Int J CARS
DOI 10.1007/s11548-010-0530-8

ORIGINAL ARTICLE

Complete fully automatic model-based segmentation of normal and pathological lymph nodes in CT data

Lars Dornheim · Jana Dornheim · Ivo Rössling

Received: 17 December 2009 / Accepted: 1 September 2010

Abstract
Purpose Exact and reproducible knowledge regarding the position, size, and type of the lymph nodes is often needed for tumor computer-aided diagnosis, treatment planning, and follow-up. An automatic segmentation method for CT data was developed that can identify and delineate normal as well as pathologically altered lymph nodes to satisfy this requirement.
Methods A semi-automatic lymph node segmentation method was developed using a 3D Stable Mass-Spring Model (SMSM), based on parallel simulation of the shape model on CT scan images. The models are started across the whole dataset at all potential lymph node positions but will only adapt to the data where a lymph node is found. The node positions can be determined by an evaluation of the model's quality of fit.
Results Systematically chosen lymph nodes in 5 CT datasets, including enlarged, necrotic, fuzzy-bounded, and deformed lymph nodes, were used to evaluate the segmentation algorithm performance. A test set of 29 lymph nodes taken from 4 typical lymph node regions were included. All lymph nodes were detected automatically, while an additional 31% false-positive ($n = 9$) candidates were detected. The average calculation time was 2 min per dataset. The segmentation accuracy was comparable to the inter-observer variance of human experts.
Conclusions Clinically relevant lymph nodes were detected within a few minutes and provided sufficient accuracy to demonstrate the feasibility of a new segmentation method. The test data were diverse, and the robust results suggest potential applicability to many kinds of lymph node abnormalities, except for extremely degenerated lymph nodes.

Keywords Dynamic models · Stable mass-spring models · Lymph node segmentation · Lymph node detection

L. Dornheim (✉)
Dornheim Medical Images GmbH,
Universitätsplatz 2, 39106 Magdeburg, Germany
e-mail: lars@dornheim-medical-images.de

J. Dornheim · I. Rössling
Institute of Simulation and Graphics, Otto-von-Guericke University,
Universitätsplatz 2, 39106 Magdeburg, Germany
e-mail: Jana.Dornheim@isg.cs.uni-magdeburg.de

I. Rössling
e-mail: iroess@isg.cs.uni-magdeburg.de

Introduction

Application

The assessment of a tumor disease (TNM classification[1]) and the following therapy decision and evaluation depend not only on the tumor itself, but also on the lymph node situation w.r.t. size, dignity, and infiltration of neighboring tissue [1]. If exact analysis, visualizations, or measurements are required, the lymph nodes have to be segmented (e.g., in CT data), which, in general, is a most time-consuming, tedious, and error-prone procedure [2,3]. A complete fully automatic segmentation of all lymph nodes in a whole body region (e.g., the neck) would be most appreciated in this case.

Technology

A complete automatic detection process consists of a detection of all lymph nodes, followed by or combined with a

[1] The international normed TNM classification describes the state of the tumor (T), the lymph nodes (N), and possible metastases (N).

Published online: 08 October 2010

segmentation of each finding. In this paper, both tasks will be addressed.

Object detection in medical images is an important and difficult problem. Due to the variability of medical structures and due to the often low image quality of medical image data, object detection tasks are currently carried out manually in most cases, as no adequate automatic techniques exist. This makes many image analysis tasks (e.g., in the preparation of datasets by segmentation of many structures) very time-consuming or may lead to increased error rates, when time is short in the clinical routine and objects are overseen (see study in [4]).

This article builds upon prior work on the segmentation with dynamic shape models (e.g., [5]) and employs this robust technique for the object detection as well. We verify these model's appropriateness for local search tasks and employ it for the development of a global search process.

In this work, we use Stable 3D Mass-Spring Models (SMSM) that were introduced by Dornheim et al. in [6]. These models are a special type of mass-spring models that have been extended by an additional torsion force that provides the direct possibility of flexible shape maintenance, in order to prevent shape collapse during the model adaptation. They are force-driven and are simulated time-discretely according to Newtonian mechanics. The model's relation to the dataset is implemented by sensors at the masses, which create forces according to the local image information, which pull the model toward specific features and structures in the dataset, while concurrently the model's shape is preserved by the torsion and spring forces.

State of the art

Lymph node segmentation

Rogowska et al. [7] analyzed several lymph node segmentation techniques and stated that reliable lymph node segmentation is only possible with the use of model knowledge. A fast-marching lymph node segmentation approach presented by Yan et al. [8] requires complex user interaction (barrier placement, etc.) to prevent leaking. The active surfaces approach proposed by Honea et al. [9] uses more complex model knowledge but was never evaluated on real CT datasets, but only on idealized phantoms.

The first relatively robust method [10] uses a Stable Mass-Spring Model (SMSM) consisting of two layers, which integrates shape and appearance knowledge and requires only a starting point close to or inside the target lymph node. The quality of the segmentation results lies slightly above the inter-personal variability. Enlarged, necrotic, or fuzzy-bounded lymph nodes were not addressed, which is however very important for tumor diagnosis. A similar approach was presented by Maleike et al. [11], who used a statistical instead of a Stable Mass-Spring Model. This statistical model uses very few shape modes to restrict the model to rather elliptical shapes. The results of this method are comparable to [10], but a lot more user interaction (rough scribbling of the lymph node interior on the dataset) is needed before the segmentation. In some cases, a manual correction is recommended.

Lymph node detection

Object detection in image data is a wide field of research. Much research in this area has been directed at the recognition and tracking of objects in projective, two-dimensional images (e.g., photos), in which occlusion and varying views of objects are a problem; however, the size of the data is rather small. These approaches are therefore focused on the identification and relation of specific, discriminable object features [12–15]. Global optimization techniques trying to fit models of different nature to the target objects are also often found in this category [16–19].

For object search in large medical volume data, those approaches are however often inappropriate, because their global optimization approach is too inefficient for large datasets. Furthermore, the object shape is not represented directly, even though it is often an essential recognition feature, as it is not affected by occlusion and different viewing directions. Besides, these models are hard to construct for specific detection tasks.

In some cases, complex filter pipelines are used to detect certain kinds of structures. E.g., [20] presents a method that employs a 3D minimum directional difference filter for enhancing blob structures with suppressing line structures, before context information is applied. But without specific knowledge about the object's shape and appearance, the detection results were not satisfying in any case (57.0% detection rate of enlarged lymph nodes with 58 false positives per case).

For search tasks in large medical volume data, three-dimensional shape models are of interest, such as they are often applied to segmentation tasks on medical data. They are able to deal with the large size of the datasets and describe the shape and gray value specifics of medical structures. Single object parts are often not specific enough to draw conclusions about the whole object. Typical shape models in this area, besides the SMSMs mentioned in the introduction, are Active Shape Models [21], Simplex Meshes [22], and super quadrics [23].

For detection tasks, SMSMs are very appropriate, as they can be created easily and efficiently for many different types of target structures [24]. Furthermore, they are locally flexible w.r.t. their shape, which complies with the nature of medical structures, and allow for a direct measurement of their quality of fit [25]. In [26], the direct predecessors of

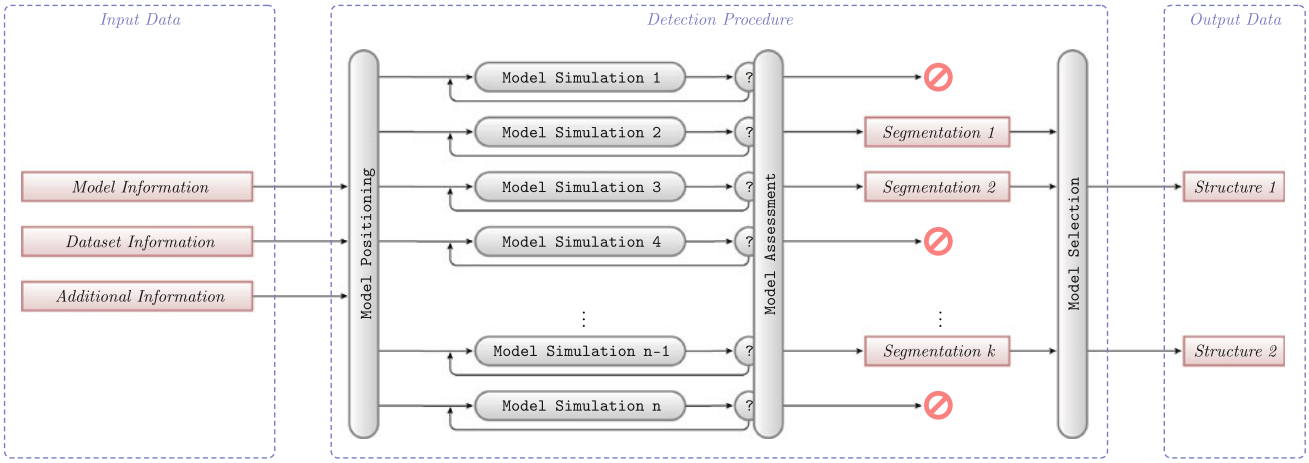

Fig. 1 Schematic overview of the complete fully automatic lymph node segmentation process

the SMSMs have been applied for object classification in the context of a stochastic search of different object parts. We do not know of any other approaches to use other shape models for object detection.

