3 mm containing no significant pathologies.



**Fig. 2.** CT image of the thorax and highlighted region containing the third, fourth and fifth thoracic vertebrae (a), the initial mesh positioning (b) and the adaptation results without connection energy (c) and with a connection energy based on affine transformation (e) and requiring a minimal distance of 2mm (f), based on connecting objects consisting of additional edges connecting two neighboring vertebral meshes (d)

A global manual positioning including rotation and translation of a mean model [8] was used for initialization of the segmentation routines. We compared the segmentation results with manually adapted meshes used as ground truth. The mean, maximum and standard deviation of the distance between the vertices of the adapted meshes to the surface of the ground truth meshes show a decrease of the segmentation error of about 10 % with the introduced methods (Tab. 1).

## 4 Discussion

As the examples in the previous section illustrate, the introduced connection energies yield an acceptable segmentation of multiple objects close to each other while an independent adaptation of the multiple meshes fails. The kind of connection energy to be applied depends on the requirements of the underlying application. A connection energy based on a rigid transformation keeps the spatial relationship of the pre-positioned meshes and hence avoids mesh overlaps if the mesh initialization is collision-free. A connection energy based on an affine transformation is more flexible allowing for a sliding of adjacent mesh surfaces. This property is important in anatomical regions where organ surfaces may slide with respect to each other. The second variant of a connection energy requiring a minimal distance between two meshes additionally can correct for initial mesh overlaps. Hence a slightly incorrect pre-positioning can be revised during the mesh adaptation process. This strategy absolves the user from a collision-free pre-positioning of the meshes or avoids a computationally expensive collision detection algorithm to be applied prior to mesh adaptation. However, further improvements may be achieved by learning the minimal distance from an ensemble of adapted models. Hence the connection energies bear the potential of correctly segmenting multiple objects with little required user interaction.

## References

1. McInerney T, Terzopoulos D. Deformable models in medical image analysis: A survey. *Med Image Anal.* 1996;1(2):91–108.
2. Kaus MR, McNutt T, Pekar V. Automated 3D and 4D organ delineation for radiation therapy planning in the pelvic area. *Procs SPIE.* 2004; p. 346–55.
3. von Berg J, Lorenz C. A statistical geometric model of the heart. *Procs BVM.* 2006; p. 116–20.
4. Weese J, Kaus M, Lorenz C, et al. Shape constrained deformable models for 3D medical image segmentation. *Procs IPMI.* 2001; p. 380–7.
5. Kaus MR, von Berg J, Niessen W, et al. Automated segmentation of the left ventricle in cardiac MRI. *Procs MICCAI.* 2003; p. 432–9.
6. Freedman D, Radke RJ, Zhang T, et al. Model-based multi-object segmentation via distribution matching. *Procs CVPR.* 2004; p. 11.
7. Costa MJ, Delingette H, Novellas S, et al. Automatic segmentation of bladder and prostate using coupled 3D deformable models. *Procs MICCAI.* 2007; p. 252–60.
8. Klinder T, Lorenz C, von Berg J, et al. Automated model-based rib cage segmentation and labeling in CT images. *Procs MICCAI.* 2007; p. 195–203.