

# Stable 3D Mass-Spring Models for the Segmentation of the Thyroid Cartilage

Jana Dornheim<sup>1</sup>, Lars Dornheim<sup>1</sup>, Bernhard Preim<sup>1</sup>, Klaus D. Tönnies<sup>1</sup>, Ilka Hertel<sup>2</sup>, and Gero Strauss<sup>2</sup>

<sup>1</sup> Otto-von-Guericke-Universität, Postfach 4120,  
D-39106 Magdeburg, Germany

{Jana.Dornheim, Lars.Dornheim, Preim, Toennies}@isg.cs.uni-magdeburg.de

<sup>2</sup> Hals-Nasen-Ohren-Universitätsklinik,  
Universitätsklinikum Leipzig,  
Liebigstr. 18a,

D-04103 Leipzig, Germany

{Ilka.Hertel, Gero.Strauss}@medizin.uni-leipzig.de

**Abstract.** The preoperative planning of primary tumor resections in the larynx region shall be supported by a 3D visualization of the patient-specific anatomy and pathological situation. This requires a segmentation of the larynx cartilage structures from CT datasets.

In our work, we apply and adapt stable 3D mass-spring models to create a 3D deformable shape model for the segmentation of the thyroid cartilage. A new concept for elastic initialization of the model is presented, allowing the deformable model to adapt specifically to patient-specific shape variations and pathological deformations.

## 1 Introduction

In the case of tumor affections in the larynx and lower hypopharynx, the patient's life expectancy and further life quality depend strongly on the required surgical treatment. Which parts of the larynx need to be resected, has a determining impact on the patient's ability to breathe, swallow and speak. For the decision on a surgical strategy, the extent of the tumor must be evaluated with respect to infiltration of the following structures:

- the vocal chords and muscles (usually judged by laryngoscopy),
- the glottis, subglottic and supraglottic space,
- and the larynx cartilages, in particular the epiglottis, thyroid cartilage, cricoid cartilage, and the two arytenoid cartilages (often judged by CT [1]).

For the assessment of the air and cartilage structures, a 3D visualization of the patient-specific anatomy and pathological situation is desirable to reduce uncertainties in the chosen surgical procedure. This requires a precise segmentation of the larynx and its substructures from CT datasets.

The inhomogenous nature of the cartilage makes its segmentation a challenging

task, for which neither simple edge-based techniques, such as LiveWire, nor gray-value-based segmentation techniques are appropriate. Since profound anatomical knowledge is needed to bridge areas of weak signal in the cartilage wall, we target at a 3D model-based segmentation of the thyroid cartilage.

## 2 State of the Art

For the segmentation of the thyroid cartilage, no specific approaches exist, to our knowledge. General methods for segmentation using 3D deformable models are known [2]: *3D Active Contours* or *Balloons* [4] incorporate rough shape knowledge by means of a viscosity condition. The use of an inflation force prevents them from shape collapse and drives them toward the desired signal. This global representation of shape does however not allow to model complex shape information. *Implicit 3D deformable models* [3] do not bear the problem of instability and need no inflation force. However, they are restricted to describing regular geometric shapes, that can be described by a simple equation. *ASMs* and *AAMs* [5] provide support for segmenting more complex shapes by means of a statistical analysis of training data. They require large amounts of training data and a very good correspondance of 3D points, which makes model creation laborious and segmentation results imprecise. For the segmentation of pathological shape variations, ASMs and AAMs are principally inappropriate. *Mass-spring models* (MSMs) have their main 3D application in the simulation of soft tissue [6], where even complex shape information is encoded in a 3D model. For 3D segmentation however, MSMs have long been neglected, due to their increased instability in 3D. In surgical simulation applications, these problems are solved by a high number of masses and dense cross-linking between them. This is reasonable for this application, since detailed modelling of the object shape is desired. For segmentation with 3D MSMs we consider different issues:

- The simulation speed is significantly slower. It depends strongly on the number of masses, since a sensor collecting and processing image information is usually attached to each mass.
- A detailed modelling of the object shape is not always desired. Instead, it is reasonable to integrate only the relevant shape features into the model.
- The forces acting at the MSM are much more uncontrollable than with a surgical simulation model. They may drag all masses into different and unforeseeable directions if the model is applied to new images.

