Noname manuscript No.
(will be inserted by the editor)

A machine learning pipeline for internal anatomical landmark embedding based on a patient surface model

Xia Zhong · Norbert Strobel · Annette Birkhold · Markus Kowarschik · Rebecca Fahrig · Andreas Maier

Received: date / Accepted: date

Abstract Purpose: With the recent introduction of FAST (fully assisting scanner technologies) by Siemens Healthineers (Erlangen, Germany), a patient surface model was introduced to the diagnostic imaging device market. Such a patient representation can be used to automate and accelerate the clinical imaging workflow, manage patient dose, and provide navigation assistance for computed tomography diagnostic imaging. In addition to diagnostic imaging, a patient surface model has also tremendous potential to simplify interventional imaging. For example, if the anatomy of a patient was known, a robotic angiography system could be automatically positioned such that the organ of interest is positioned in the system's iso-center offering a good and flexible view on the underlying patient anatomy quickly and without any additional X-ray dose. **Method:** To enable such functionality in a clinical context with sufficiently high accuracy, we present an extension of our previous patient surface model by adding internal anatomical landmarks associated with certain (main) bones of the human skeleton, in particular the spine. We also investigate different approaches for positioning of these landmarks employing CT datasets with annotated internal landmarks as training data. The general pipeline of our proposed method comprises the following steps: First, we train an active shape model using an existing avatar database and segmented CT surfaces. This stage also includes a gravity correction procedure, which accounts for shape changes due to the fact that the avatar models were obtained in standing position, while the CT data was acquired with patients in

Xia Zhong · Andreas Maier
Pattern Recognition Lab, Department of Computer Science, Friedrich-Alexander-Universität Erlangen-Nürnberg, Martensstr. 3, 91058 Erlangen, Germany E-mail: xia.zhong@fau.de

Norbert Strobel
Fakultät für Elektrotechnik, Hochschule für angewandte Wissenschaften Würzburg-Schweinfurt, Schweinfurt, Germany

Annette Birkhold · Markus Kowarschik · Rebecca Fahrig
Siemens Healthineers, Advanced Therapies, Forchheim, Germany

supine position. Second, we match the gravity-corrected avatar patient surface models to surfaces segmented from the CT datasets. In the last step, we derive the spatial relationships between the patient surface model and internal anatomical landmarks. **Result:** We trained and evaluated our method using cross-validation using 20 datasets, each containing 50 internal landmarks. We further compared the performance of four different generalized linear models set ups to describe the positioning of the internal landmarks relative to the patient surface. The best mean estimation error over all the landmarks was achieved using lasso regression with a mean error of 12.19 $\pm$ 6.98 mm. **Conclusion:** Considering that interventional X-ray imaging systems can have detectors covering an area of about 200 mm $\times$ 266 mm (20 cm $\times$ 27 cm) at iso-center, this accuracy is sufficient to facilitate automatic positioning of the X-ray system.

Keywords Patient modelling · anatomical landmark · statistical shape model · interventional · X-ray · imaging

1 Introduction

With better availability of affordable yet sufficiently accurate 3D RGB-D cameras, estimating patient models in a diagnostic or interventional imaging setting has become feasible. The availability of patient models now triggers new research question such as how to use them to optimize existing workflows, e.g., with respect to radiation safety or ease of use. For example, in the diagnostic environment, Singh et al. [16] demonstrated that a patient surface model can be fitted to the patient ahead of CT scans using a RGB-D camera. Knowing the position of the patient model in the table coordinate system, an algorithm was designed to optimally adjust the table height and scanning area and thus to automate and accelerate the diagnostic CT imaging workflow. In the interventional environment, a patient surface model can, for example, be used to improve X-ray dose management or to simplify system use. Johnson et al. evaluated the accuracy of skin dose estimations using different patient models [11]. The type of patient models varied from a stylized model (lowest accuracy) to patient-specific models (highest accuracy). Their experiments demonstrated that more accurate patient models led to better skin dose estimates. Furthermore, Bednarek et al. [5] showed that reliable, real-time skin dose estimation results can be obtained by using an anthropomorphic patient model and careful modeling of the underlying X-ray imaging physics. Patient surface models were also employed to analyze breathing motion [20], or to estimate internal organ motion from surface motion [17]. Further interventional applications include C-arm exposure control in knee imaging [12], navigation assistance [4] in radiation therapy, and anatomy lessons using augmented reality [6]. However, in these applications, most patient models were limited to surface models providing information about the body shape only. No clinically relevant internal anatomical objects have been included in these surface models yet. Since organ positions do, however, correlate with human outer