Methods

The process of automatically detecting all structure of a certain type (lymph nodes in our application case) in a whole dataset or a dataset region is schematically depicted in Fig. 1. The search process needs three kinds of inputs:

1. The dataset or dataset region, which is searched for the specific structures.
2. The specific structure type to search for in form of a dynamic model describing its shape and appearance.
3. Additional search information like the number of expected structures and relevant sub-regions.

As the search process starts, the whole input data are analyzed and an adequate partitioning of the dataset into cells is determined. The dynamic model is cloned, once for each created cell of the dataset. The cells are chosen small enough, so that a single dynamic model is always able to find any target structure in its associated cell.

Now, in each cell, a complete local search is performed by the associated clone of the dynamic model. Using additional search information, some of the cloned models can be skipped for performance reasons. If a dynamic model has adapted successfully to the data, then the resulting segmentation describes a detected structure; otherwise no target structure was found in the corresponding cell. The search process can run in parallel and has finished, when all cloned models have been skipped or finished their own adaptions to the data, either successful or without success. At last, only a subset of the successful adapted models can be selected to avoid duplicate findings to ensure a certain number of detected structures.

In the following sections, first the used lymph node model and its dynamics and components are introduced (“Lymph node model”). After that, the whole search process (“Complete automatic segmentation process”), which this model is used in, is described.

Lymph node model

In this section, we first present a segmentation model for lymph nodes in CT data, which is based on the method from [10], but extends it considerably w.r.t. accuracy and support of enlarged, necrotic, fuzzy-bounded, and deformed lymph nodes. This lymph node segmentation technique requires an initialization step, which can be carried out manually or automatically. In this work, a fully automatic initialization is of interest, which will be addressed later in “Complete automatic segmentation process” on the description of the object detection framework, which will use the following method for the segmentation of the single lymph nodes.

Model design

We use a three-dimensional Stable Mass-Spring Model (SMSM) for the segmentation of the lymph nodes, which is a mass-spring model with extended dynamics to achieve a more controllable and stable behavior in 3D (see “Model dynamics”). This model consists of two layers and is initially sphere-shaped (Fig. 2). Due to its physics-based nature, it can deform elastically to model the ellipsoidal or bean shape of real lymph nodes. The outer model layer has edge sensors at the mesh vertices, which strive toward edges in the CT dataset that have an orientation parallel to the model’s surface. In contrast to this, the vertices of the inner model layer wear

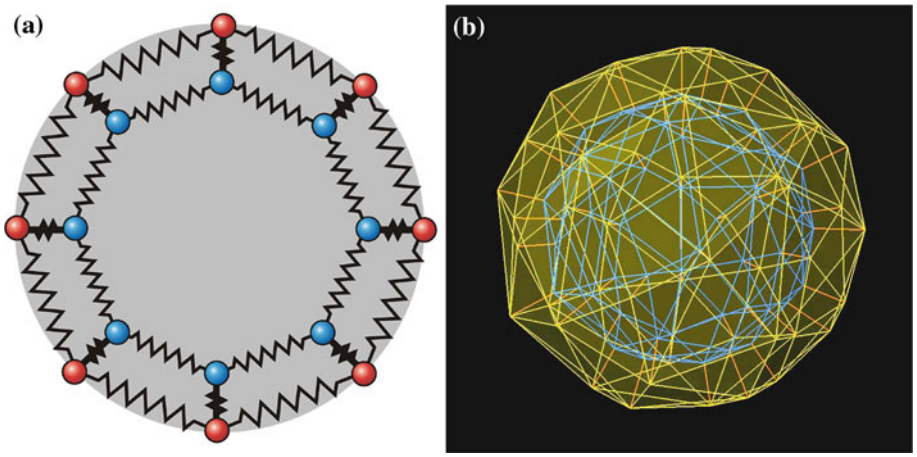

Fig. 2 **a** Schematic 2D view of the two layers of the 3D lymph node model (SMSM). Outside, the edge sensors (*red*) and inside the gray value sensors (*blue*) are shown. Bold springs are especially stiff. **b** 3D view of the real lymph node model

intensity sensors, which strive toward the typical gray value range of lymph nodes in CT datasets.

During the model simulation, the model adapts to the lymph node's edges by means of the gradient sensors on its outer layer, while the intensity sensors on its inner layer keep the model on the lymph nodes, so it cannot drift away or leak out. Meanwhile, the shape-preserving torsion forces of the model ensure that the model keeps a rough lymph node shape according to the chosen weighting of dataset fitting versus shape maintenance.

Model dynamics

We use Stable Mass-Spring Models (SMSMs, see [6]), an extended kind of mass-spring models, that enables a stable dynamic behavior in 3D. They adapt to structures in image data by balancing their internal and external forces during a discrete time simulation. While the internal forces sustain the model's size and shape knowledge, the external forces drive the model to certain image features. This way, a weightable compromise between the image data and the model knowledge is achieved, so that missing image information can be replaced by appropriate expectations.

In contrast to conventional mass-spring models, they use an additional kind of internal force (the torsion force), which can be used to explicitly control the model's shape, while the spring forces control the scaling of the model. Furthermore, measurement of these forces allows separated conclusions about the model's current shape and scale deviation. In the following paragraphs, the model mechanics are described in more detail.

Mass-spring models

Mass-spring models are dynamic, physics-based models known from image analysis literature (e.g., [26]). They are a theoretical model representing a dynamic system of *mass points* interconnected by elastic *springs*. Besides the physical parameters, i.e., the masses m_i of the mass points i, and the rest lengths $l_{0_{ij}}$, and spring constants k_{ij} (force exerted per length difference) of the single springs, the connection topology of the masses and springs plays a central role.

The dynamics of such a system can be described by Newtonian Mechanics. The motion of the mass points is thus solely influenced by the forces acting on the mass points. In this simple theoretical and ideal system, only *internal forces* exist, which are the spring forces $\mathbf{F}_{ij}$ (Eq. 1) being exerted on the mass points by the elastically deformed springs. Their value depends on the positions of the incident mass points i and j (here represented by position vectors s_i and s_j).

$$\mathbf{F}_{ij} = k_{ij} \cdot \left(\left\| \mathbf{s}_j - \mathbf{s}_i \right\| - l_{0_{ij}} \right) \cdot \frac{\mathbf{s}_j - \mathbf{s}_i}{\left\| \mathbf{s}_j - \mathbf{s}_i \right\|} \tag{1}$$

As it is common in literature (e.g., in [27]), the model dynamics shall not be calculated exactly by differential equations. Instead, it is simulated in discrete time steps of distance $\triangle t$. Starting from the speed $\mathbf{v}_{i_t}$ of mass i at time t, its speed at the time $t + \triangle t$ is calculated. Equation 2 shows the motion equation involving the spring forces summed up for all masses j that are connected to the mass i. This notation of the summation of connected masses shall be kept throughout this paper. In addition, a damping factor d exists in this equation, which is motivated by the time-discrete simulation of the system. This significantly reduces problems in reaching a stable equilibrium.

Image information is integrated through *external forces* exerted on the mass points through sensors. Each sensor i creates a force $\mathbf{F}_i$ depending on the image data. Thereto, the position of mass $\mathbf{s}_i$ as well as the type and parametrization of the mass' sensor is relevant. Furthermore, this motion equation contains weight coefficients, where w_f weights the spring force component and $w_s(i)$ is the coefficient for the sensor force component. The latter depends on i insofar,

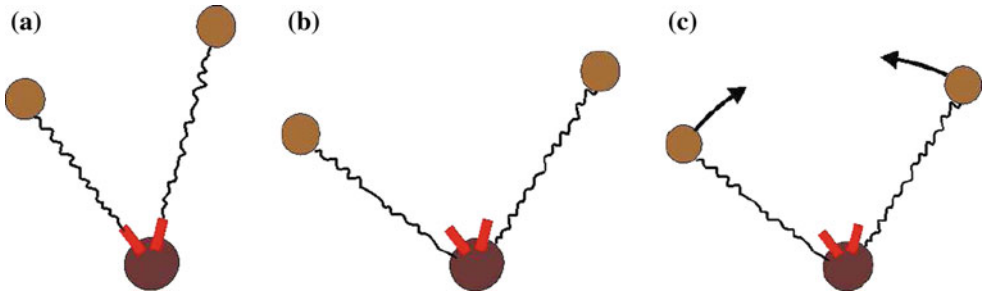

Fig. 3 **a** shows the springs in their rest directions (marked at the mass). In **b**, these springs are contorted by some external force. As a result, in **c**, a torque moment acts on them in direction of their rest directions

as different types of sensors suggest different and independent weight coefficients.

$$\mathbf{v}_{i_{t+\triangle t}} = \left(\mathbf{v}_{i_t} + \frac{w_f \cdot \sum_j \mathbf{F}_{ij} + w_s(i) \cdot \mathbf{F}_i}{m_i} \cdot \triangle t \right) \cdot d \qquad (2)$$

Torsion forces

In [26], Bergner already recognized the instability problem of explicit dynamic 2D models and suggests to its solution, besides a dense cross-linking, the use of an angle force which he refers to as *torque force*. Dense cross-linking (especially in higher dimensions) solves the instability problem only at the expense of high model complexity and only indirectly, for no additional control over the degree of stability is gained. The *angle forces* in [26] do only consider angles between two springs with a common mass. In higher dimensions, this allows the angle plane to turn, so that a multitude of additional angle forces would be necessary, leading to the same complexity problems as the dense cross-linking.

The solution to this problem is the introduction of (normalized) *rest directions* $r_{0_{ij}}$ of the springs starting from one mass i to all its adjacent mass j along the lines of the spring rest lengths. Spring contortions, i.e., deviations from their rest directions, can then be compensated by opposed torque moments as Fig. 3 illustrates.