These issues make model stability coupled with a small number of masses and weak cross-linking the main requirements for using 3D MSMs for segmentation. *Stable 3D MSMs* with additional *torsion forces* for stabilization, introduced by [7], offer a mechanism to fulfill these requirements.

### 3 Method

In our work, we create two stable 3D MSMs [7] of the thyroid cartilage semi-automatically from a manual segmentation. We thereto adapt and refine the model creation introduced in [8], which proposes a model topology consisting of two parts created independantly from each-other:

1. an (outer) surface submodel containing masses with *gradient sensors* and the contour facets, and
2. an (inner) volume submodel containing *intensity sensors*.

Both submodels are connected by 1:n spring connections. *Intensity sensors* operate directly on the input dataset and drag the attached masses toward neighbouring voxels of a certain gray value interval. *Gradient sensors* create a force dragging the attached mass toward neighbouring voxels with a high gradient magnitude. The gradient sensors work on a direction-filtered gradient image, ensuring that contour masses are only attracted by gradients of the same direction as the incident surface normal. This prevents the model from being distracted by neighbouring but irrelevant gradients.

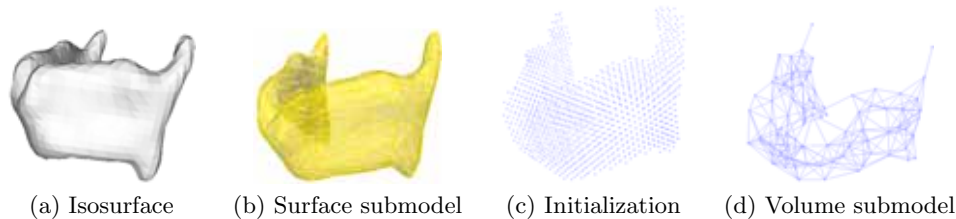
### 4 Model Generation for the Thyroid Cartilage

The semi-automatic model generation for the thyroid cartilage is based on a manual segmentation created from a dataset with a visually average-shaped larynx. The sample segmentation is available as a binary volume dataset. For our examination, two different models were created:

1. One *volumetric MSM*, according to the model creation of [8], consisting of a surface submodel and a volume submodel connected by 1:1 connections.
2. For comparison, a pure *surface MSM* is created, consisting of the surface submodel only. This model was employed to evaluate whether segmentation based on edge detection alone is more appropriate than using gray value information.

#### 4.1 Generation of the Surface MSM

The sample segmentation was resampled down to an isotropic voxel size for efficient model generation. From this resampled segmentation, an isosurface was generated using the marching cubes method (Fig. 1(a)), smoothed and simplified with Quadric Error Metrics [9]. The resulting number of triangles (50–200) provides an appropriate modelling of the cartilage shape while still allowing for realtime model simulation (Fig. 1(b)). The resulting surface was used to create the surface MSM: For each vertex of the mesh, a mass point was created. For each edge in the mesh, a spring connecting the incident vertices, resp. mass points, was created. All masses and spring constants throughout the model were set to 1.0. A gradient sensor was attached to all mass points of the surface MSM.



**Fig. 1.** Stages of model creation for the thyroid cartilage.

## 4.2 Generation of the Volumetric MSM

For the generation of the volumetric model, inner masses with intensity sensors have to be created and combined with the surface submodel. For this, an initial set of mass points is created by placing one mass point at each voxel of the resampled manual segmentation (Fig. 1(c)). Then, the initial point set is reduced iteratively.

**Reduction of Inner Masses.** For each mass point, all mass points inside a neighbourhood of radius  $r$  are moved to their common center of mass and merged. By iterating this relaxation, the initial point set is reduced considerably, and fills the manual segmentation evenly.