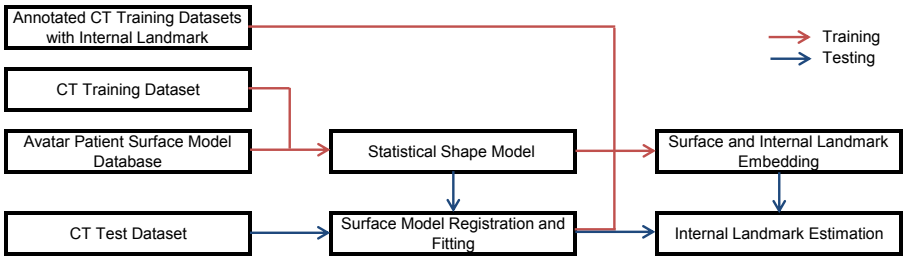

Fig. 1: The workflow of the proposed method during training (red) and testing (blue) phase. The CT training and testing datasets are CT volumes in supine position. The annotated CT datasets are CT volumes with annotated landmark position in supine position. The avatar surface database comprises 3D surface meshes acquired in standing position.

shape, it should be possible to use surface models to obtain at least a priori information about where certain organs are most likely to be found. To reliably embed internal organs into surface models, we propose to use internal landmarks providing an internal frame of reference relative to which organ positions are to be determined. A natural frame of reference is the spine. In the related field of computer vision, deriving articulated models or skeletons of a surface mesh is well studied. One of the most widely used methods to predict the skeleton of a 3D image (represented either as a surface model or as a depth image) was proposed by Anguelov et al. [2]. At first, a template is fitted to the target surface mesh using non-rigid mesh registration. Then, the fitted template is partitioned and an expectation maximization algorithm is used to determine the skeleton. Another approach to estimate the skeleton of a 3D mesh was proposed by Au et al. [3]. This method gradually shrinks the surface mesh using Laplacian based operations. Afterwards, the connectivity of the contracted mesh is adjusted to obtain a final skeleton of the 3D mesh. While these methods can be used for animation of 3D human models as, e.g., used in movies, they fail to compute a realistic human skeleton when applied to patient surface models. This is why we propose a new method for finding anatomical relevant internal landmarks. We start with a surface model calculated based on a patient's height and weight and involves estimated relationships between internal landmark positions relative to the outer body shape.

2 Materials and Methods

To extend the surface model with the internal landmarks, we use CT datasets with annotated landmarks. The learning based approach has three main steps. The workflow of the proposed method is shown in Fig. 1. In the training phase, we learn a statistical shape model using the avatar database first. The shape deformation due to the pose differences (avatars: standing vs. CT voxel volumes: suppine position) was also taken into account in this step using CT train-

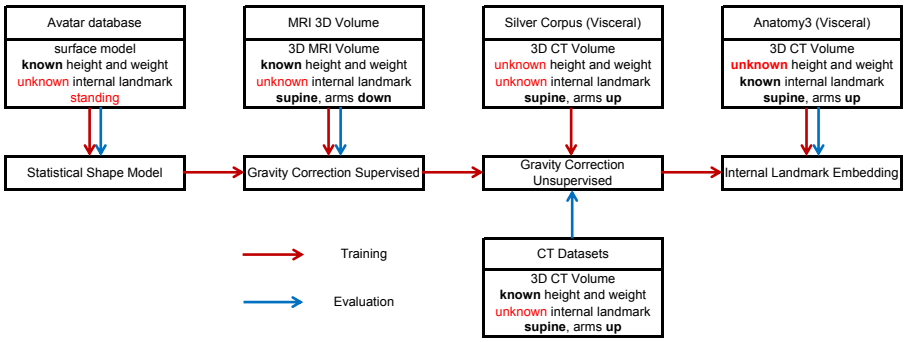

Fig. 2: An illustration of the datasets used and associated step in the pipeline.

ing datasets. Second, for each CT volume, a corresponding trained shape model (surface) was determined. Third, we learned the mapping function $\mathcal{M}_i(\cdot)$ to establish a relation between the t^{th} model $\hat{\mathbf{x}}_t$ and the associated i^{th} internal landmark $\mathbf{l}_{t,i}$, such that

$$\mathcal{M}_i = \arg\min_{\mathcal{M}_i} \sum_t \|\mathbf{l}_{t,i} - \mathcal{M}_i(\hat{\mathbf{x}}_t)\|_2^2 \tag{1}$$