These torque moments manifest themselves in *torsion forces* $\mathbf{F}_{(i,j)}$ (Eq. 3), whose values depend on the mass-specific torsion constant t_i (torque moment per torsion angle).

$$\mathbf{F}_{(i,j)} = \frac{t_i \cdot \left| \angle(\mathbf{r}_{ij}, \mathbf{r}_{0_{ij}}) \right|}{\left\| \mathbf{r}_{ij} \right\|} \cdot \frac{\mathbf{n}_{ij}}{\left\| \mathbf{n}_{ij} \right\|} \quad \text{with} \quad \mathbf{r}_{ij} = \mathbf{s}_j - \mathbf{s}_i \qquad (3)$$

They act upon the mass j tangentially to the motion curve in order to compensate the torsion. The working direction of the torsion forces $\mathbf{n}_{ij}$ is calculated as Eq. 4 shows.

$$\mathbf{n}_{ij} = \mathbf{r}_{0_{ij}} - \frac{\langle \mathbf{r}_{ij}, \mathbf{r}_{0_{ij}} \rangle}{\left\| \mathbf{r}_{ij} \right\|^2} \cdot \mathbf{r}_{ij} \qquad (4)$$

Stabilization works successfully in arbitrarily high dimensions. It allows an exact control of the shape stability via a weight coefficient w_t of the torsion forces. It requires as many torsion calculations per mass, as springs are connected to this mass. Thanks to the absolute rest directions of all springs, there is no danger of mutually shifting angle planes. So, the problem of shape collapse and contortion in higher dimensions is no longer relevant (for details see [6]). Equation 5 shows the motion equation extended by the torsion forces.

$$\mathbf{v}_{i_{t+\triangle t}} = \left(\mathbf{v}_{i_t} + \frac{w_f \cdot \sum_j \mathbf{F}_{ij} + w_t \cdot \sum_j \mathbf{F}_{(j,i)} + w_s(i) \cdot \mathbf{F}_i}{m_i} \cdot \triangle t \right) \cdot d \qquad (5)$$

The *torsion forces* belong to the *internal forces*, as they code knowledge about the model's shape. Mass-spring models having these forces are called *Stable Mass-Spring Models* (SMSMs). With sparse cross-linking, it can be assumed that the springs define the size of the modeled object, whereas the spring directions model its shape. Shape and size of a model can therefore be weighted and measured individually with respect to their influence on the model adaptation. Furthermore, the existence of rest directions for each mass allows easily for direction-dependent sensors. Therefore, edge sensors of SMSMs are normally direction-weighted, which leads to lower edge sensor values the more the edge's normal direction differs from the sensor's mass' rest direction (see [28] for further details).

External forces

The external forces $\mathbf{F}_{\mathbf{ext}}$ are induced by sensors k attached to the masses of the model. In general, many kinds of different sensors are thinkable, but on voxel image data, normally voxel-based sensors are used, which compute their forces as gradients on the original or preprocessed variants I of this image data on the position $\mathbf{s}_i$ of the mass i they are attached to, as Eq. 6 shows.

$$\mathbf{F}_{\mathbf{ext}_k} = \nabla I(\mathbf{s}_i) \qquad (6)$$

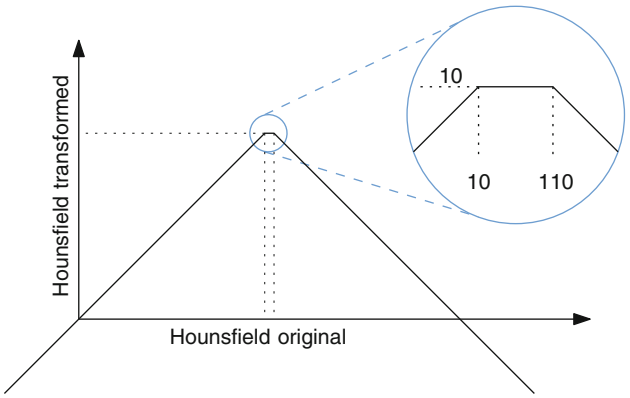

Fig. 4 Schematic course of a plateau filter function

This preprocessing can depend on the sensor status, e. g. its direction $\mathbf{d}_k$. In the case of edge sensors, this direction is the normalized normal direction of the model border. A border offset parameter b_k was introduced within this work to move the sensor position along its parent mass' normal back and forth, as can be seen in Eq. 7. This way, the model border can be pushed along its normal directions to grow or shrink the model to compensate systematical over or under segmentations. This is an important control option, since ground truth segmentations normally do not have their borders exactly on the highest gradient values in the image data.

$$\mathbf{F}_{\mathbf{ext}_k} = \bigtriangledown I\,(\mathbf{s}_i + b_k \cdot \mathbf{d}_k) \tag{7}$$

Sensor filter design

In contrast to the technique in [10], we filter the CT datasets (their intensities and gray values, respectively) by means of plateau transformations, before the sensors of the model access them. In the case of the intensity sensors, a transformation variant is used, which sets all pixels within the lymph node Hounsfield range [2] to its base value 10. All values below 10 are not changed, and all values above 110 are flipped on the center value of the lymph node intensity interval, in our case 60 (see Fig. 4).

This way, it is guaranteed that all gradients in the dataset point in direction of the lymph node intensity range, which has no further gradients inside. So all gradient-based intensity sensors strive into the lymph nodes but create no disturbing and unwanted forces when they are inside the lymph nodes.

Especially for necrotic lymph nodes (like in Figs. 5 and 6), this behavior is very important and a main improvement compared to the base method, because they have strong inner gray value variations, which normally would strongly disturb the model. If otherwise the gray value interval is filtered using rigid limits (rectangle-shaped function) like in the base method, then strong forces that deform the model too much will appear due to the very strong gradients at the filtered, binary gray value interval borders. Furthermore, there will be no forces outside the lymph nodes, which could pull the intensity sensors back into the lymph node, if they have driven away. Also, if the gray value interval does not fit perfectly to an individual lymph node, its borders cannot be found exactly by the model if they lie slightly outside this interval, since the model will be kept strongly inside this interval behind the binary, filtered gray value interval border (see Fig. 5b).

[2] From 10 to 110 in our datasets, as determined by our clinical partners.

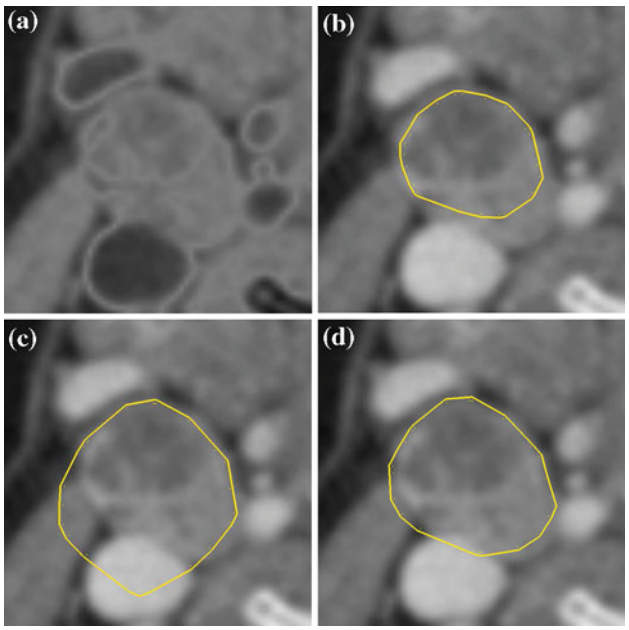

Fig. 5 **a** plateau-transformed dataset. **b** model adaption using a rectangle-filtered gray value range for the intensity sensors. model adaption using plateau-filtered intensity sensors in **c** without and in **d** with plateau filtering for the edge sensors

Also the edge sensors use a special, plateau-filtered dataset instead of the original dataset. This plateau transformation is different from the one used for the intensity sensors: The plateau is degenerated to a peak at the top of the lymph node gray value interval, which is 110 in our case. So, this special case of the plateau filtering is carried out with a triangle-shaped function. By this technique, it is achieved that the orientation of all edges in the dataset within the gray value range above 110 is flipped, but their size and strength is not changed, so that no edges disappear or show gray value interruptions caused by the filtering.

After the plateau filtering, the lymph node's gray value range is the most intense range in the dataset. Therefore, it can be assumed that all lymph node border edges are oriented (w.r.t. their gradient) toward the inside of the lymph node, even if they were adjacent to a brighter, contrasted vessel in the original dataset (see the plateau-transformed dataset in Fig. 5a in contrast to the original dataset in b).

Furthermore, it can be assumed that all relevant lymph node border edge information does not lie outside the lymph node Hounsfield interval, extended by a certain tolerance for the border value range. We use 50% of the lymph node interval from 10 to 110 as this tolerance, leading us to our edge Hounsfield interval of −15 to 135. All edges outside this interval will be filtered out.

This way, the edge sensors on the outer model layer can be adjusted to detect only border edges oriented inside, discarding most of the useless, distracting edge information in the dataset without losing any lymph node border edges. Without plateau filtering, either the outside-oriented edges would also have to be considered, which would additionally distract the model, or the edges to brighter neighboring structures would be ignored and not found by the model. In the last case, an unwanted drifting toward these brighter structures is often observed (like in Fig. 5c).

