Assessing breathing motion by shape matching of lung and diaphragm surfaces*

Martin Urschler, Horst Bischof

Institute for Computer Graphics and Vision, Graz University of Technology,
Inffeldgasse 16/2, A-8010 Graz, Austria
{urschler,bischof}@icg.tu-graz.ac.at

ABSTRACT

Studying complex thorax breating motion is an important research topic for accurate fusion of functional and anatomical data, radiotherapy planning or reduction of breathing motion artifacts. We investigate segmented CT lung, airway and diaphragm surfaces at several different breathing states between Functional Residual and Total Lung Capacity. In general, it is hard to robustly derive corresponding shape features like curvature maxima from lung and diaphragm surfaces since diaphragm and rib cage muscles tend to deform the elastic lung tissue such that e.g. ridges might disappear. A novel registration method based on the shape context approach for shape matching is presented where we extend shape context to 3D surfaces. The shape context approach was reported as a promising method for matching 2D shapes without relying on extracted shape features. We use the point correspondences for a non-rigid thin-plate-spline registration to get deformation fields that describe the movement of lung and diaphragm. Our validation consists of experiments on phantom and real sheep thorax data sets. Phantom experiments make use of shapes that are manipulated with known transformations that simulate breathing behaviour. Real thorax data experiments use a data set showing lungs and diaphragm at 5 distinct breathing states, where we compare subsets of the data sets and qualitatively and quantitatively asses the registration performance by using manually identified corresponding landmarks.

Keywords: Breathing motion, 3D shape matching, 3D volume registration

1. INTRODUCTION

According to the European Respiratory Society, lung diseases rank second behind cardiac diseases in terms of mortality and cost of treatment†. Computerized methods for objective, accurate and reproducible analysis of lung structure and function can provide important insights into these problems. However, due to the complexity of the breathing motion, investigations and applications working with thoracic but also abdominal images are often very complicated. There are several areas of application that would benefit from a proper treatment of problems due to breathing motion. Accurate fusion of functional (i.e. lung perfusion, lung ventilation) and anatomical data sets[1] requires knowledge about the data sets' positions in the breathing cycle and about the possible soft-tissue deformations that are induced by diaphragm and rib cage movement. Problems with breathing motion artifacts that occur due to time-consuming image acquisition protocols (e.g. SPECT imaging) demand methods to reduce these artifacts in a post-processing step for example by using a pre-defined model of breathing motion. In radiotherapy planning, which is an essential technique for treatment and cure of tumors confined in closed regions, breathing motion leads to problems with patient positioning and target volume determination.[2] Breathing motion estimation and compensation tools are therefore invaluable for the planning step. Registration of pre- and intra-operative thoracic and abdominal images is nowadays being used more often by image-guided intervention and surgery systems. An accurate registration of thoracic and abdominal soft-tissue regions has to take breathing motion into account. An example for such a registration system used for liver interventions is described in Blackall et al.[3,4]

*We gratefully acknowledge the support of Prof. Eric Hoffman, Department of Physiologic Imaging, University of Iowa, Iowa City, IA 52242 for providing the CT image data.

†http://www.ersnet.org

The work presented in this paper is part of a long-term research effort to investigate breathing motion in thoracic and abdominal tomographic images given intermodal CT data sets at several different breathing states between Total Lung Capacity (TLC) and Functional Residual Capacity (FRC). In this context, non-linear registration of soft tissue organs (lung, liver) over the breathing states will be used to derive motion models from the data. Statistical models will be established, optionally using biomechanical constraints, for tasks like e.g. breathing artifact reduction or breathing motion compensation in image fusion applications.

This paper specifically focuses on the non-linear registration task that has to be performed before one can build statistical models. Basically the body of literature distinguishes between intensity- and feature based non-linear registration methods. A survey on non-linear registration methods in medical image analysis can be found in Maintz et al.[5] Audette et al.[6] give an algorithmic overview of surface registration techniques for medical imaging. Zitova et al.[7] recently published an overview of image registration techniques. The work in this paper concentrates on non-linear registration based on a shape descriptor, more specifically finding correspondences via shape matching[8] and registering identified correspondences in the context of a thin-plate spline framework.[9] In many cases shape-based methods provide a more accurate solution since only the organ of interest is considered as opposed to intensity-based methods. On the other hand, segmentation of the organs of interest is not always an easy task and inaccuracies in the segmentation process have effects on the subsequent registration step.

Feature-based registration starts with a feature extraction step. After analyzing lung and diaphragm surfaces over the breathing cycle one will notice that the shapes undergo certain deformations which make it hard to robustly identify correspondences. Especially the deformations of the surface of the elastic lung tissue are hard to describe. The lung surface is attached to the rib cage by adhesive forces and the diaphragm-induced cranio-caudal forces move the lung tissue while at the same time the bones of the rib cage remain comparatively rigid. As a consequence the rib cage forces its shape on the moving elastic lung tissue and it is not necessarily the case that classical 3D features like ridges (or valleys) on the lung surface stay ridges (or valleys) during breathing. For this reason a shape matching algorithm has to be developed that is independent of points of interest like curvature maxima or similar 3D features. After reviewing the body of literature the shape context approach reported by Belongie et al.[10] was identified as a reasonable and promising approach for this purpose. This algorithm was extended to 3D and adapted for the task at hand. The idea of extending the shape context descriptor to 3D has been independently proposed by Frome et al.[11] for recognizing 3D objects in range scans. Given the correspondences from the shape matching an estimate of the breathing motion in form of a displacement field can be calculated in a non-linear registration step. Belongie et al. propose the thin-plate spline framework for this task. Davis et al.[12] argued that in the context of medical applications the elastic body spline is a more accurate model for displacement field calculation.

This paper is organized as follows. Section 2 briefly describes the 3D image data used for this research. In Section 3 the basic techniques and their extensions which were used to identify shape correspondences and non-linearly register the image data are presented. Section 4 introduces setup and results of the synthetic and real data experiments, while Section 5 discusses the obtained results and gives a conclusion.

2. 3D IMAGE DATA

Image data for this study comes from high-speed multi-detector spiral CT sheep scans. The sheep CT data was provided by Prof. Eric Hoffman, University of Iowa, IA. The data was acquired at five different breathing states between TLC and FRC by a protocol where breath is held at fixed inspiration levels during the 30 sec scan time. This leads to a static breathing scheme, which has to be considered for the interpretation of derived motion models from matched and registered shapes. A CT protocol to scan thorax anatomy at different breathing states with high spatial resolution during *dynamic* (normal) breathing is not yet used routinely. The image dimensions per breathing state are approximately 512x512x550 with voxel dimensions of 0.52mm x 0.52mm x 0.6mm. Fig. 1 shows some typical slices of the 3D image data. In this work a single data set composed of five different breathing states was used.