The model $\hat{\mathbf{x}}_t$ was matched to the segmented CT surface $\mathbf{x}_t$ using the trained active shape model. We estimated the mapping function $\mathcal{M}_i(\cdot)$ for each landmark separately. Each step of the training pipeline employs different types of datasets containing complementary anatomical information. Fig. 2 highlights the connection between used datasets and associated step in the pipeline. First, to train the active shape model of the patient surface model, an avatar database derived from the Civilian American and European Surface Anthropometry Resource Project (CAESAR) [13] is used (surface shape, height and weight). Second, 25 MRI datasets with patients' height and weight are used to calculated the deformation correction in a supervised manner. Third, the SilverCorpus dataset of the Visceral datasets [19] (60 full body CT datasets, without height and weight) is used to refine the deformation field using an unsupervised method. Fourth, the Anatomy3 dataset of the Visceral dataset [19] (20 full-body CT scans, with 50 associated internal landmarks in each dataset, without height and weight) comprising internal landmarks is used to train the internal landmark embedding network. For the evaluation additional five CT datasets without internal landmarks, but with known height and weight are used.

In the testing phase, a model is registered and fitted to the Anatomy3 CT dataset using the trained shape model and gravity correction field. Using the learned mapping function $\mathcal{M}_i(\cdot)$ and fitted model, the position of the internal landmark can be estimated. Then we can assess how close it is to the actual landmark of this CT test dataset.

2.1 Statistical Shape Model

We first trained a statistical shape model for the patient model using an avatar database. All of the avatar surface models have been derived from a common standard template. As a consequence, they all have the same number of vertices. The vertices across the patient models share the same correspondences. For example, a vertex positioned on the tip of the nose will have the same index in each patient model. Using this database, a statistical shape model of the surface model can be trained and expressed as

$$\mathbf{x}_t = \bar{\mathbf{x}} + \mathbf{D}\mathbf{b}_t \tag{2}$$

where $\bar{\mathbf{x}}$ denotes the mean surface model, matrix $\mathbf{D}$ denotes the modes of variation, and the vector $\mathbf{b}_t$ denotes the associated weighting vector of ground truth surface mesh $\mathbf{x}_t$. However, this trained statistical shape model represents a surface model of a person in standing pose. In most medical applications, the patient is, however, in supine position. We proposed a shape correction term for this pose difference, called gravity-deformation term $\mathbf{G}$, in our previous work [21].

$$\hat{\mathbf{x}}_t = \bar{\mathbf{x}} + (\mathbf{D} + \mathbf{G})\mathbf{b}_t \tag{3}$$

This term $\mathbf{G}$ was determined in a supervised manner. To this end, MRI data with known gender, height, and weight was used.

In this work, we extended our previous approach to an unsupervised learning method as information about gender, height, and weight of the patient is often not available in clinical datasets due to anonymization. Since we neither knew height and weight of the new (CT) datasets nor was the gravity-deformation term available to us, we used an expectation maximization (EM) algorithm to iteratively solve for them. Before starting with the EM scheme, we first segmented the surface model of each CT dataset with unknown height and weight information. In the E step, we estimated for each segmented CT surface the shape parameter vector $\mathbf{b}$ with given modes of variation $\mathbf{D}$ and fixed gravity-deformation term $\mathbf{G}$. This estimation can be done by minimizing

$$\mathbf{b}_t = \arg\min_{\mathbf{b}} \mathcal{D}\left(\mathbf{x}_t, \bar{\mathbf{x}} - (\mathbf{D} + \mathbf{G})\mathbf{b}_t\right)^2 \tag{4}$$

where $\mathcal{D}(\cdot, \cdot)$ denotes the mean surface distance between the two meshes. In the M step, we first calculated the deformation vector between the estimated surface $\hat{\mathbf{x}}_t$ derived from the E step and the ground truth $\mathbf{x}_t$. This is an non-rigid mesh registration problem and can be solved by estimating the system matrix $\mathbf{A}_t \in \mathbb{R}^{N\times 4\times 4}$

$$\mathbf{A}_t = \arg\min_{\mathbf{A}} \mathcal{D}(\mathbf{x}'_t, \mathbf{A}\hat{\mathbf{x}}'_t)^2 + \lambda \sum_{j\in\mathcal{N}(i)} \|\mathbf{A}_i - \mathbf{A}_j\|_2^2 \tag{5}$$