Multi-model strategy

When trying to segment lymph nodes in general, it will be noticed that lymph nodes can have very different appearances. They can be small or large, elliptical or severely deformed, evenly gray or strongly textured, etc.. Especially lymph nodes that are relevant for tumor assessment are often enlarged, deformed, or even necrotic (i.e., strongly textured or at least with dark interior areas). From a computer vision point of view, normal and necrotic lymph nodes must be handled as different classes of objects as respects scaling, shape variation, and texture.

SMSMs are prototypical models, which are verifiably very well suited for the segmentation of such a class of objects that are characterized by a specific shape, appearance, and also scaling, if wanted. In contrast to this, statistical models (ASMs, etc.) can represent different object classes at the same time, but in practice they also have problems, if these classes differ too much from each other. In general, for single specific classes, specifically designed models perform better than general models, which also model other classes at the same time.

Hence, to adequately deal with the different classes of lymph node appearances, we use a set of single lymph node models as a virtual overall model in a concurrent segmentation strategy, where each single model is specifically designed for a certain type of lymph node. For the lymph node segmentation, one of the most important class partitions is the one by size. So, we use a set of models for lymph nodes of different sizes (see Fig. 6). Furthermore, there are also models thinkable for different lymph node textures, shapes, etc.

For the multi-model segmentation strategy, all models of the set will be started at the same position (the start position of the virtual overall model) in parallel. After the individual model adaptions have converged and finished, the best fitting model of all started models is chosen as the segmentation result of the virtual overall model. So, besides the pure segmentation result, it is also possible to determine the class of the segmented lymph node, because it is known which lymph node class the chosen model belongs to.

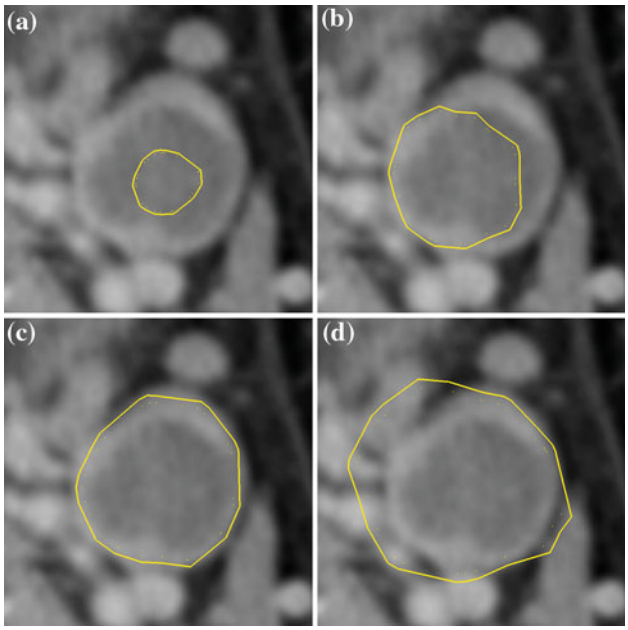

Fig. 6 Results of the multi-model segmentation using modes of different scaling level **a** factor 3, **b** factor 5, **c** factor 7, **d** factor 9

The best fitting model is determined using a quality of fit calculation, which considers both the correspondence of the adapted model with the dataset (by evaluating the current sensor values of the adapted model) and its deformation (by using the still existing shape forces of the adapted model). The result is calculated as a value between 0 (not fitting) and 1 (very well fitting). In [25], it is explained how the quality of fit can be calculated for SMSM, which are used in our method. Furthermore it states that these models are especially suited for the calculation of such a quality of fit, which was one important reason to choose this model type for our segmentation task.

Complete automatic segmentation process

In this section, we present the overall search process. It is a general method, which can be used to detect arbitrary known structures in medical volume data safely and automatically. Thereto, these structures need to be adequately describable by a dynamic shape model, which we will explain in "Requirements". For our application task, the lymph node shape model described in "Lymph node model" is used.

The search itself is based on a parallel adaptation of several instances of this shape model to the image data. All model instances are systematically distributed across the complete

dataset at all potential lymph node positions. This fully automatic distribution serves as the required initialization step for the models, as introduced in "Lymph node model".

Requirements

The global search process consists of independent local search processes, whose individual action ranges are ideally a partition of the search space. It is essential for the correctness of the search result that the complete search space is covered without gaps by this partition, whereas the pairwise disjunctiveness is only relevant for the efficiency of the search. For the sake of simplicity and efficiency, it is additionally desirable that all individual action ranges have the same shape, size, and orientation, so that they can be treated uniformly. For this reason, we chose a partition of the volumetric search space into congruent, axis-parallel cuboids.

In order for a model to successfully perform a local search in its action range, it must fulfill three requirements: consistency with its target structure, monotonicity of the search space, and measurability of the quality of fit. These are explained in the following sections.

Consistency with the target structure

The model must consistently represent its target structure, i.e., the target structure must be a local minimum in the search range of the model, w.r.t. the objective function given implicitly by its motion equations. This ensures that the target structure is always found, if the search initialization is sufficiently close. Thereto, the model must reliably represent the shape and gray value features of the target structure in order to discriminate it from other structures. At the same time, it must be flexible enough to model unexpected occurrences of the target structure (e.g., by noise, soft tissue deformation, natural shape variability, or smaller pathological variations).

Dornheim et al. [24] describes the construction of SMSMs from sample segmentations, which quasi represent the expectancy of the model knowledge, but do not limit it w.r.t. its system-imposed local variability, which is necessary for modeling the mentioned occurrences. It shows how SMSMs can efficiently be constructed for target structures, so that they are adequate for segmentation. The consistency with the target structure is a prerequisite for that, as otherwise the model would not converge to the target structure, even for optimal initial placement, which will necessarily lead to a bad segmentation.

Monotonicity of the search space

The model should reliably find a target structure in its action range, if it is placed inside this structure. This is only the case, if the target structure is the monotonously closest local minimum of the objective function of the model in its search range [3]. This means in particular that there must not be more than one target structure inside the action range of one model, because in this case at least one target structure cannot be found by the model. This monotonicity only has to be given for an action range, if indeed a target structure is present in this action range.

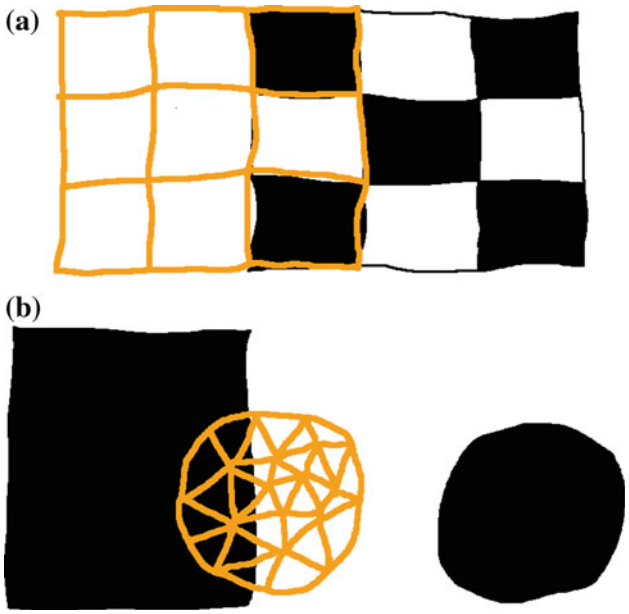

Fig. 7 **a** The orange chess board model has converged to a false, internal local minimum of the objective function. **b** The orange model of the *black circle* was attracted by the neighboring *black quadrilateral*

If only one target structure exists in one action range, there may be multiple reasons for the occurrence of monotonously closer local minima. On the one hand, there may be additional local minima located between the starting point of the model and the target structure, which are also caused by the target structure (e.g., self-repeating patterns like in a chess board, see Fig. 7a) Besides these internal minima, there can be external ones, which result from other structures in the action range, distracting the model completely or partly, due to the close neighborhood to the starting point (see Fig. 7b for an example of a prominent adjacent structure).

For an efficient global search, a partition of the complete search space is desirable, which partitions it in few large local search ranges. However, the larger each action range is, the more local minima can occur inside it. In reality, for SMSMs describing compact objects, the biggest search range without local minima lies normally around the half of the target structure size. It means that such a model finds its target easily, if it is started somewhere in the inner half of its target structure, as [10] states.

[3] For the sake of simplicity, we identify the position of a target structure with its barycenter.