3. METHODS

In the literature there are several related methods for registration and warping of lung surfaces. Betke et al.[13] reported an automatic 3D registration technique for lung surfaces in CT scans. They use static anatomical lung

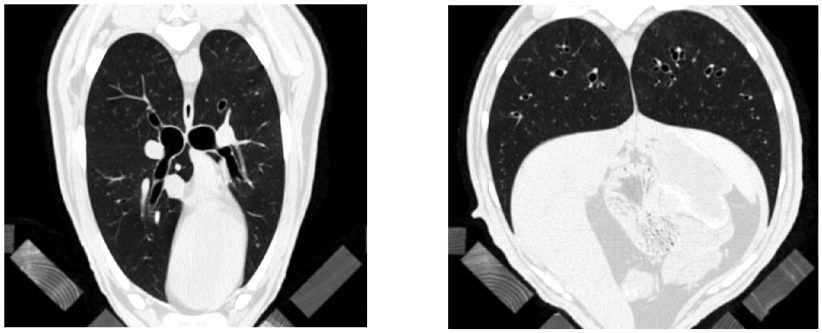

Figure 1. Typical slices of the 3D image data.

landmarks to obtain an initial registration and refine it with an iterative surface-to-surface registration method. Both steps solely use an affine transformation model to describe the deformation. Fan et al.[14] developed a technique for motion field derivation from lung images by using an intensity-based 3D optical flow scheme following a coarse pre-registration from corresponding airway tree branching points. Li et al.[15] proposed an inter-subject warping and registration scheme for lung CT images using a set of reproducibly extractable corresponding feature points and a landmark and intensity-based consistent image registration algorithm. Tschirren et al.[16] presented a registration scheme based on a branchpoint matching of segmented airway trees.

In this work the shape context approach was used to find correspondences on lung and diaphragm surfaces followed by a non-linear thin-plate spline registration. This is a more general approach than the ones stated above, since basically all kinds of extracted surfaces (e.g. segmented airway or vessel tree surfaces as well as skeletonized trees) can be used in the same matching framework. Related work to the shape context approach can be found in Sclaroff et al.[17] who perform a modal matching algorithm based on the idea of describing objects in terms of generalized symmetries. Another approach is the TPS-RPM (Thin-Plate Spline - Robust Point Matching) method developed by Chui et al.[18] who estimate the parameters of the correspondence problem and the non-linear registration transformation simultaneously in a deterministic annealing driven expectation-maximization scheme.

3.1. Shape Context Matching and Non-Rigid Registration

The shape context approach[10] was reported as a reasonable and promising method for matching 2D shapes (especially hand-written digits and letters) and 2D object recognition without relying on extracted features. It combines global (by regarding all points of a shape) and local (by storing information about the relation of all possible point pairs) shape knowledge in a clever way. Objects are treated as (possibly infinite) point sets and it is assumed that the shape of an object is captured by a finite subset of its points, giving a set $P = \{p_1, ..., p_n\}$. The points can be obtained as locations of edges from an edge detector or from another method to sample contour/surface points from a shape. The points need not and typically will not correspond to key points or structures such as maxima of curvature, inflection points or surface ridges. If one looks at the set of vectors emitted from one point p_k to all other points p_i of a shape with $i \neq k$, this set can be interpreted as a rich description of the shape configuration relative to p_i. Since this description is much too detailed, the *relative distribution* of this set of vectors is taken as a compact, yet highly discriminative descriptor instead. Therefore, for each point p_i a histogram h_i of the relative positions of the remaining points is calculated which is called the *shape context*. This histogram uses bins that are uniform in a three-dimensional spherical coordinate system (θ,ϕ,r). The r coordinate axis is logarithmically sampled, such that positions of nearby sample points have stronger influence on the descriptor than ones located farther away. Fig. 2a shows the bin structure of 3D shape context histograms.
Now for each point p_i on the first shape, the "best" matching point q_j on the second shape has to be located.

For a point tuple $< p_i, q_j >$, let

$$C_{ij} = C(p_i, q_j) = \frac{1}{2} \sum_{k=1}^{K} \frac{[h_i(k) - h_j(k)]^2}{h_i(k) + h_j(k)}$$

denote the cost of matching these two points. This cost function is simply based on the χ^2 test statistic which is a natural choice for comparing histograms. Given the set of costs C_{ij} between *all* pairs of points on first and second shape, one wants to relate each point from the first shape with *exactly* one point from the second shape under the constraint that the total cost of this mapping is minimized. This problem can be formulated in a graph framework by taking the sample points from both shapes as graph nodes in a bipartite graph, i.e. there are only graph edges *between* the two sets of sample points, no edges between sample points of a single set. Edges in the graph are weighted with the cost function C_{ij}. This graph setup is illustrated in Fig. 2b. Finding a cost-minimizing mapping between point sets now transforms to a graph matching problem which is an instance of the weighted bipartite assignment problem. It can be solved in polynomial time, e.g. with the Hungarian algorithm in $O(N^3)$ time, with N being the number of nodes in the bipartite graph. The final result of the graph matching step is a one-to-one mapping of corresponding points from the two shapes.

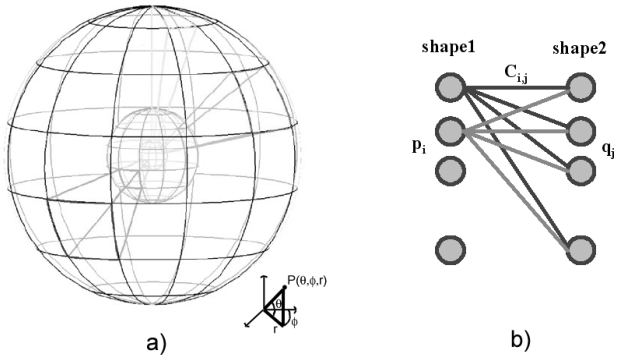

Figure 2. Components from the 3D shape context descriptor calculation. a) shows the histogram bin structure embedded in a spherical coordinate system (θ,ϕ,r). b) depicts the bipartite graph structure used for the graph matching algorithm, with nodes p_i and q_j representing all sample points from first and second shape, respectively, and edges weighted with $C_{i,j}$ representing the cost function to match a pair of points.