where the scalar N denotes the number of datasets and vector $\mathbf{x}'_t$ and $\hat{\mathbf{x}}'_t$ denote the homogeneous coordinates of vector $\mathbf{x}_t$ and $\hat{\mathbf{x}}_t$, respectively. Afterwards, we calculated the residual discrepancy vector $\mathbf{r}'_t = \mathbf{A}_t\hat{\mathbf{x}}'_t - \hat{\mathbf{x}}'_t$. This residual error

is related to estimation errors of the gravity deformation field $\mathbf{G}$. To arrive at a better gravity deformation matrix, we can update this term by minimizing

$$\delta\mathbf{G} = \arg\min_{\delta\mathbf{G}} \sum_t \|\mathbf{r}_t - \delta\mathbf{G}\mathbf{b}_t\|_2^2 \tag{6}$$

This problem can be solved using SVD and we update the term $\mathbf{G} = \mathbf{G} + \alpha \cdot \delta\mathbf{G}$ where α is the learning rate. To tackle this highly non-convex optimization problem, we use $\mathbf{G}$ from the supervised method in our previous work as initialization and use the EM algorithm until convergence. The system matrix $\mathbf{A}$ is obtained by calculating the derivative and using a BroydenFletcherGoldfarbShanno (BFGS) optimizer [9]. We set the learning rate α as 0.1.

2.2 Surface Model Registration and Fitting

To determine how to embed landmarks into our patient surface models, we started by matching surface models to those CT volumes where landmarks were present (Anatomy3 of Visceral). Since height and weight of our landmark-annotated datasets were again unknown, we needed to find the best shape parameter vector $\mathbf{b}_t$ and associated rigid transformation $\mathcal{T}_t$. This was accomplished by minimizing

$$\mathbf{b}_t, \mathcal{T}_t = \arg\min_{\mathbf{b},\mathcal{T}} \mathcal{D}\left(\mathbf{x}_t, \mathcal{T}(\bar{\mathbf{x}} + (\mathbf{D} + \mathbf{G})\mathbf{b})\right)^2 \tag{7}$$

This optimization problem can be solved efficiently using alternating minimization. The challenge in estimating the transformation $\mathcal{T}$ is that the available CT may have less body coverage than our patient surface models. In addition, the estimated transform should be robust such that two meshes which are related by non-rigid deformation can be registered. To arrive at such a transform, we used mesh feature descriptors and feature matching. To establish correspondences between feature points robustly, we first sampled the key points as proposed by Sahillioğlu et al. [15], such that the sampling is almost uniform. In this way, we avoid accumulation of feature points in high curvature regions when identifying key points. We used the LDSIFT [7] feature descriptor to calculate the feature map for each key point. This descriptor has shown robustness to non-rigid deformation between source and target meshes. As the human body is symmetric, the feature matching potentially finds false matches (mirrored at a patient's longitudinal axis of symmetry). To rule them out, we relied on the fact that for patients in prone position the transformation $\mathcal{T}$ is primarily a translation with very limited rotation. In our actual implementation, we divided the surface mesh into eight primary districts using the principal axis of the mesh, and rejected any cross-district matching. Due to the non-rigid deformation between the two meshes, we may, however, still encounter ambiguous solutions for the rigid transformation. To ensure robustness, we considered all feasible results using RANSAC [8], put them into motion clusters, and used a mean shift algorithm to find the center

of the biggest cluster. This cluster center is used as our motion estimation $\mathcal{T}_t$. After each iteration, the origin of the district's coordinate system will be updated accordingly. Knowing the transformation $\mathcal{T}_t$, the weighting factor $\mathbf{b}_t$ can be estimated by minimizing

$$\mathbf{b}_t = \arg\min_{\mathbf{b}} \mathcal{D}\left(\mathbf{x}_t, \mathcal{T}_t(\bar{\mathbf{x}} + (\mathbf{D} + \mathbf{G})\mathbf{b})\right)^2 + \lambda \mathcal{R}\left(\mathbf{x}_t, \hat{\mathbf{x}}\right) \tag{8}$$

where the scalar λ is a parameter, and the function $\mathcal{R}(\cdot, \cdot)$ calculates the symmetric overlap ratio between the bounding box of the segmented and the estimated surface mesh. This function regularizes the minimization problem such that the estimated surface will not be degenerated. After this step, the residual error between estimated and segmented surface was minimized using a non-rigid mesh registration, as proposed by Allen et al. [1].