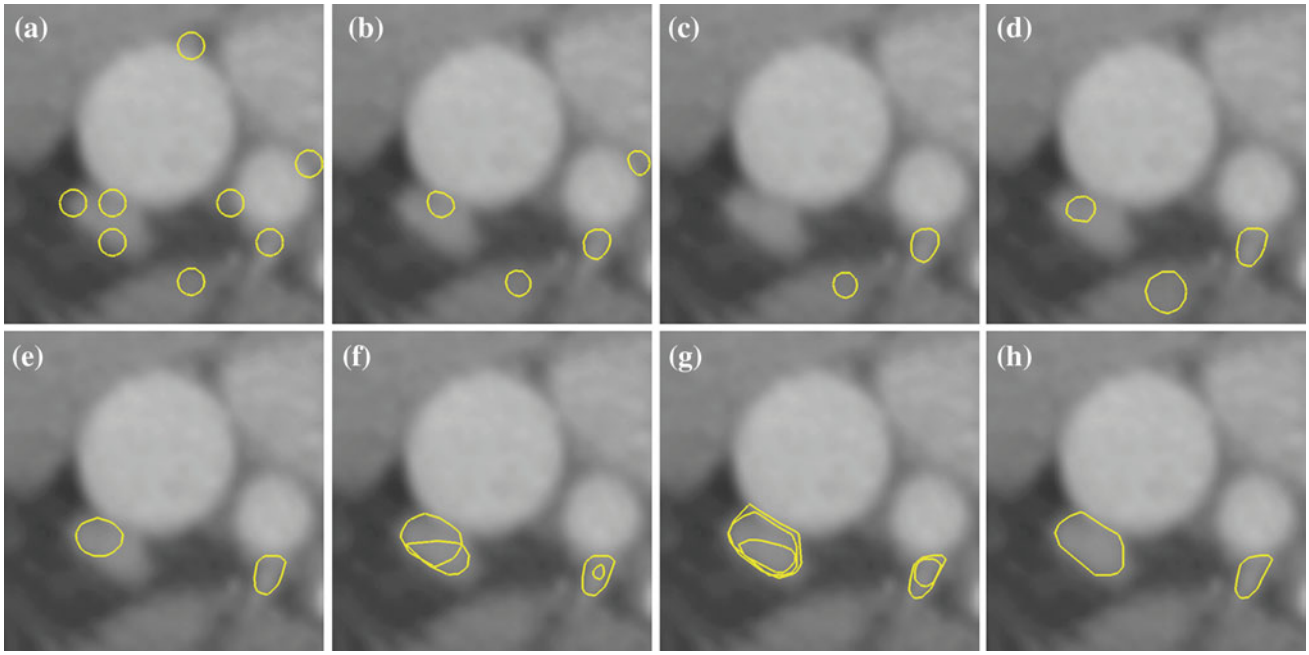

Fig. 8 Lymph node detection process. **a** Initial placement of the model population according to the expectation map ("Reduction in the population"). **b–f** Progressing simulation, models are rejected and extend from other slices into the current slice, respectively. **g** End of the search, two structures have been detected by more than one model. **h** Extraction of the best adapted representative of each cluster of structure models

Measurability of the quality of fit

Another prerequisite for the applicability of the models to local search is the possibility to measure the quality of the model's adaptation to the target structure in the dataset. This quality of fit expresses how well the shape and gray value properties expected by the model were found in the dataset, i.e., how much the structure the model has adapted to resemble the target structure. Thereto, information about the compliance of the adapted model with the dataset, as well as information about the occurred deformation, must be accounted for.

In [25], this information is efficiently computed for SMSMs, which are pointed out to be especially suited for the calculation of the quality of fit. These pieces of information are combined to a quality of fit function, which can take values between 0 (bad quality of fit) and 1 (perfect quality of fit). The parameters for this function can be estimated automatically from a successful adaptation of the model to its target structure. This quality of fit function was evaluated by means of a plausibility consideration of the possible model behavior and a series of experiments.

The search process

The model-based object search technique presented here uses the capability of SMSMs for a local search. By means of a population of several model instances with individual action ranges, a global search is achieved. For the success and efficiency of the search, in particular the composition and subsequent reduction of this population are critical.

The search procedure

The search process starts with the generation and placement of the model population. Thereto, we distribute instances of the search model across the dataset in a rectangular, axis-parallel grid, so that their local search ranges cover the complete dataset without gaps (see "Requirements").

Now the simulation of the dynamic models is started simultaneously and is continued, until their movements converge and the models come to a rest. Fig. 8 gives an impression of this population life cycle. After the simulation, all models that do not bear a given quality of fit are rejected. All other models have adapted successfully to a target structure.

Grouping of the results

Especially in the case of larger target structures, it often happens that a structure lies in the action range of more than one model and can therefore be detected by more than one model. In this case, models that detected the same structure need to be grouped after the simulation and reduced to one final candidate. As an indication for the adaptation of two

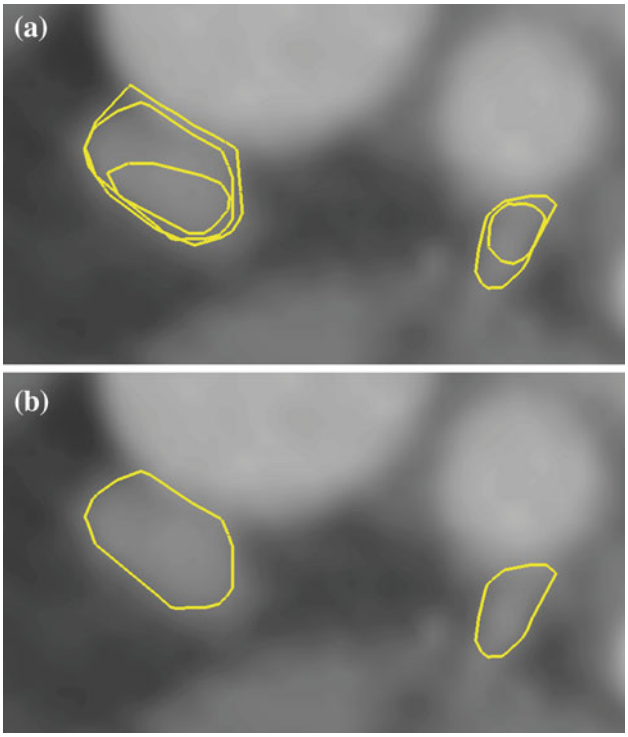

Fig. 9 **a** Two structures are each found by several models started from different grid points. **b** After grouping of the models, for each structure, the model with the best quality of fit was selected

models to a common structure, a volume-based subset coefficient k is introduced (see Eq. 8), which is 1 in case of a full subset relation of the models M_1 and M_2 and 0 in the case of no intersection of these models. Here, the subset relation is used intentionally instead of the equality relation for the description of the model grouping, because a subset leads to an unambiguous grouping.

$$k = \max\left\{\frac{V_{M_1 \cap M_2}}{V_{M_1}}, \frac{V_{M_1 \cap M_2}}{V_{M_2}}\right\} \tag{8}$$

After the grouping of the successfully adapted models on behalf of a high threshold k, the individual with the best quality of fit is chosen as representative out of each group and the rest of the group is discarded. This way, besides a good search result, a segmentation result as good as possible is ensured, because high qualities of fit are normally connected to good segmentation results (see [25]). Figure 9 shows the effect of this model grouping.

If the number of target structures in a dataset is known (e.g., exactly one left ventricle is expected in a thorax dataset), then this number of expected findings can exactly be chosen from the set of individuals by only keeping the best fitting individuals until this expected number is reached.

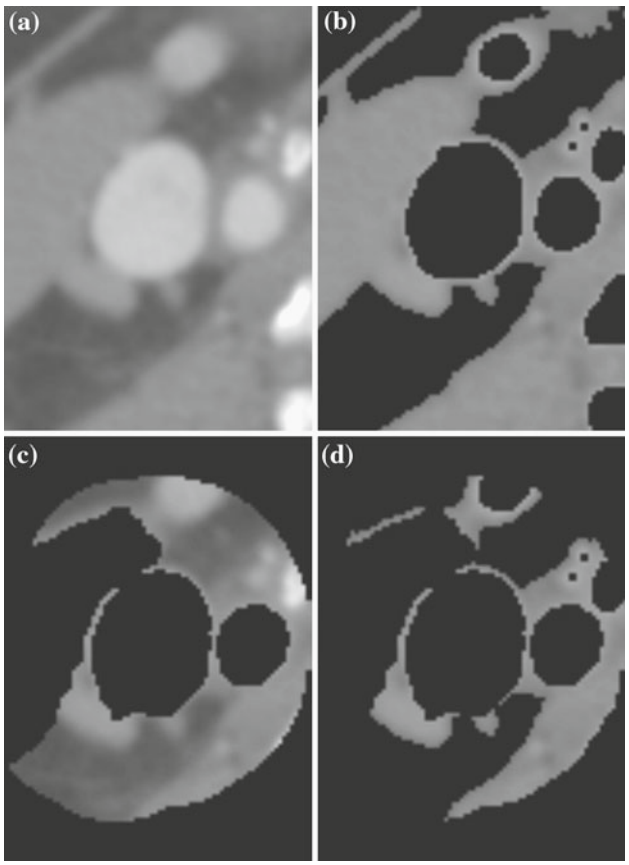

Fig. 10 Original CT slice of the neck and three variants filtered by different expectation maps. **a** Original. **b** *Gray value* expectation map of the lymph nodes. **c** Anatomic (positional) expectation maps of the lymph nodes. **d** AND-combined expectation maps (**b**) and (**c**)

Reduction of the population

Due to the necessary grid size for very small target structures (e.g., lymph nodes) on real medical volume data, very large model populations may be necessary, which cause very high calculation times for the above-described search process. However, the majority of all model instances do not find a target structure. The earlier this is identified, the earlier these models should be removed from the model population in order to increase the efficiency of the whole search process. This circumstance can be identified on the basis of several criteria that are now described in more detail.