Belongie et al. give some arguments about invariance and robustness of the shape context approach. It is implicitly invariant to translation, since all measurements are taken with respect to object points. Scale invariance can easily be achieved by normalizing the radial distances by the mean distance. If this is desirable for an application, rotation invariance can also be achieved by incorporating local coordinate systems based on the tangent vector at each point instead of using a common global coordinate system. Rotation invariance is not an issue for this work, since data sets are either already in a similar position and orientation, or will be brought into similar orientation by a pre-processing step. Outlier handling can be introduced by using "dummy" nodes attached to each point set, whose cost is chosen in a way that outliers have larger cost values.

After establishing the point correspondences Belongie et al. make use of the thin-plate spline framework[9] for non-linear registration. The thin-plate spline approach leads to a transformation that consists of an affine part and a non-linear deformation part depending on the identified correspondences. The parameters of the thin-plate spline model are calculated from the constraints that the displacement at corresponding points is zero (exact interpolation of the displacement field) and that the spline model between corresponding points yields regular and smooth displacements tending to zero the further away from the corresponding points they are.

3.2. The Shape Registration Pipeline

A shape registration pipeline was implemented able to derive a displacement field given two image data sets at two distinct states of the breathing cycle. Some parts of the pipeline make use of the popular ITK[‡] (Insight Segmentation and Registration Toolkit) software library. A coarse overview of this pipeline is given in Fig. 3.

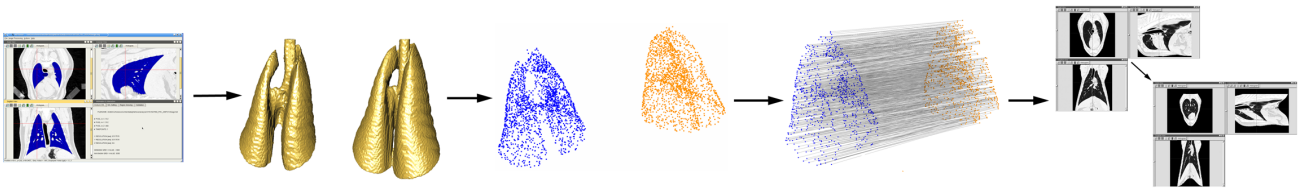

Figure 3. Shape Registration Pipeline. From left to right, volume segmentation, surface extraction, surface point sampling, shape matching to get point correspondences and non-linear registration and volume warping based on the point correspondences are depicted.

The first step in the registration pipeline is an automatic segmentation of the lung surface. A simple algorithm that incorporates a region grower for background removal, a grey value thresholding step and a connected component labeling gives a rough segmentation of the airway tree and the lung tissue. An airway tree segmentation algorithm that was developed in our group[19] is used to get an accurate airway tree with the rough segmentation as a constraint. The lung segmentation is separated into left and right lobe by a labeling step based on the main branching point of the airway tree, using the same method presented in Beichel et al.[20] for liver partitioning. This results in a binary volume for left and right lung lobe. From these binary volumes a diaphragm point set is extracted by regularily sampling points from the bottom part of the lung surface. It should be noted that this simple diaphragm segmentation scheme is working well for sheep CT data but can not so easily be used for human diaphragm segmentation, since the human heart sits on top of the diaphragm, while in sheep data the heart is surrounded by lung tissue.

The next step is the extraction of a point set from the segmented lung surfaces. In contrast to the original shape context paper, discretization is an important issue in this work since 3D discretizations easily produce very large numbers of points. Further, some methods might produce more points in areas of high curvature which is a disadvantage for calculating the shape context descriptor. In this case the descriptor would lose its ability to describe the small lung deformations due to breathing. To avoid these difficulties the binary volumes are triangulated using the Marching Cubes algorithm. This gives a dense and regular triangle mesh, which is sampled in a *regular* way. In all experiments the number of sampling points per lung surface was varied between 200 and 3400 points.

The calculation of the shape context descriptor was extended to incorporate local grey value information. At each extracted surface point a second term of the cost function based on the normalized cross correlation is calculated. The extended cost function C'_{ij} has the following form

$$C'_{ij} = C'(p_i, q_j) = \alpha C(p_i, q_j) + \beta NCC(p_i, q_j)$$

with $C(p_i, q_j)$ being the original shape context descriptor cost function and $NCC(p_i, q_j)$ being the normalized cross correlation evaluated at 5x5x5 neighborhoods of the original grey value CT image centered around p_i and q_j respectively. This additional term adds robustness against segmentation errors and makes use of the otherwise neglected intensity information in the data sets.

Since the 3D extension of the shape context approach requires a larger number of sample points than the simpler 2D case, it is crucial to utilize an efficient algorithm for the bipartite graph matching problem. In this work the minimum weight assignment algorithm from the LEDA software library[§] was used, which solves the

[‡]http://www.itk.org
[§]http://www.algorithmic-solutions.com

graph matching in $O(N(M+NlogN))$ time, with N being the number of nodes and M the number of edges in the bipartite graph. This algorithm proved itself as reasonably fast, since all practical experiments showed that the calculation of the cost functions at the N^2 points took more time than the graph matching step. After finding the one-to-one correspondences a certain percentage of matched points were removed. More specifically matched edges get sorted by their weights and the ones with largest weights get excluded. This adds some robustness against errors introduced in the preceding segmentation and point sampling steps. The subsequent registration step does not need the full number of correspondences since it interpolates the displacements between missing correspondences according to the chosen interpolation model. However, the percentage of correspondences to remove has to be chosen small enough to prevent entire loss of information in certain image regions.

The final step in the registration pipeline is the non-linear registration of the matched points in a thin-plate spline framework. A thin-plate spline interpolation function $f(x, y, z)$ has the form

$$f(x,y,z) = a_1 + a_x x + a_y y + a_z z + \sum_{i=1}^{n} w_i U(\parallel (x_i, y_i, z_i) - (x, y, z) \parallel)$$

with kernel $U(r) = r$. Davis et al.[12] have reported that the elastic body spline (EBS), which solely differs from the thin-plate spline in the kernel function $U(r)$, is a more accurate model for medical imaging applications. They state that this kernel is actually better suited for modeling tissue deformations, since it is derived from Navier's partial differential equations that model the equilibrium of an elastic body subjected to forces. It incorporates a parameter that models the tissue elasticity derived from Poisson's ratio. In this implementation both types of spline kernels were used since the non-linear registration framework is the same in both cases. Another aspect in the registration framework is the large number of sample points that have to be extracted for a reasonable shape approximation. The thin-plate spline framework resembles an interpolation of the displacement field between corresponding points, therefore a large number of correspondences tend to introduce overfitting to still fulfill the interpolation requirements. The group around Belongie published a paper[21] dealing with this problem. In this work the findings of Rohr et al.[22] to approximate thin-plate spline mappings were considered. Rohr et al. proposed to add a regularization term to the formulation, which is steered by a parameter λ weighting the tradeoff between interpolation and smoothness of the solution. λ ranges between 0 for exact interpolation to 0.1 for an approximated purely affine transformation with hardly any local deformations.