The convergence of the relaxation toward a reasonable point set depends on the choice of  $r$ . According to our experiments, a radius of half the desired minimum distance of two mass points leads to a convergence representing the original shape well. A dense placement of the volume masses has the advantage that the inner properties of the segmentation target structures are measured at more positions, which is equivalent to a higher sampling rate. We therefore always chose  $r$  to be within  $[\text{voxelsize}; \sqrt{2} \cdot \text{voxelsize}]$ .

**Cross-Linking of the Inner Masses.** Each of the resulting volume mass points is linked with each neighbouring mass point within a user-defined radius  $p$ . In our model of the thyroid cartilage, a radius of  $p = 10\text{mm}$  (for a voxel size of 2.148 mm) led to good results in all cases (Fig. 1(d)).

**Connecting inner and outer submodel.** Both the volume and the surface submodel are interconnected by springs and merged to one volumetric MSM. A 1:1 interconnection (each point of the surface submodel is connected to the closest mass point of the volume submodel) has proved to be appropriate for good model stability.

## 5 Model Placement and Adaptation

### 5.1 Initial Model Placement

A good initial adjustment of the deformable model is needed, so that the adaptation of the model to the dataset will not be distracted by ‘wrong’ gray value information of adjacent structures.

For the initialization, the model’s position, rotation and scaling have to be adjusted. Thereto, we derived a simple initialization method from the model specification already available: For the deformable model, a mass point at each of the most prominent landmarks is marked as a *key mass* at the end of the model creation process. The user can specify the positions of these key masses by clicking into the dataset, where they are fixed. Then, the model simulation is started with only the spring and torsion forces active, but all sensor input ignored. The internal forces adapt the model’s shape to the key positions (Fig. 2(a),(b)).

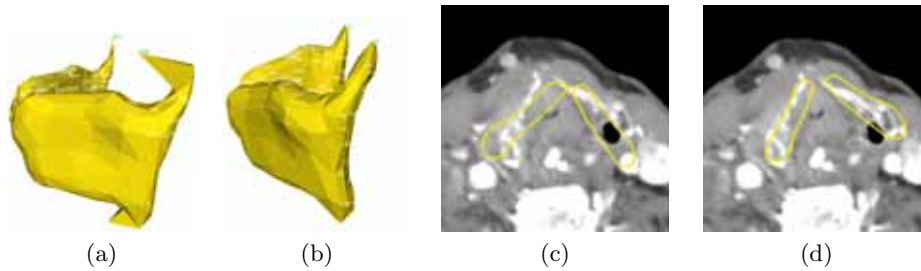
The 6 key masses for the model of the thyroid cartilage are positioned at the *cornu superius left* and *right*, *cornu inferius left* and *right*, as well as the *upper (Adam’s apple)* and *lower* end of the *larynx front side* (Fig. 2(a)). With this method the model adjusts itself to the specified key positions easily (spring force weighting  $w_f = 5.0$ , torsion force weighting  $w_t = 10.0$ , damping factor  $d = 0.001$ , simulation time step  $\Delta t = 0.05$ ).

After complete adaptation, the current shape deformation is *learned* by the model. That means the rest lengths and rest directions of all springs are set to their current (deformed) length and direction values. In comparison to an affine transformation matrix used for model initialization, our method of initial placement has the following advantages:

- The user can select the number of key positions. Even the choice which landmarks to use as key positions could be left to the user.
- Elastic deformations of the model are possible for the model initialization, allowing a more detailed initial placement.
- Even in cases with a displaced or deformed larynx due to a tumor affection, the model could be initialized very well (Fig. 2(c),(d)).
- Learning the deformed model shape after the initialization phase allows to use more specific shape knowledge for each patient, instead of using one average model for all patients. This is particularly beneficial in cases with a pathological larynx shape.

### 5.2 Model Adaptation to the Thyroid Cartilage

After initialization, the precise model adaptation is started (sensor force weighting  $w_s = 0.05$  for gradient sensors,  $w_s = 0.001$  for intensity sensors, spring force weighting  $w_f = 1.0$ , torsion force weighting  $w_t = 2.0$ , damping factor  $d = 0.001$ , simulation time step  $\Delta t = 0.05$ ). All key masses are left fixed, so that the model adaptation occurs within their frame of reference. This way, the model is kept at the correct position. Besides, the lengths of the cornu superius and inferius vary widely among different patients. By keeping the top of the cornu fixed, we can ensure that the whole cornu is found.