Expectation Maps The target structures in medical datasets are in general not evenly distributed across the dataset. Almost always they are only found in specific anatomical regions. These regions can be described by a certain gray value range, or a spatial relationship to certain landmarks or anatomical structures (see Fig. 10). In addition, knowledge from preceding segmentations of other structures can be used to exclude regions from the area of potential target structure locations. The acceptable search regions can be described

Table 1 Parameter of the available neck CT datasets

Dataset	Voxel count		Voxel size			Contrast agent	Manufacturer
	X/*Y*	*Z*	*X*/*Y* (mm)	Slice distance (mm)	Slice thickness (mm)		
1	512	65	0.28	3	3	yes	Siemens
2	512	61	0.45	3	3	no	Siemens
3	512	63	0.42	3	3	yes	GE
4	512	262	0.47	0.7	1	yes	Philips
5	512	161	0.41	1.5	3	yes	Philips

The datasets were chosen to be as divers as possible

by binary *expectation maps*. Expectation maps can be calculated in different ways, depending on the kind of knowledge they are based on. For instance, the region inside a certain distance around a given, already segmented structure can be marked as an expectation map (Fig. 10). By a binary combination (AND) of several expectation maps achieved in this way, the global search space can be reduced significantly, by immediately removing models from the population that leave this expectation range during the search.

Quality of fit Besides the final determination of the successfully fitted models, the quality of fit can also be used to reduce the population during the search process. As successful models have in most cases been placed close to the target structure (often with significant overlapping), they have already an increased adaptation quality at the beginning of the simulation process, which increases during their adaptation. We can therefore specify a lower threshold for the quality of fit, below which models are removed from the population. This lower acceptance threshold on the quality of fit can be increased progressively during the search process, up to the final threshold on the quality of fit, which is applied after the simulation process ("The search procedure"), so that after a certain number of simulation steps only successfully adapted models remain in the population.

Scaling After the successful adaptation, the scaling of the adapted model can also be used to reduce the population size, if it is known that the target structures do not exceed or go below a certain size threshold. However, in this case, the population reduction has the goal of excluding false-positive candidates, instead of increasing the calculatory efficiency. Along these lines, several other features are thinkable, which could be used to identify false-positive results after the end of the search process.

Evaluation

In this section, the methods described in this paper will be evaluated. Before the complete automatic segmentation process as a whole is evaluated, the relevant parts of it will be examined separately to gain more detailed insight into the complete method. To get reliable results, we used very diverse data.

Test environment

The experiments were carried out on 5 neck CT datasets from scanners of three different manufacturers. Some of the datasets were acquired with a contrast agent and some without. The slice distances varied from 0.7 to 3 mm, and the resolution was between 0.28 and 0.47 mm. For details, see Table 1. An isotropic resampling (to the slice resolution) was necessary for the Stable Mass-Spring Models, because their sensors rely on gradient computations, which need cubic voxels for equidistant sensor range in every direction. The gold standard were expert segmentations additionally approved by radiologists.

All experiments were carried out on a modern standard PC (Intel Core2 Quad Q9550, 2.83 GHz, 4 GB RAM). For the single-model segmentation evaluation, one processor was used. In contrast, for the search process tests, all four processors were used in parallel to simulate the different models for performance reasons.

Single lymph node segmentation

Table 2 shows the results of the direct evaluation of the single lymph node model. Here, the segmentations of the human experts of 40 different lymph nodes from 5 neck CT datasets are compared to the results of the base method from [10] on the same lymph nodes and the newly developed method in this paper. The lymph nodes were equally chosen from 4 categories addressing the standard case and the typical problems from an image analysis point of view to assess the potential of the evaluated method more differentiated. These categories were as follows:

- isolated lymph nodes (clear borders, no direct neighbor structures)
- lymph nodes with weak gradients at the border (in direct vicinity to soft tissue with similar intensity)

Table 2 Overall comparison of the segmentation results of all 40 lymph nodes by two human experts, the base method from [10] and the here presented method

	Expert 1	Expert 2	Base [10]	New
Tanimoto[a]	0.678	0.684	0.560	0.721
Over seg.	33.3%	38.2%	7.5%	24.5
Under seg.	11.6%	8.0%	40.0%	11.0%
Surf. dist.	0.378	0.356	0.465	0.280
Hausdorff	2.57	2.39	2.74	2.44

The measures are the Tanimoto coefficient, the relative over and under segmentation in percent, the average surface distance in mm, and the Hausdorff distance in mm

[a] The Tanimoto coefficient t is a volumetric similarity measure of two segmentations A and B, which is defined as follows: $t = \frac{V_A \cap V_B}{V_A \cup V_B}$

- lymph nodes with strong gradients at the border (in direct vicinity to bones or contrast agent showing higher intensity)
- deformed lymph nodes (with no rough elliptical shape)

The calculation never took more than 2 s per lymph node (average: 0.52 s). In contrast to the base method from [10], the newly developed technique performs 40% better in average (regarding the average surface distance) and even improves the results of the human experts between 21 and 26% regarding this measure, meaning that it is lying clearly within the range of the inter-personal variance of these human experts. The improved results of the new method are caused by the two techniques presented in "Lymph node model". The sensor plateau filters enhance the results in general, which is shown in Fig. 5. In contrast to that, the improvements on the subset of significantly enlarged lymph nodes are bigger, due to the multi-model segmentation. At last, a chosen border offset parameter value of −0.8 mm balances over and under segmentation and avoids large over or under segmentations, which would lead to worse results.

Table 3 shows the detailed results for the different lymph node categories, and Fig. 11 shows sample segmentation results for each category. As expected, the best results are achieved on isolated lymph nodes, followed by lymph nodes with strong gradients at the border, because of the clear borders in both cases. Lymph nodes with weak gradients at the border or deformed shape lead to worse results, because of the partly useless border or shape information. Interesting is that the new method improves the results in all categories. This is due to several effects:

1. the overall more stable fitting of the model to the lymph node border , due to the use of plateau filtering techniques in all sensor types,

Table 3 Comparison by lymph node category (10 lymph nodes each) of the segmentation results of all 40 lymph nodes by two human experts, the base method from [10] and the here presented method

	Expert 1	Expert2	Base [10]	New
Isolated				
Tanimoto	0.654	0.643	0.512	0.720
Over seg.	29.7%	50.8%	9.8%	16.9%
Under seg.	17.5%	5.2%	44.3%	15.7%
Surf. dist.	0.253	0.273	0.408	0.176
Hausdorff	1.99	2.15	2.25	1.67
Weak border				
Tanimoto	0.648	0.718	0.589	0.705
Over seg.	43.8%	32.5%	11.7%	28.8%
Under seg.	9.7%	7.8%	34.2%	10.0%
Surf. dist.	0.610	0.367	0.544	0.367
Hausdorff	3.67	2.69	3.47	3.15
Strong border				
Tanimoto	0.698	0.666	0.536	0.726
Over seg.	30.2%	42.3%	6.6%	30.5%
Under seg.	10.5%	8.3%	42.4%	7.3%
Surf. dist.	0.281	0.391	0.434	0.240
Hausdorff	2.18	2.22	2.37	2.28
Deformed				
Tanimoto	0.712	0.707	0.599	0.732
Over seg.	29.6%	27.0%	1.8%	21.9%
Under seg.	8.8%	10.6%	38.9%	11.1%
Surf. dist.	0.366	0.394	0.474	0.337
Hausdorff	2.44	2.49	2.88	2.64

The measures are the Tanimoto coefficient, the relative over and under segmentation in percent, the average surface distance in mm, and the Hausdorff distance in mm

2. the improved segmentation of the significantly enlarged lymph nodes, which are contained in each lymph node category,
3. the model's border offset, which leads to an even proportion of over and under segmentation.

The effect of the multi-model segmentation technique was separately quantitatively evaluated using models of different scales (base model with 3 mm diameter and bigger scales of this model) on all 19 existent enlarged lymph nodes (diameter > 10 mm) in the given datasets. Table 4 shows the results in the case of an especially large, necrotic lymph node (diameter: approximately 22 mm) exemplarily. Here, the positive effect of the new technique can be seen very clear, since the base scale model does not work a bit. On the smaller but still enlarged lymph nodes, the results are very similar, but on smaller scales, of course. Here, it can be seen that a high quality of fit corresponds very well with a good segmentation.

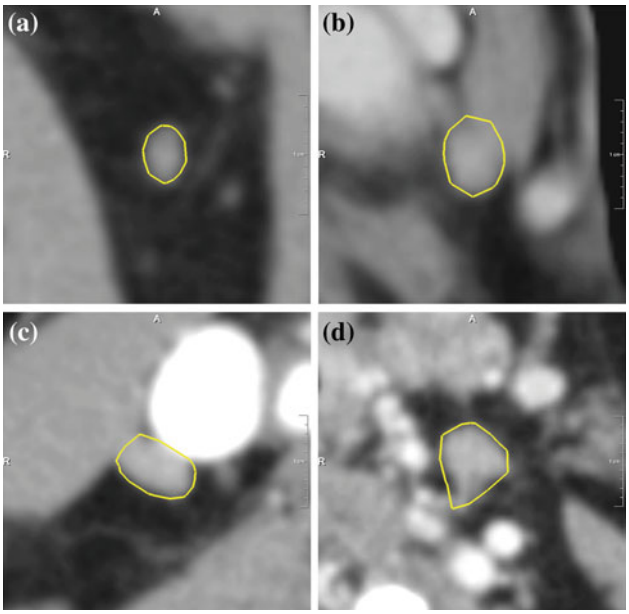

Fig. 11 Sample segmentation results from the different lymph node categories. **a** Isolated lymph node. **b** Lymph node with weak gradients at the border. **c** Lymph node with strong gradients at the border. **d** Deformed lymph node

To get a better impression of this values, Fig. 6 shows some sample results of this test.