2.3 Surface and Internal Landmark Embedding

After fitting the surface model $\hat{\mathbf{x}}$ to the segmented CT surface $\mathbf{x}$, we search for the mapping $\mathcal{M}_i(\cdot, \cdot)$ between $\hat{\mathbf{x}}$ and the annotated internal landmarks $\mathbf{l}_i$ in the CT datasets. For the i^{th} landmark in the t^{th} CT data, $\mathbf{l}_{t,i}$, we assume that the landmark position (coordinates) can be expressed as a generalized linear combination of neighboring surface vertices $\hat{\mathbf{x}}_{t,j} \in \mathcal{N}(\mathbf{l}_i)$. The neighborhood was defined using the k-nearest neighbors based on the Mahalanobis distance. We used the covariance matrix of the surface model $\mathbf{x}$ to calculate the Mahalanobis distance. We selected a large neighborhood such that the residual error due to non-linearity in the mapping was minimized. The number and indices of the neighborhood vertices can vary in each dataset. We used the unions according to the indices as the neighborhood for each internal landmark. To describe this more formally, we introduced the matrix $\mathbf{X}_{\mathcal{N}(\mathbf{l}_i)} = [X^T_{1,\mathcal{N}(\mathbf{l}_i)}, \cdots, X^T_{N,\mathcal{N}(\mathbf{l}_i)}]^T$ and the matrix $\mathbf{L}_i = [\mathbf{l}^T_{1,i}, \cdots, \mathbf{l}^T_{N,i}]^T$ where matrix $X_{t,\mathcal{N}(\mathbf{l}_i)} = [\mathbf{x}_{t,1}, \cdots, \mathbf{x}_{t,M}]$ comprises all vertices in an associated neighborhood for the t^{th} dataset and the landmark $\mathbf{l}_{t,i}$. The scalar M denotes the number of neighboring vertices. At this point, we can formulate a generalized linear mapping with a cost function for the linear mapping weighting vector $\mathbf{w}_i$ for landmarks $\mathbf{l}_{t,i}$.

$$\mathbf{w}_i = \arg\min_{\mathbf{w}_i} \|\mathbf{X}_{\mathcal{N}(\mathbf{l}_i)}\mathbf{w}_i - \mathbf{L}_i\|_2^2 + \lambda \mathcal{P}(\mathbf{w}_i) \tag{9}$$

In Eq. 9, the scalar λ is a Lagrange multiplier for the penalty function $\mathcal{P}$. The penalty function $\mathcal{P}$ is different for each generalized linear mapping. In ridge regression [10], the L2 norm of $\mathbf{w}$ was used while in lasso regression [18], we relied on the L1 norm. For elastic net regression [22], a mixture of L1 and L2 norm was employed, while in linear embedding [14] the sum of $\mathbf{w}$ was used. Afterwards, we compared the results obtained using either linear embedding, or ridge regression, or lasso regression or elastic net regression when estimating the landmark positions. Note, that in case of linear embedding regression, the minimization problem can be solved in closed form.

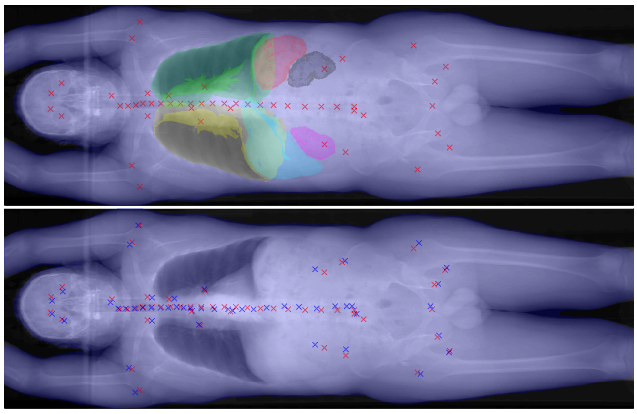

Fig. 3: A sample of the Anatomy3 training set is shown on the top. The annotated internal landmarks are indicated using red crosses and the segmented internal organs are shown in different colors. The bottom figure shows the output of our algorithm, i.e., internal landmarks (in blue) derived based on the surface.

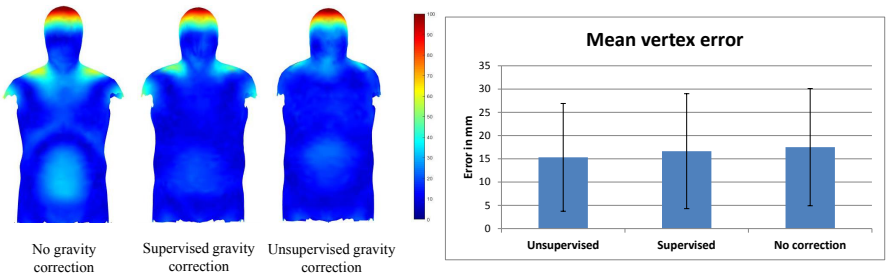

Fig. 4: Qualitative (left) and quantitative (right) evaluation of the estimation error using unsupervised, supervised and no gravity correction.