4. EXPERIMENTS AND RESULTS

To assess the validity of the shape context matching and registration approach qualitative and quantitative evaluations were performed on synthetically transformed and real thorax data sets. In this work three different kinds of shapes were produced in the segmentation steps and taken as input for the validations, i.e. diaphragm, lung lobe and airway tree surfaces. Fig. 4 shows examples for each kind of shape.

4.1. Matching Experiments on Synthetically Transformed Data

The basic procedure for the synthetic matching experiments is to provide a data set $A(x, y, z)$ and a transformed version of the data set $B(x, y, z)$ as input to the shape matching algorithm. The synthetic transformation $T : \{B(x, y, z) = T(A(x, y, z)), 1 \leq x \leq n_x, 1 \leq y \leq n_y, 1 \leq z \leq n_z\}$ is defined by the evaluator and unknown to the matching algorithm, therefore B serves as a "gold standard" data set for the shape matching algorithm. The algorithm computes from its inputs $A(x, y, z)$ and $B(x, y, z)$ a list of point correspondences $< p_i, q'_i >$, mapping a point set $p_i : \{p_i(x, y, z) \in A(x, y, z), 1 \leq i \leq n\}$ to a point set $q'_i : \{q'_i(x, y, z) \in B(x, y, z), 1 \leq i \leq n\}$ with n being the number of surface sample points. By applying the transformation T on the point set p_i a set of synthetically transformed points $q_i : \{q_i(x, y, z) = T(p_i(x, y, z)), 1 \leq i \leq n\}$ is calculated. If the shape matching algorithm would be perfect q_i and q'_i would be identical for all values of i. Since in practice this is not the case, for each tuple of point sets q'_i and q_i minimum, mean and maximum distances are calculated and interpreted as an accuracy measure for the shape matching algorithm.

Some of the parameters involved in the matching and registration algorithm remain fixed during the experiments. The weighting parameters of shape context and normalized cross correlation cost function are set to $\alpha = 0.7$ and $\beta = 0.3$ to give more importance to the shape context cost function term. The percentage of point

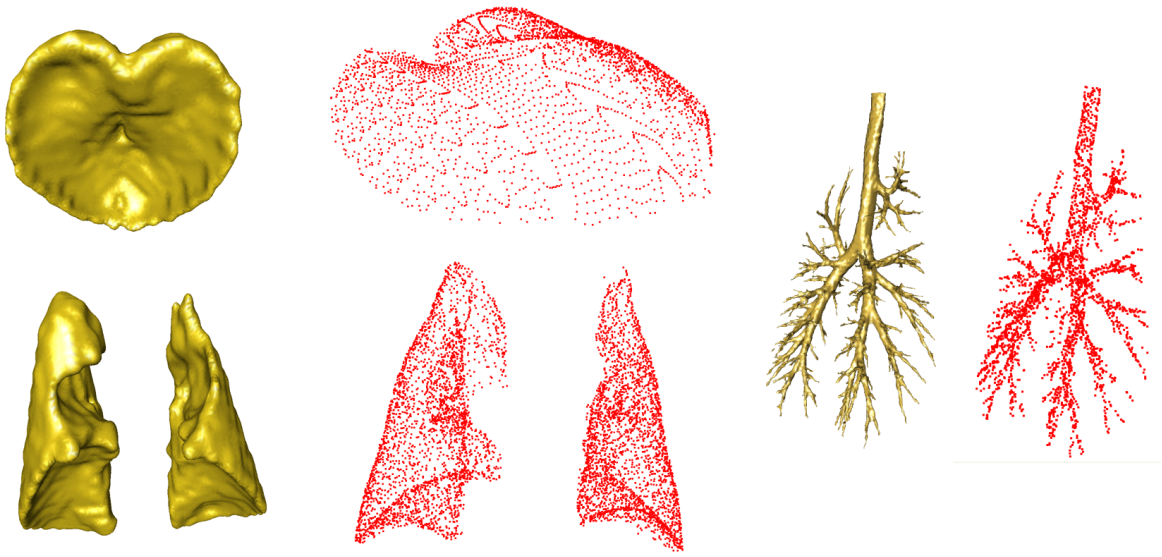

Figure 4. The different kinds of data used in the experiments. Segmented surfaces to the left and sampled points to the right, respectively. Right column shows airway tree data, while left column shows diaphragm at the top and lung lobes at the bottom.

correspondences removed from the shape matching results is set to 20% and it was ensured that there are no regions lacking sample points after removal by visual inspection. The elastic body spline kernel is not used for the final validations. Compared to the thin-plate spline kernel its need to tune the parameter resembling Poisson's ratio and its higher computation time were not worth the only slightly better results. The regularization factor for the thin-plate spline interpolation is set to 0.01.

The shape context approach is used to find shape correspondences of thorax structures *over breathing*. Therefore a synthetic transformation has to simulate breathing behaviour. A non-linear transformation $T : \{B(x, y, z) = T(A(x, y, z))\}$ was designed to approximate a breathing-like deformation by simulating diaphragm and rib cage movement. According to the literature[23] diaphragm movement for deep breathing ranges between 25 to 40 mm in the vertical direction during a breathing cycle. This movement is simulated by applying a translational force to the diaphragm surface in the data sets negative z direction. A non-linear force is calculated by weighting the constant translation $t_{vertical}$ with a two-dimensional Gaussian distribution that depends on the $x-$ and $y-$ coordinates of the data set, i.e. the further away from the center of the diaphragm surface a point is, the smaller is the negative z translation. Mathematically a displacement vector $\vec{d_1} = (0, 0, z')$ is applied to each point $(x, y, z)^T$ that maps it to $(x, y, z')^T$ with

$$z' = z - t_{vertical} e^{-\frac{(x-\mu_x)^2+(y-\mu_y)^2}{2\sigma^2}}$$

where (μ_x, μ_y) corresponds to the $x-$ and $y-$ coordinates of the center of gravity of the diaphragm points and σ is chosen such that points lying at the exterior of the diaphragm surface (where the diaphragm is attached to the rib cage) nearly remain fixed.