Initialization robustness of the lymph node model

On the same 40 lymph nodes from "Single lymph node segmentation", we also tested the robustness of the initialization, that is the dependency on the lymph model's starting position. Thereto, we systematically compared all possible starting positions on a given grid, so that the placed model still touches a part of the target structure. This is necessary, because otherwise the model gets no sensor input from the target structure and cannot perform a successful local adaption. Thereto, the size of this grid was chosen as the bounding box of each lymph node, extended on each side by half the diameter of the basis model, which has a diameter of 3 mm. As the grid resolution, we used 1.2 mm, which is more than twice as fine as the resolution of our search grid (see "Complete automatic lymph node segmentation"). A segmentation from a grid starting point was considered as successful, with a robust starting point, when it had a Tanimoto coefficient of 0.66 or more compared to the reference segmentation from the model that was successfully started at the center of each lymph node. This conforms to the average Tanimoto coefficient of inter-personal manual comparison segmentations. All used parameters can be seen in Table 5.

In Table 6, the results of the initialization robustness tests are shown. As can be seen, a robust starting point was always found outside a circle around each lymph node center with a radius of 40% of the respective lymph node diameter, which is 80% of its radius. This value is noticeably higher than the expected 50% reasoned from the theory in "Requirements", because of the fact that the plateau-filtered inner intensity sensors seem to steer the model during the model adaption even from starting positions outside the lymph node border reliably inside the lymph node, which forms their local attracting plateau maximum. In this context, Fig. 12 illustrates this test in pictures.

Complete automatic lymph node segmentation

We selected the four typical lymph node regions in one of the datasets of the neck used in the previous tests. They contain 29 lymph nodes with a diameter of minimal 8 mm, so that they could be potentially enlarged and clinically relevant for that reason. Furthermore, we had complete gold standards for them. The chosen regions represent the two typical settings for lymph nodes from an image analysis point of view, because regions 1 and 2 have many lymph nodes directly between other soft tissue structures (muscles, vessels, cartilage etc.), and regions 3 and 4 contain lymph nodes in the direct neighborhood of bone structures. The specific information on the four regions can be found in Table 7.

The lymph node model developed in "Lymph node model" (see Fig. 2) was used in the search process, which is suitable for the adequate semi-automatic segmentation (after manual initial positioning) of the lymph nodes, as seen in Table 2. This makes it suitable for the local search of these structures, as stated in Sect. "Requirements". Furthermore, as the lymph node model is an SMSM, a method exists to efficiently

Table 4 Results of the multi-model segmentation with models of different scaling factors (base model: 3 mm diameter) on a large, necrotic lymph node (diameter: ca. 22 mm)

Scaling factor	1	2	3	4	5	6	7	8	9	10
Quality of fit	0.572	0.570	0.572	0.757	0.769	0.834	0.835	0.841	0.791	0.763
Visual acceptable	no	no	no	no	no	yes	yes	yes	no	no
Tanimoto coefficient	0.012	0.012	0.046	0.428	0.445	0.733	0.731	0.760	0.620	0.591
Average surface distance (mm)	8.85	8.80	8.35	3.65	3.53	1.19	1.23	1.03	1.98	1.97

Best results achieve the models of scaling factor 6, 7, and 8, which have an initial diameter of 18, 21, and 24 mm, respectively

Table 5 Parameter values used for the initialization robustness test

Test parameter	Value
Test grid size	lymph node size + 3 mm
Test grid resolution	1.2 mm
Tanimoto coeff. threshold	0.66

Table 6 Maximal successful model initialization distances from the respective centers of the successful reference models for each lymph node and resulting Tanimoto coefficients

	Distance (mm)	Distance fraction of lymph node diameter (%)	Tanimoto coefficient
Minimum	2.40	40.2	0.676
Average	6.69	73.6	0.898

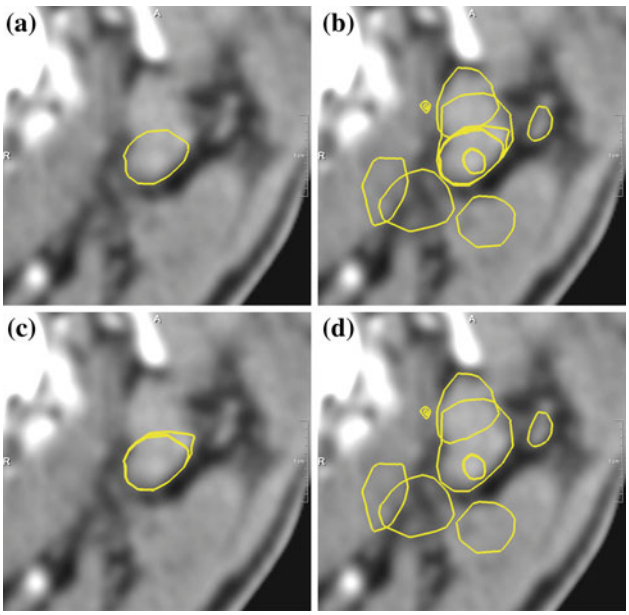

Fig. 12 Initialization robustness tests on a sample lymph node. **a** Correctly adapted reference model. **b** All models finally adapted. **c** All successfully adapted models. **d** All not successfully adapted models

Table 7 Overview of the 4 chosen typical sub-regions from the examined neck CT dataset (rough locations, numbers of contained lymph nodes, and sizes of the sub-regions)

Region	Location description	Lymph nodes	Size in voxels		
			X	*Y*	*Z*
1	V. jugularis right	7	80	113	233
2	V. jugularis left	12	110	137	222
3	Mandible right	4	142	110	76
4	Mandible left	6	106	110	80

Table 8 Parameter values used for the search process

Search parameter	Value
Search grid resolution	3.5 mm
Quality of fit (Init.)	0.6
Quality of fit (End)	0.8
Maximum step count	250
Minimal scaling	2.0
Overlap coefficient	0.7

calculate its quality of fit (see [25]), which is another requirement according to Sect."Requirements". With a diameter of 3 mm, the lymph node model was significantly smaller than its target structures in order to be able to initialize it completely inside the target structure, which is necessary for a reliable segmentation, before it adapts its size by its dynamic adaption to a target lymph node.

For the search process described in "The search process", the density of the starting grid for the initial model population has to be determined. It was according to "Requirements" chosen slightly below half the expected minimal object diameter of 8 mm of potential clinically relevant lymph nodes. For the lymph node detection, we therefore chose a grid distance of 3.5 mm. This grid size can easily be chosen like that, as in the lymph node application case we are dealing with a compact model (the lymph node model includes the object's interior), so that no different internal local minima exist for the objective function of the model. This value derived from the theory in "Requirements" has furthermore been confirmed by practical examinations about the robustness w.r.t. the initial placement of the model. "Initialization robustness of the lymph node model" shows that the segmentation accuracy of the lymph node models only decreases at an initial placement, which is more than half the object's radius away from the center of the lymph node. All other parameters could be determined easily by experiments, as the search is robust w.r.t. to them, so that they could be chosen with a certain tolerance. These parameters are listed in Table 8.

The expectation map was a combination (Boolean AND) of two maps. One map represented the anatomical regions of the neck lymph node groups, and the other map represented the expected gray value range of the lymph nodes.

The results from the search process are listed in Table 9. Example results are depicted in Figs. 13 and 14. All 29 lymph nodes were found in 2 min in all of the tested regions, so there exist no false negatives. Nine false positives (ca. 31% of the correct lymph node count) were detected, which must be considered as very few, compared to the manual expert search (compare [4]) and other automated approaches [20].

The false-positive lymph node candidates (in most cases slightly contrasted blood vessels) are similar to lymph nodes

Table 9 Results of the lymph node detection

Sub-region	Models total	Models started	Models successf.	Detected cand. (detect. rate)	False pos.	Runtime in sec
Region 1	3,920	294	7	7 (100%)	0	28
Region 2	5,967	346	18	12 (100%)	6	50
Region 3	1,989	184	7	4 (100%)	3	25
Region 4	1,690	134	6	6 (100%)	0	17
Sum	13,566	958	38	29 (100%)	9	120

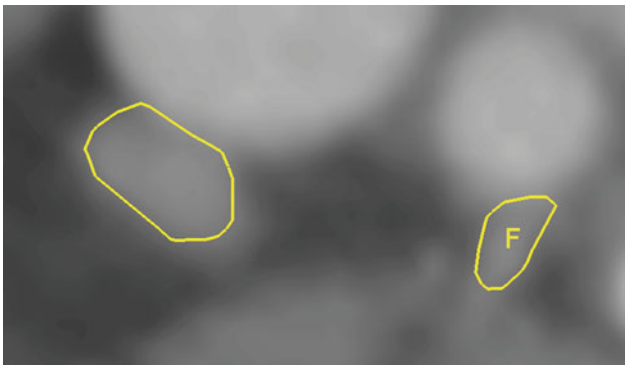

Fig. 13 Two detected target structures for the lymph node detection, the false positive is marked ("F")

with respect to the modeled shape and gray value properties but differ from the latter by their context (see Fig. 13). Additional information about the location of the blood vessels, as it is often available, would allow to exclude these false-positive results.

Table 9 also shows the calculation times of the search process in the range of 2 min for all relevant lymph node regions in a dataset, which is a necessary condition for a practical, or even clinical applicability of the method. Besides the use of 4 processors for the straightforward parallel calculation, one reason for the low calculation times is the high percentage of individuals that were rejected early, which is shown in detail in Table 10.

Limitations

Beside the positive results, there exist some limitations of the presented method. At the moment, 3D data from a CT are required, because of the comparability of the Hounsfield values. If other image modalities (e.g., MRI) would be calibrated in a defined way, the application of the method on them would also be thinkable.