3 Evaluation and Results

We trained our algorithm using the Silvercopus and Anatomy3 from the Visceral dataset [19]. The Silvercopus comprises 60 full body CT datasets. It was used to train the active shape model with the gravity-deformation for the CT datasets using unsupervised learning. We evaluate the model estimation error using five CT datasets with known height and weight. We estimate the model with the patient metadata from the CT dataset using the statistical shape model, the shape model with supervised gravity correction, and the shape model with unsupervised gravity correction. We compared the estimation error of these three different methods using mean vertex distance. The error in head and shoulder is not considered in the mean vertex calculation, as the head is only partially visible in the CT datasets, and the arms are placed

upwards. Fig. 4 show the qualitative (left) and quantitative (right) result of these three variations. We projected the mean vertex difference of each vertex using color encoding on the estimated model. As we can see, without gravity correction, most estimation errors occur in the thorax and abdominal region. Using supervised and unsupervised gravity correction, the estimation errors in the thorax and abdominal region were notably reduced. The mean vertex error without gravity correction is 17.5 ± 12.6 mm. Using supervised gravity correction, the mean vertex error is reduced to 16.6 ± 12.4 mm. The estimation error is further reduced to 15.3 ± 11.6 mm by using unsupervised gravity correction.

Subsequently we evaluate the internal landmark estimation using different model estimation methods. Therefore, we use the Anatomy3 dataset comprising 20 full-body CT scans with 50 associated internal landmarks in each dataset. An example of the dataset can be found in Fig. 3. This dataset was used to learn the landmark embedding weighting $\mathbf{w}_i$ for each landmark. We evaluated our algorithm with a leave-one-out cross-validation. The estimation error is given in mean $\pm$ standard deviation in mm.

We first compared the estimation error using different model estimation methods. For each estimation method, we trained the mapping error using locally linear embedding (LLE) and compared the mean estimation error. Without gravity correction, the mean estimation error of all landmarks is 16.72 ± 12.28 mm. The error reduced with supervised gravity correction to 14.95 ± 13.37 mm and further reduced with unsupervised gravity correction to 13.43 ± 12.00 mm.

Furthermore, we compared the estimation result using LLE, ridge regression (Ridge), lasso regression (Lasso) and elastic net regression (Elastic). The results are shown in Fig. 5. We can see from the results that in most of the cases the ridge regression has the worst performance with an overall estimation error of 22.19 ± 12.36 mm. The closed-form solution of the LLE method outperformed ridge regression and reduced the mean error to 13.43 ± 12.00 mm. Using the elastic regression method, both the standard deviation and the mean estimation error improved to 12.71 ± 7.32 mm. The best performance was achieved using the lasso regression. With this method the overall landmark estimation error was 12.19 ± 6.98 mm. We also performed a paired t-test to all of our methods of the estimation error in spine positions and other landmarks. We also performed statistically analysis of the performance differences between the methods using students t-test ($p < 0.05$) with Bonferroni correction. Therefore a p-value below 0.0083 was considered statistically significant. We found, that using the lasso method resulted in 38 landmarks in significantly smaller errors compared to the ridge regression. Using the lasso method resulted additionally compared to the elastic regression in a smaller error for three lumbar spine and one thoracic spine landmarks. Compared to the LLE method one landmark (trochanter major left) had significant smaller errors.