In a similar fashion, simulation of rib cage behaviour during breathing leads to the development of another displacement force component. A radial, center-directed translation t_{inward} is used to form a second displacement $\vec{d_2} = (x', y', 0)$ that maps points $(x, y, z)^T$ to $(x', y', z)^T$ with

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} \mu_x \\ \mu_y \end{pmatrix} + t' * \frac{\vec{c}}{|\vec{c}|}$$

where

$$\vec{c} = \begin{pmatrix} x - \mu_x \\ y - \mu_y \end{pmatrix} \qquad and \qquad t' = |\vec{c}| - t_{inward} * (1 - e^{-\frac{(x-\mu_x)^2+(y-\mu_y)^2}{2\sigma^2}})$$

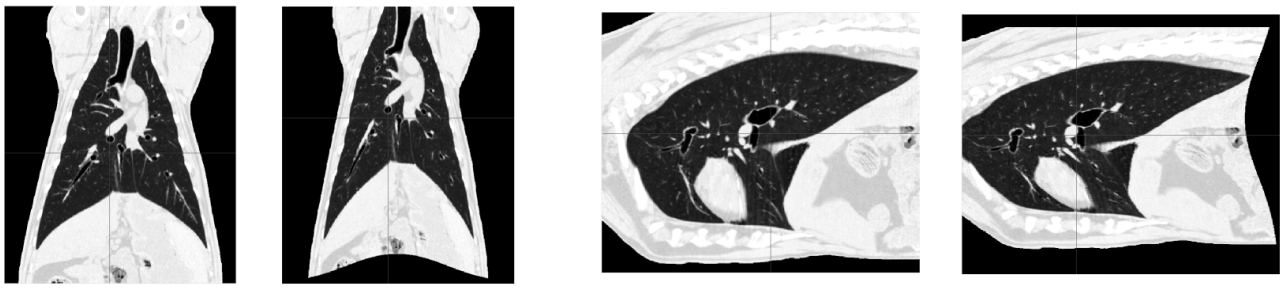

Figure 5. Sagittal and coronal views of a data set demonstrating the synthetic transformation that is used to simulate breathing behaviour. Left column shows original, right column transformed data, respectively.

Combining displacements $\vec{d_1}$ and $\vec{d_2}$ gives a total displacement $\vec{d}$ that is equivalent to a non-linear transformation $T : \{B(x, y, z) = A(x, y, z) + d(x, y, z)\}$. An example for this transformation using $t_{vertical}$ of 25 mm and t_{inward} of 10 mm on a thorax data set is shown in Fig. 5.

The first synthetic experiment operates on a segmented diaphragm. Diaphragm surface points are extracted from a segmented lung surface data set at TLC. The vertical translation $t_{vertical}$ and the inward translation t_{inward} remain fixed to 25 and 10 mm, respectively. The number of sample points used to represent the diaphragm surface are varied between 400 and 3600 sample points in steps of 400. To assess the validity of the shape matching method based on the shape context descriptor (without dependencies on segmentation and shape sampling errors) sampled point sets are transformed according to the simulated breathing transformation. In this case an exact correspondence of points is given initially and an ideal shape matching method would identify 100% correct correspondences. The percentage of correctly identified correspondences over different numbers of sampling points is given in Table 1. The expected 100% correspondence was achieved in most cases. The higher number of sampling points show a performance decrease, since a very high number of shape context histogram bins would be necessary to prevent discretization artifacts.

[%] / Samples	**400**	**800**	**1200**	**1600**	**2000**	**2400**	**2800**	**3200**	**3600**
Correspondences	100	100	100	100	99.6	99.1	97.3	94.5	89.2

Table 1. Percentage of correctly identified correspondences from synthetically transformed diaphragm point sets.

To assess the dependencies on sampling inaccuracies the segmented volumes are then transformed instead of the sampled point sets. Each transformed volume is sampled independently. First of all the number of sampled points is varied while keeping the translation parameters at the same value as in the previous experiment. This demonstrates the matching algorithms behaviour with increasing sampling rate, which clearly an increase using more sample points with a saturation at very high numbers. Table 2 shows these results.

[mm] / Samples	**400**	**800**	**1200**	**1600**	**2000**	**2400**	**2800**	**3200**	**3600**
Min Distance	0.3091	0.2709	0.2161	0.1288	0.1007	0.1351	0.0639	0.1033	0.0580
Mean Distance	8.9919	5.8678	6.0271	5.5008	5.2024	4.9839	4.7611	4.8170	4.4415
Max Distance	25.9215	24.3896	22.4572	22.6455	23.3288	21.2271	18.1501	21.2802	18.4788

Table 2. Distance error measures for synthetically transformed diaphragm volumes over different sampling sizes.

From now on the number of diaphragm sample points is fixed to 2000 sample points. Following the previous results this is a good compromise between running time and matching quality, especially since the quality does not increase drastically for higher number of sample points. t_{inward} is set to 10 mm and $t_{vertical}$ is varied between 0 and 45 mm. Table 3 depicts these results, the same experiment with $t_{vertical}$ set to 25 mm and t_{inward} being varied show a similar result. Performance decreases when deformations get too large.

Another set of experiments is performed on airway tree and lung lobe surfaces. Airway tree and lung lobe

[mm] / t	0	5	10	15	20	25	30	35	40	45
Min Dist	0.1212	0.05007	0.1358	0.0342	0.1539	0.1845	0.1463	0.1040	0.0974	0.2626
Mean Dist	3.4065	3.4537	3.5498	3.8951	4.5062	4.9031	5.3673	6.0291	6.5062	7.4041
Max Dist	12.1399	11.6625	12.2154	14.3470	17.2173	17.5739	19.1093	21.5376	26.4628	27.8111

Table 3. Distance error measures for synthetically transformed diaphragm volumes over different transformations.

segmentations from a data set at TLC are used as input to the shape matching algorithm. Again, the vertical translation $t_{vertical}$ and the inward translation t_{inward} remain fixed to 25 and 10 mm, respectively. The number of sample points is varied between 200 and 3400 sample points in steps of 400. Table 4 summarizes the achieved results which are similar to the results from the diaphragm experiments.

	[mm]/Samples	200	600	1000	1400	1800	2200	2600	3000	3400
Airway	Min Dist	1.7055	0.3713	0.3554	0.1322	0.1694	0.1642	0.1923	0.1314	0.0696
	Mean Dist	11.7425	8.3019	7.3544	5.9857	5.8020	5.7074	5.5174	5.0402	4.5784
	Max Dist	32.7307	28.5479	31.1092	24.9632	25.3906	26.3004	27.5073	27.1693	25.6884
Lobe	Min Dist	1.9441	0.5800	0.2449	0.3710	0.2271	0.1464	0.3372	0.2448	0.0835
	Mean Dist	15.8415	9.4848	7.9323	7.1462	6.7032	6.4448	5.7276	5.9852	5.9906
	Max Dist	38.9555	26.5770	28.3671	24.0893	19.1600	21.7416	22.1864	27.2684	26.1502

Table 4. Distance error measures for synthetically transformed airway & lung volumes over different sampling sizes.