In our tests, the result quality decreases significantly, if the X-Y resolution falls below 1 mm and the Z resolution falls below 3 mm. However, the quality of the results does not seem to depend on the manufacturer of the scanner or the use of contrast agent.

Furthermore, it was noted that extremely degenerated pathological lymph nodes (frayed shape, more than 5 cm diameter, necroses at the border, etc.) differ so much from

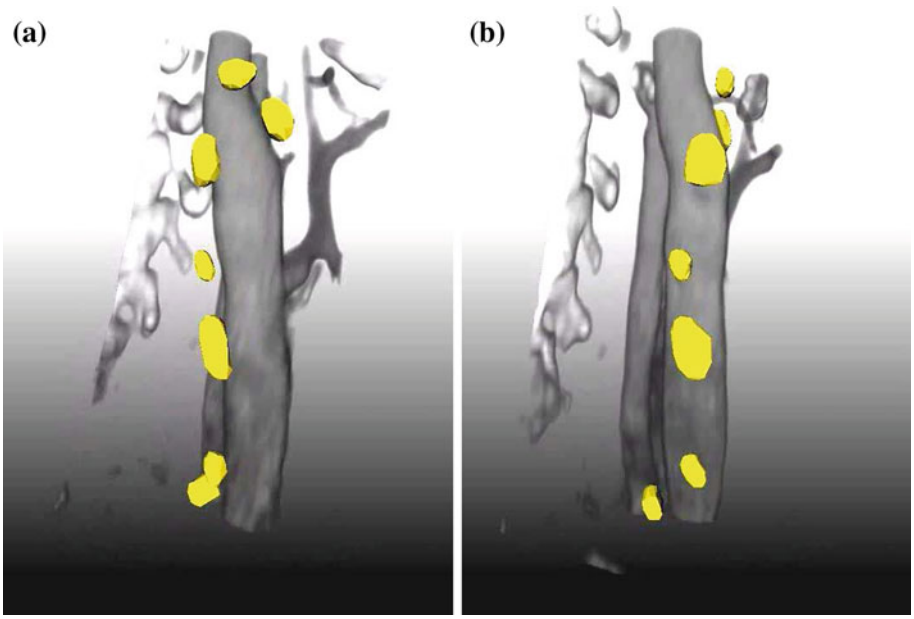

Fig. 14 All lymph nodes were correctly detected in region 1, which are shown here in context to the V. jugularis

Table 10 Details of the model reduction during the lymph node detection, distinguished by reduction criteria and process phase

Sub-region	Exp. map (Init.)	Qual. of fit (Init.)	Exp. map (Sim.)	Qual. of fit (Sim.)	Scaling (End)	Grouping (End)
Region 1	3,388	238	26	255	11	8
Region 2	5,151	470	22	279	11	23
Region 3	1,507	298	4	141	22	13
Region 4	1,348	208	6	103	8	14
Sum	11,394	1,214	58	778	52	58

Double countings were possible

the shape and appearance knowledge coded in the used lymph node model because of their individuality that no reliable automatic segmentation is possible. Here, methods for the segmentation of individual tumors address these special structures better. But for the vast majority of possible lymph node characteristics, the present method performs very reliable, as the tests show.

Conclusion

The presented partial technique for the segmentation of single lymph nodes in CT datasets is the first to handle lymph nodes of different classes (normal, enlarged, necrotic, fuzzy-bordered, deformed). Furthermore, it is very robust regarding its initialization. An evaluation on a set of 40 systematically chosen lymph nodes of 4 different image analysis categories from 5 very different neck CT datasets showed that the quality of the results improves the former techniques significantly and lies clearly in the range of the inter-personal variance of human experts. The running times of about half a second of the developed method for the single lymph node segmentation qualifies it for use within other complex methods. Anyway, complicated deformations or large necroses cannot be segmented satisfying in any case, because too much dataset information (artifacts, poor dataset quality, etc.) is missing to be completed by the model knowledge.

The introduced plateau filtering technique could also be interesting for other sensor-based segmentation models, because general problems are addressed with them (object internal gray value range and useless, distracting edge information). Also interesting in this way could be the principle of the multi-model segmentation, whose effect was theoretically motivated and practically clearly confirmed. Here, a transfer to other segmentation problems with different object classes is also considerable.

The presented novel completely automatic segmentation method for structures in medical volume data based on a population simulation of dynamic shape models (SMSMs) consists of a global search, which is achieved by a distributed, local search of the individual model instances with a quality of fit estimation. The lymph node model (an SMSM) developed in the first part of this paper has proved to be especially suited for this purpose.

All 29 lymph nodes contained in 4 selected typical lymph node regions of a neck CT dataset were detected correctly. The 31% false-positive results can be attributed to the (also for a human observer) difficult distinguishability of the target structures from other medical structures. In the examined cases, the presented method performed within 2 min, which complies with the time requirements for an application in the clinical setting of our clinical partners, since it can be run as a background task, while other work is done by the user.

Beyond the presented application case, our object detection approach is a general and well-manageable possibility to search for compact structures in medical volume data. The requirements developed in this work allow for a judgment on the suitability of the presented detection technique for further application cases.

References

1. Sobin LH, Wittekind C (2002) TNM: classification of malignant tumours. Wiley-Liss
2. Hintze J, Cordes J, Preim B (2005) Bildanalyse für die präoperative Planung von Neck Dissections. In: BVM
3. Cordes J, Dornheim J, Preim B, Hertel I, Strauss G (2006) Preoperative segmentation of neck ct datasets for the planning of neck dissections. In: SPIE: Medical Imaging
4. Schreyer A et al (2005) Vergleich der Genauigkeit und Effizienz koronarer und axialer Rekonstruktionen zur Lymphknotendetektion bei einem 16-Zeilen-Spiral-CT am Beispiel von Halslymphknoten. RöFo: Fortschritte auf dem Gebiete der Röntgenstrahlen und der Nuklearmedizin 177:1430–1435
5. Dornheim L, Tönnies KD, Dixon K (2005) Automatic segmentation of the left ventricle in 3D SPECT data by registration with a dynamic anatomic model. In: MICCAI
6. Dornheim L, Tönnies KD, Dornheim J (2005) Stable dynamic 3D shape models. In: ICIP
7. Rogowska J et al (1996) Evaluation of selected two-dimensional segmentation techniques for computed tomography quantification of lymph nodes. Invest Radiol 13:138–145
8. Yan J et al (2004) Lymph node segmentation from CT images using fast marching method. Comput Med Imaging Graph 28: 33–38
9. Honea D, Snyder WE (1999) Three-dimensional active surface approach to lymph node segmentation. In: SPIE: Medical Imaging 3361:1003–1011

10. Dornheim J et al (2007) Segmentation of neck lymph nodes in ct datasets with stable 3d mass-spring models. Academic Radiology (Elsevier) MICCAI 2006 Special Issue
11. Maleike D et al (2008) Lymph node segmentation on ct images by a shape model guided deformable surface method. In: SPIE: Medical Imaging
12. Felzenszwalb PF, Huttenlocher DP (2005) Pictorial structures for object recognition. Int J Comput Vis 61:55–79
13. Carbonetto P, de Freitas N, Barnard K (2004) A statistical model for general contextual object recognition. In: ECCV
14. Ferrari V, Tuytelaars T, Gool LV (2004) Simultaneous object recognition and segmentation by image exploration. In: ECCV
15. Belongie S, Malik J, Puzicha J (2002) Shape matching and object recognition using shape contexts. IEEE Trans Pattern Anal Mach Intell 24:509–522
16. Coughlan J, Yuille A, English C, Snow D (2000) Efficient deformable template detection and localization without user initialization. Comput Vis Image Underst 78:303–319
17. Jain AK, Zhong Y, Lakshmanan S (1996) Object matching using deformable templates. IEEE Trans Pattern Anal Mach Intell 18:267–278
18. Keren D, Subrahmonia J, Cooper DB (1992) Robust object recognition based on implicit algebraic curves and surfaces. In: ICPR 791–794
19. Ballard DH (1981) Generalizing the hough transform to detect arbitrary shapes. Pattern Recognit 13:111–122
20. Kitasaka T, Tsujimura Y, Nakamura Y, Mori K, Suenaga Y, Ito M, Nawano S (2007) Automated extraction of lymph nodes from 3-d abdominal ct images using 3-d minimum directional difference filter. In: MICCAI
21. Cootes TF, Taylor CJ, Cooper DH, Graham J (1995) Active shape models—their training and application. CVIU 61:38–59
22. Delingette H (1994) Simplex meshes: a general representation for 3D shape reconstruction. In: CVPR 856–857
23. Bardinet E, Cohen LD, Ayache N (1998) A parametric deformable model to fit unstructured 3D data. CVIU 71:39–54
24. Dornheim L, Dornheim J, Tönnies KD (2006) Automatic generation of dynamic 3d models for medical segmentation tasks. In: SPIE: Medical Imaging
25. Dornheim L (2008) Quality of fit of stable mass-spring models. Pattern Recognit Image Anal 18:359–364
26. Bergner S, Al-Zubi S, Tönnies K (2004) Deformable structural models. In: ICIP
27. Cohen LD (1991) On active contour models and ballons. CVGIP: Image Underst 53:211–218
28. Dornheim L, Dornheim J, Seim H, Tönnies KD (2006) Aktive Sensoren: Kontextbasierte Filterung von Merkmalen zur modellbasierten Segmentierung. In: Bildverarbeitung für die Medizin