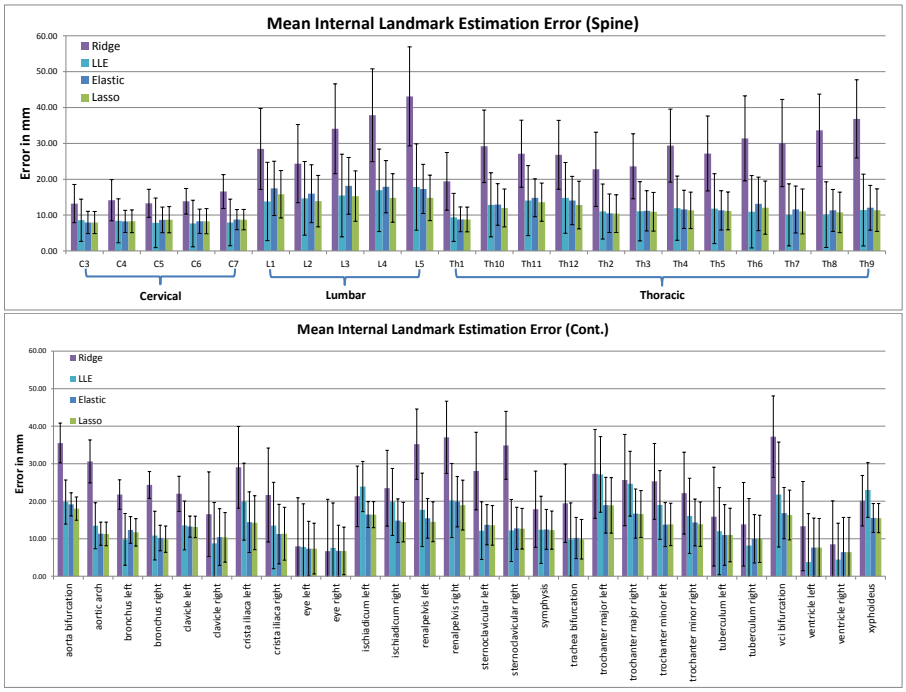

Fig. 5: Mean internal landmark estimation error using linear embedding (LLE), ridge regression (Ridge), elastic net regression (Elastic) and lasso regression (Lasso)

4 Discussion

The main contribution of this study is to introduce a pipeline for the training and estimation of internal anatomical landmarks based on a given patient surface model. The introduced pipeline takes the deformation caused by gravity as well as the positions of the arms of the patient for the landmark estimation into account. Therefore it allows to reliably estimate internal landmark positions of patients during interventional procedures, which may contribute to improving navigation and guidance during image-guided procedures. Considering that interventional X-ray imaging systems can have detectors covering an area of about 200 mm $\times$ 266 mm (20 cm $\times$ 27 cm) at iso-center, the achieved accuracy is sufficient to facilitate automatic positioning of the X-ray system.

The goal of our work was to locate anatomical landmarks inside a patient surface model based on surface vertices close to associated internal landmarks. We found that our approach based on a linear mapping relative to a template surface yielded good matches between predicted landmark positions and actual positions. The results also indicate that our template fitting method is robust as the accuracy for the template fitting is crucial for the subsequent internal landmark estimation. We can also see from the results that the ridge and LLE

methods have a higher estimation error than Elastic and Lasso regression. One of the reasons for this difference could be that by introducing the L1 norm as a regularizer, the Elastic and Lasso regression impose the sparsity of weighting vector $\mathbf{w}$. This might be an indication that we could improve our neighborhood selection method by including only those vertices which are significant for the estimation. In this work, we did not investigate non-linear mapping methods as the number of our datasets is still limited.

Some assumptions were made in the model implementation. For the registration we divided the mesh in eight districts. Theoretically, a minimum of two districts are needed for a plane symmetric mesh, four districts are needed for axisymmetric mesh, and eight districts are need for a point symmetric mesh in 3D space, assuming only very small rotation is present between two meshes. As the patient surface model is plane symmetric (left and right part of the body), axissymmetric (part in abdomen), and point symmetric (e.g. head), we use eight districts. Furthermore, we assumed a translation with limited rotation. As the patient is lying on his back during the CT scan and the image data of the scan is saved in a coordinate system, where the y-axis aligns with the table movement direction only small rotation is possible between different datasets. Also, the datasets we use comprises only patients lying on their back and we estimate our model in DICOM coordinate, therefore, we assume that a translation with limited rotation.

So far, we estimated internal landmarks independently and have not yet taken any constraints regarding their positions relative to each other into account, but this could be a rewarding next step as human anatomy follows certain rules. For certain groups of landmarks, e.g., the spine, an active shape model enforcing neighborhood constraints could potentially further improve the estimation results. We see our work of embedding internal landmarks into a surface model as adding prior information to our surface models. Once internal landmarks are found, it is possible to embed an appropriately adjusted spine model into the surface model. As most organs have a fixed position relative to the spine, fitting a spine model to a surface model can, for example, be used for organ-specific positioning of a robotic angiography system, such as the Artis pheno (Siemens Healthineers, Erlangen, Germany). In Fig. 6, we show an example of fitting an adapted spline model using the estimated internal landmarks. In future work, we will look into refining these initial (a priori) estimates for patients where pre-operative image data is available. In other words, we will try to calculate a posteriori internal landmark estimates based on our a-priori estimates.