Fig. 6a,b,c shows matching results for synthetically transformed diaphragm, lung lobe and airway tree surfaces transformed with $t_{vertical}$ of 25 mm, t_{inward} of 10 mm and sampled with 2000 sample points, respectively.

4.2. Registration Experiments on Synthetically Transformed Data

The synthetic registration experiments use a data set A and its transformed version B as input to the non-linear registration algorithm which is composed of the shape matching and the displacement field interpolation component. Experiments are solely performed on lung lobe surface segmentations. Again a transformation $T : \{B(x, y, z) = T(A(x, y, z))\}$ is defined by the evaluator and unknown to the algorithm. The point correspondences $< p_i, q_i' >$ that are created by shape matching of an original and a transformed data set are used to calculate a displacement field by regularized spline-based interpolation of the point correspondences within the thin-plate spline framework. This displacement field is utilized to warp the original data set A to a data set B'. By comparing B' with the synthetically transformed data set B the two quantitative error measurements relative volume overlapping error (RVOE) and target registration error (TRE) are calculated. RVOE is defined as $1 - \frac{B \cap B'}{B \cup B'}$, perfect registration would result in a value of 0 while 1 resembles total misregistration. The target registration error is calculated by manually identifying a number of corresponding points in the lung lobes. Therefore the original grey value data set is transformed according to the synthetic transformation T and afterwards the airway tree is segmented in both the original and the transformed image. Both airway tree segmentations are skeletonized and 20 branch points are extracted, respectively. Fig. 6f,g shows the graph-based airway tree representations and the manually labeled airway tree branch points. By warping identified airway tree branch points of data set A according to the displacement field from the registration step and calculating their distance from the corresponding branch points in synthetically transformed data set B, the minimum, mean and maximum target registration error is computed.

Five thorax data sets are used in this experiment. The first data set is examined at varying numbers of sample points between 600 and 2000, while the other four data sets are solely examined with 1000 sample points. RVOE (in %) and TRE results (in mm) for data set *T32* are given in Table 5. The mean RVOE of all 5 data sets is 10.166% while mean target registration errors over all 5 data sets are 1.576mm, 6.734mm and 16.502 for minimum, mean and maximum TRE, respectively. $t_{vertical}$ is 25 mm and t_{inward} is 10 mm. Fig. 6 shows one lung lobe of data set *T32* before (d) and after (e) registration.

4.3. Registration Experiments on Real Thorax Data

The first validation experiment on the real thorax data makes use of a sheep thorax data set at five different breathing states between TLC and FRC, these data sets are called *T32*, *T24*, *T16*, *T8* and *F*. Each of them

	Measure/Samples	200	600	1000	2000
T32	RVOE [%]	12.47	10.89	9.54	8.26
	Min/Mean/Max TRE	1.81/7.39/14.03	1.70/6.06/11.09	1.79/5.37/10.08	1.613/5.1672/9.891

Table 5. RVOE and TRE for synthetically transformed lung volume *T32* over different sampling sizes.

contains a lung lobe segmentation from which 1000 points are sampled. Four data subsets are built, with the first subset consisting of states {*T32*, *T24*, *T16*}, the second subset consisting of {*T24*, *T16*, *T8*}, third subset {*T16*, *T8*, *F*} and fourth subset {*T32*, *T16*, *F*}. For each of these subsets the transformation T' relating first and second state of the subset and T'' relating second and third state of the subset is calculated by using the shape context matching approach. Further the transformation T''' is calculated relating first and third state of the subset. By comparing the results of applying T''' and $T''(T')$ on the points of the first state of each subset minimum, mean and maximum distances are calculated and shown in Table 6.

[mm] / Data Subset	{T32,T24,T16}	{T24,T16,T8}	{T16,T8,F}	{T32,T16,F}
Min Distance	0.1538	0.2732	0.2458	0.1997
Mean Distance	7.76223	7.9934	8.2067	8.2582
Max Distance	24.6284	24.2931	23.9196	30.1443

Table 6. Real data experiment showing accuracy of the interpolation of intermediate breathing states for five data sets.

For the interpolation of the displacement fields these validations inherently assume a linear relationship between the different states in the breathing cycle, which is not necessarily the case for real thorax breathing motion. Therefore, a second validation experiment is performed on the real thorax data by using the airway tree branching points as corresponding landmarks again to calculate target landmark registration errors. The same procedure as described in the previous section is used. Registering two real data sets results in a displacement field which is applied to the manually identified airway tree branching points of the first data set and compared to the manually identified airway tree branching points of the second data set. In this way minimum, mean and maximum target registration errors are computed (see Table 7). Fig. 6 shows the result of overlaying data sets *T8* and *T32* before (f) and after (g) registration.

[mm] / Data Set	T32/T24	T32/T16	T32/T8	T32/F	T24/T16	T24/T8	T24/F
Min Distance	0.9465	0.3671	0.5833	2.4355	1.4852	1.3285	1.8347
Mean Distance	5.6503	4.9746	4.8215	6.2446	7.1077	7.4101	7.9520
Max Distance	22.2019	18.8932	14.1134	17.2176	22.1003	22.5588	20.8391

Table 7. Real data experiment showing target registration errors of pairs of data sets.

5. DISCUSSION AND CONCLUSION

Experiments show that the non-linear registration method based on the shape context matching approach is a well-suited method for a variety of soft tissue organ surfaces. Shape matching and subsequent volume registration and warping was successfully performed using airway tree, lung, lung lobe and diaphragm segmentations. Experiments on synthetic and real data sets show mean registration errors in the range of 5 to 8 mm. Since there is a sophisticated image processing pipeline necessary for the whole registration process, the final error is composed of several components. First of all segmentation and surface point sampling introduce errors that have an effect on the registration result. The matching validation experiments using synthetically transformed point sets assess the effect of segmentation errors. They show that the shape context descriptor based matching approach is a suitable method if there are no segmentation errors present, since the percentage of correctly identified correspondences stays stable at almost 100% over a wide range of different numbers of sample points. However, a very large number of sample points introduces inaccuracies due to the discrete nature of the shape context histogram. For a stable matching quality histogram size would have to be increased at larger numbers of sample points which leads to high computation times. The matching validation experiments using synthetically

transformed input volumes and independently sampled point sets clearly show the dependency of matching quality on the number of sample points. In accordance to common insights from sampling theory an increase in the number of sample points increases matching quality until it stays stable between 4 and 5 mm starting around 2000 sample points. This is consistent with the former experiment, since at higher sampling rates histogram discretization issues come into play. The error of 4 to 5 mm can be explained by point sampling issues. The matching method uses independently sampled points of two shapes, the established correspondences implicitly carry an error in the sampling positions which becomes relevant in the graph-based optimization to find the one-to-one correspondences.