**Fig. 2.** Elastic model initialization: (a) Key masses placed, (b) MSM after adaptation to key positions, (c) Affine Initialization, (d) After elastic initialization

## 6 Results and Evaluation

### 6.1 Data Material and Gold Standard

12 CT datasets of the neck were acquired for preoperative planning, containing the larynx. The slice thickness of the datasets ranged from 1.5 mm to 6.0 mm. The datasets varied significantly w.r.t. signal-to-noise ratio, contrast and motion artifacts. In 3 datasets, the larynx was displaced or partially destroyed due to tumor affection. On all 12 datasets, a manual segmentation of the thyroid cartilage was created by an experienced user and controlled by a radiologist. These verified manual segmentations were used as a gold standard for the evaluation. For comparison, a manual segmentation of all datasets was created by an experienced user.

### 6.2 Experiments

From one dataset with an average-shaped larynx, a volumetric MSM and a surface MSM were generated as described above. The key masses were marked manually and used throughout all experiments. The two models were then applied to the remaining 11 datasets (leave-one-out-test) in the following manner:

1. The user marks the key positions in the dataset (6 markers).
2. The model is automatically positioned and scaled according to the bounding box of the key positions.
3. The model is adapted to the key positions (1 click for stopping this phase).
4. The newly adapted shape is automatically learned by the model.
5. The model adaptation to the dataset is performed with the key masses still fixed. (1 click for model stopping).

The segmentation results for both models, as well as the manual segmentation results of an experienced user were compared with the given gold standard by different evaluation measures (Tab. 1).

### 6.3 Results

In all 11 datasets, the thyroid cartilage could be robustly segmented with an average border distance of 1.064 mm to the gold standard. No significant loss of segmentation quality could be found in the cases of pathological larynx shapes (Fig. 3(a)). Weak-signal holes in the cartilage were successfully bridged by the model’s intrinsic shape knowledge (Fig. 3(b)). The model adaptation time needed for elastic initialization was 0.5–1.5 minutes for all datasets, the model adaptation to the datasets took 2–4 minutes per dataset.

Our results show, that the volumetric MSM is not superior to the pure surface

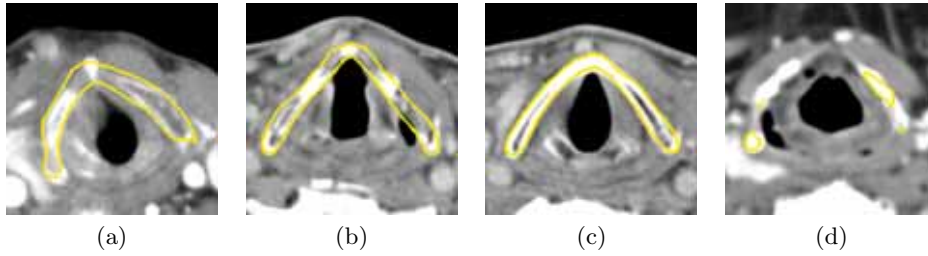
**Table 1.** Average segmentation results of the 2 models and 1 experienced user compared with the gold standard, measured for 11 CT datasets of the neck

Evaluation Measure	Volumetric MSM	Surface MSM	Experienced User
Oversegmentation	25.582 %	20.998 %	13.220 %
Undersegmentation	56.723 %	68.444 %	18,413 %
Hausdorff Distance	11.108 mm	9.843 mm	5,225 mm
Average Distance	1.195 mm	1.064 mm	0,174 mm

MSM. In fact, with intensity sensor weighting  $w_s > 0.001$ , the intensity sensors tend to be attracted by false gray values in neighbouring structures. This leads to strong model instability and significantly worse segmentation results. We therefore recommend using a pure surface model for the segmentation of the thyroid cartilage, which will be less affected by gray value inhomogeneities. The model’s adaptation to the datasets is significantly better in the lower part of the thyroid cartilage (i.e. below the adam’s apple) than near the upper border, which may cause up to 50 % of the undersegmentation. This can be attributed to several circumstances:

- The lower part of the thyroid cartilage tends to be signal-intensive, while the MSM may often lack signal support in the upper part. This often leads to strong undersegmentation of this area (Fig. 3(d)). This explains the high values for undersegmentation and the Hausdorff distance. A simple user interaction might prevent this behaviour.
- In some cases where the hyoid bone is very close to the thyroid cartilage, the upper sensors of the model are attracted by its strong signal. In order to prevent leaking, the os hyoideum may be subtracted from the dataset first. Another possibility is to integrate the hyoid bone into the model to ensure that the masses for the thyroid cartilage are kept at appropriate distance.

Except for the hyoid bone, no other neighbouring structures distracted the model. This must be attributed to the new method for initialization followed by shape learning. Without these techniques, the thyroid cartilage could not be separated from the thyroid gland, blood vessels and the trachea robustly.



**Fig. 3.** Results of the model adaptation to the datasets.

## 7 Discussion

A deformable 3D model has been created and adapted for the segmentation of the inhomogeneous and complex-shaped thyroid cartilage. The quality of the segmentation can probably compete with that of ASMs, but not yet with that of an experienced user. This can be expected to change with a solution for the undersegmentation of the upper border. Our experiments show that our method is principally very well suited to segment the thyroid cartilage:

Compared with a manual or LiveWire segmentation, the model offers a drastic reduction of interaction effort. Already now, the model can be used as a presegmentation of the cartilage, which needs only be corrected at 2–3 positions by the user. In contrast to other 3D models, such as Implicit Models and ASMs, interaction is supported by the explicit shape representation of our model. During all phases of the adaptation process, the user can drag the model’s masses to the right positions. This is facilitated by the intuitive concept of masses and springs, which make the model behaviour predictable even for novice users.

This model is adaptable to different shape variations by means of an elastic initialization followed by shape learning. In contrast to statistical shape models however, our method is not limited to a pre-learned range of shape variations. This makes it especially suited for segmenting pathological structures.

## References

1. Eugene N. Myers (ed.): Operative Otolaryngology. Volume I. Head and Neck Surgery. Publ.: W.B. Saunders Company, pp. 403–443
2. McInerney, T., Terzopoulos, D.: Deformable Models in Medical Image Analysis: A Survey. *Medical Image Analysis*, Vol. 1, No. 2, pp. 91–108, 1996.
3. Bardinet, E., Cohen, L.D., Ayache, N.: A Parametric Deformable Model To Fit Unstructured 3D Data. *CVIU*, Vol. 71, No. 1, pp. 39–54, 1998.
4. Cohen, I., Cohen, L.D., Ayache, N.: Using Deformable surfaces to Segment 3D Images and Infer Differential Structures. *CVGIP: Image Understanding*, Vol. 56, No. 2, pp. 242–263, 1992.
5. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Comparing Active Shape Models with Active Appearance Models. In *Proc. Br. Machine Vision Conf.*, 1999, pp. 173–182.

6. Bianchi, G., Solenthaler, B., Szekely, G., Harders, M.: Simultaneous Topology and Stiffness Identification for Mass-Spring Models Based on FEM Reference Deformations. In: MICCAI 2004, pp. 293–301.
7. Dornheim, L., Tönnies, K.D., Dornheim, J.: Stable dynamic 3d shape models. In: International Conference on Image Processing (ICIP 2005).
8. Dornheim, L., Tönnies, K.D.: Automatische Generierung dynamischer 3D-Modelle zur Segmentierung des linken Ventrikels in 3D-SPECT-Daten. In: H.-P. Meinzer, H. Handels, A. Horsch, and T. Tolxdorff, editors, Bildverarbeitung für die Medizin, Informatik aktuell, pages 138-142. Springer, 2005.
9. Garland, M., Heckbert, P.S.: Surface Simplification Using Quadric Error Metrics. In Proc. SIGGRAPH 1997, pp. 209–216.