The validation experiments with varying synthetic deformation forces show that larger deformations induce a degrading matching accuracy, this indicates that the matching method is not suitable to model extremely large movements. Synthetic and real-data registration experiments also show acceptable behaviour. Mean target registration errors are in the range of 5 to 9 mm. Here the effects of the preceding image processing pipeline steps have to be taken into account. Segmentation inaccuracies, surface sampling issues and discretization effects in the shape context matching all sum up to form an error which gets slightly smoothed by the thin-plate spline interpolation process at those points that are interpolated, but not at the correspondences. In addition, calculation of target registration errors is an error-prone process by itself, especially since for these validation studies the branchpoint labeling was not performed by a medical expert but by the first author of this paper.

The outcome of this work is a method that allows to calculate deformation fields capturing organ motion. In this paper the focus lay on breathing motion but the basic concepts can be used for other kinds of motion as well. Displacement fields may be derived from organs like diaphragms or lung surfaces which enables one to build statistical models of organ motion, e.g. by using Active Shape Models[24] in subsequent stages. A statistical model of breathing motion resembles a very useful tool for segmentation and registration applications in medical imaging areas that suffer from motion artifacts.

In this paper a 3D extension of the shape context approach for matching and registering 3D surfaces was presented and experiments on lung, lung lobe, airway tree and diaphragm surfaces were shown. Shape context based non-linear registration is a promising technique which has to be studied further to be able to lower the TRE. Future work will include algorithm fine-tuning on the one hand and more validation experiments on the other hand. Fine-tuning might be performed by using the Normalized Mutual Information measure instead of Normalized Correlation Coefficients for the intensity-based cost function and by finding a better way to get rid of segmentation and discretization errors and outliers. The thin-plate spline displacement field interpolation might be fine-tuned by using the weights from the matched points for regularization instead of simply having a single regularization parameter. A more appropriate validation study should include a large number of data sets and use manual correspondences identified by experts for the target registration error calculation. Further, the method should be compared to other state-of-the-art techniques. Based on a large number of validated data sets, a statistical model of the displacement fields that describes breathing motion could be established.

ACKNOWLEDGMENTS

Parts of this work were supported by the Austrian Science Foundation under grant P14897 and grant P17066-N04.

REFERENCES

1. E. A. Hoffman, J. K. Tajik, and S. D. Kugelmass, "Matching Pulmonary Structure and Perfusion via combined Dynamic Multislice CT and Thin-Slice High-Resolution CT," *Computerized Medical Imaging and Graphics* **19**(1), pp. 101–112, 1995.
2. L. Weruaga, J. Morales, L. Nunez, and R. Verdu, "Estimating volumetric motion in human thorax with parametric matching constraints," *IEEE Transactions on Medical Imaging* **22**, pp. 766–772, June 2003.
3. T. Rohlfing, C. R. Maurer, W. O. O'Dell, and J. Zhong, "Modeling liver motion and deformation during the respiratory cycle using intensity-based free-form registration of gated MR images," in *Proceedings of SPIE: Medical Imaging: Visualization, Display and Image Guided Procedures*, **4319**, pp. 337–348, 2001.
4. J. Blackall, G. Penney, A. King, A. Adam, and D. Hawkes, "Tracking alignment of sparse ultrasound with preoperative images of the liver and an interventional plan using models of respiratory motion and deformation," in *Medical Imaging 2004: Visualization, Image-Guided Procedures and Display*, *Proceedings of SPIE* **5637**, pp. 218–227, 2004.

5. J. B. A. Maintz and M. A. Viergever, "A survey of Medical Image Registration," *Medical Image Analysis* **2**(1), pp. 1–36, 1998.
6. M. A. Audette, F. P. Ferrie, and T. M. Peters, "An algorithmic overview of surface registration techniques for medical imaging," *Medical Image Analysis* **4**, pp. 201–217, 2000.
7. B. Zitova and J. Flusser, "Image registration methods: A survey," *Image and Vision Computing* **21**, pp. 977–1000, October 2003.
8. R. C. Veltkamp and M. Hagedoorn, "State of the Art in Shape Matching," Tech. Rep. UU-CS-1999-27, Utrecht, 1999.
9. F. L. Bookstein, "Principal Warps: Thin-Plate Splines and the Decomposition of Deformations," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11**, pp. 567–585, June 1989.
10. S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**(4), pp. 509–522, 2002.
11. A. Frome, D. Huber, R. Kolluri, T. Buelow, and J. Malik, "Recognizing Objects in Range Data Using Regional Point Descriptors," in *Proceedings of the European Conference on Computer Vision (ECCV) 2004*, *LNCS*(3023), pp. 224–237, (Springer Verlag Heidelberg-Berlin), 2004.
12. M. H. Davis, A. Khotanzad, D. P. Flamig, and S. E. Harms, "A Physics-Based Coordinate Transform for 3-D Image Matching," *IEEE Transactions on Medical Imaging* **16**, pp. 317–328, June 1997.
13. M. Betke, H. Hong, and J. P. Ko, "Automatic 3D Registration of Lung Surfaces in Computed Tomography Scans," in *Proceedings of the 4th Int Conf on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, **4**, pp. 725–733, 2001.
14. L. Fan, C. W. Chen, J. M. Reinhardt, and E. A. Hoffman, "Evaluation and Application of 3D Lung Warping and Registration Model Using HRCT Images," in *Proceedings of SPIE: Medical Imaging: Physiology and Function from Multidimensional Images*, **4321**, pp. 234–243, 2001.
15. B. Li, G. E. Christensen, J. Fill, E. A. Hoffman, and J. M. Reinhardt, "3D Inter-Subject Warping and Registration of Pulmonary CT Images for a Human Lung Model," in *Proceedings of SPIE: Medical Imaging: Physiology and Function from Multidimensional Images*, **4683**, pp. 324–335, 2002.
16. J. Tschirren, K. Palagyi, J. M. Reinhardt, E. A. Hoffman, and S. M., "Segmentation, Skeletonization, and Branchpoint Matching - A Fully Automated Quantitative Evaluation of Human Intrathoracic Airway Trees," in *Proceedings of the 5th Int Conf on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, (Tokyo, Japan), 2002.
17. S. Sclaroff and A. Pentland, "Modal Matching for Correspondence and Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence* **17**, pp. 545–561, June 1995.
18. H. Chui and A. Rangarajan, "A new point matching algorithm for non-rigid registration," *Journal of Computer Vision and Image Understanding* **89**, pp. 114–141, 2003.
19. T. Pock, C. Janko, R. Beichel, and H. Bischof, "Multiscale Medialness for Robust Segmentation of 3D Tubular Structures," in *Proceedings of the Computer Vision Winter Workshop 2005*, p. accepted, 2005.
20. R. Beichel, T. Pock, C. Janko, R. Zotter, B. Reitinger, A. Bornik, K. Palagyi, E. Sorantin, E. Werkgartner, H. Bischof, and M. Sonka, "Liver segment approximation in CT data for surgical resection planning," in *SPIE Medical Imaging 2004: Image Processing*, J. Fitzpatrick and M. Sonka, eds., **5370**, pp. 1435–1446, 2004.
21. G. Donato and S. Belongie, "Approximate Thin Plate Spline Mappings," in *Proceedings of the European Conference on Computer Vision (ECCV)*, *Lecture Notes in Computer Science* **2352**, pp. 21–31, Springer Verlag, (Heidelberg), 2002.
22. K. Rohr, H. S. Stiehl, R. Sprengel, T. M. Buzug, J. Weese, and M. H. Kuhn, "Landmark-Based Elastic Registration Using Approximating Thin-Plate Splines," *IEEE Transactions on Medical Imaging* **20**, pp. 526–534, June 2001.
23. S. C. Davies, A. L. Hill, R. B. Holmes, M. Halliwell, and P. Jackson, "Ultrasound Quantitation of Respiratory Organ Motion in the Upper Abdomen," *British Journal of Radiology* , p. 1096, 1102 1994.
24. T. F. Cootes and C. J. Taylor, "Statistical models of appearance for medical image analysis and computer vision," in *Proceedings of SPIE: Medical Imaging: Image Processing*, (4322), pp. 236–248, 2001.

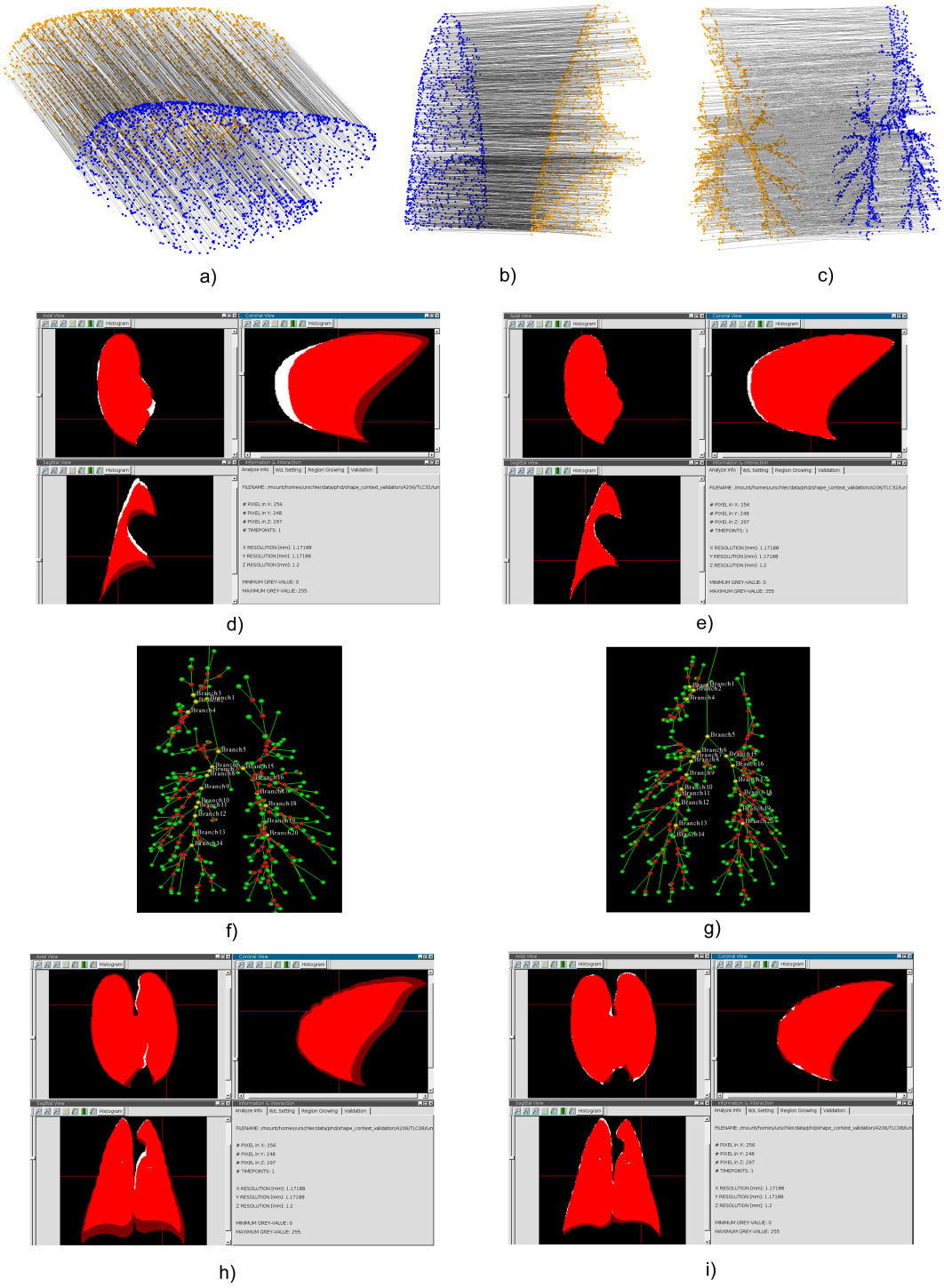

Figure 6. Some shape matching results are depicted in a), b) & c). d) & e) show registration results of synthetically transformed data set *T32*. f) & g) give an example of corresponding airway branch points that were used to determine target registration error. h) & i) show results from registering data sets *T32* & *T8*. In d), e), h) and i) the white volumes are the original and dark volumes are the overlaid warped volumes.