Acknowledgements We gratefully acknowledge the support of Siemens Healthineers, Forchheim, Germany. We thank Siemens Corporate Technology for providing the avatar database. Note that the concepts and information presented in this paper are based on research, and they are not commercially available.

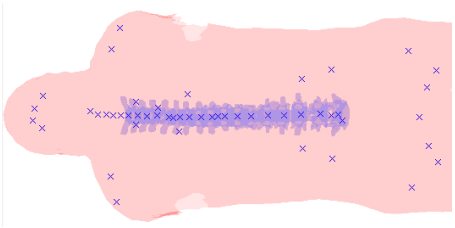

Fig. 6: An example demonstrates the estimated internal landmark of a patient surface model generated based on patient meta data such as height and weight. A spine model is placed according to the landmark position and shown as the overlay on the patient surface model.

Disclaimer

The authores X. Zhong, N. Strobel, A. Birkhold, M. Kowarschik, R. Fahrig, A. Maier declare that they have no conflict of interest. This article does not contain any studies with human participants performed by any of the authors.

References

1. Allen, B., Curless, B., Popović, Z.: The space of human body shapes: reconstruction and parameterization from range scans. In: ACM Trans Graph, vol. 22, pp. 587–594 (2003)
2. Anguelov, D., Koller, D., Pang, H.C., et al.: Recovering articulated object models from 3d range data. In: Uncertain Artif Intell, pp. 18–26 (2004)
3. Au, O.K.C., Tai, C.L., Chu, H.K., Cohen-Or, D., Lee, T.Y.: Skeleton extraction by mesh contraction. In: ACM Transactions on Graphics (TOG), vol. 27, p. 44. ACM (2008)
4. Bauer, S., Wasza, J., Haase, S., et al.: Multi-modal surface registration for markerless initial patient setup in radiation therapy using microsoft's kinect sensor. In: Proc IEEE Int Conf Comput Vis, pp. 1175–1181 (2011)
5. Bednarek, D.R., Barbarits, J., Rana, V.K., Nagaraja, S.P., Josan, M.S., Rudin, S.: Verification of the performance accuracy of a real-time skin-dose tracking system for interventional fluoroscopic procedures. In: Medical Imaging 2011: Physics of Medical Imaging, vol. 7961, p. 796127. International Society for Optics and Photonics (2011)
6. Bork, F., Barmaki, R., Eck, U., Yu, K., Sandor, C., Navab, N.: Empirical study of non-reversing magic mirrors for augmented reality anatomy learning. In: Mixed and Augmented Reality (ISMAR), 2017 IEEE International Symposium on, pp. 169–176. IEEE (2017)
7. Darom, T., Keller, Y.: Scale-invariant features for 3-d mesh models. pp. 2758–2769 (2012)
8. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM **24**(6), 381–395 (1981)
9. Fletcher, R.: Practical methods of optimization. John Wiley & Sons (2013)
10. Hoerl, A.E., Kennard, R.W.: Ridge regression: Biased estimation for nonorthogonal problems. Technometrics **12**(1), 55–67 (1970)
11. Johnson, P.B., Borrego, D., Balter, S., et al.: Skin dose mapping for fluoroscopically guided interventions. Med Phys **38**(10), 5490–5499 (2011)
12. Rausch, J., Maier, A., Fahrig, R., Choi, J.H., Hinshaw, W., Schebesch, F., Haase, S., Wasza, J., Hornegger, J., Riess, C.: Kinect-based correction of overexposure artifacts in knee imaging with c-arm ct systems. Journal of Biomedical Imaging **2016**, 1 (2016)

13. Robinette, K.M., Daanen, H., Paquet, E.: The caesar project: a 3-d surface anthropometry survey. In: 3-D Digital Imaging and Modeling, 1999. Proceedings. Second International Conference on, pp. 380–386. IEEE (1999)
14. Roweis, S.T., Saul, L.K.: Nonlinear dimensionality reduction by locally linear embedding. science **290**(5500), 2323–2326 (2000)
15. Sahillioğlu, Y., Yemez, Y.: Minimum-distortion isometric shape correspondence using em algorithm. IEEE Trans Pattern Anal Mach Intell **34**(11), 2203–2215 (2012)
16. Singh, V., Chang, Y., Ma, K., et al.: Estimating a patient surface model for optimizing the medical scanning workflow. In: Med Image Comput Comput Assist Interv, pp. 472–479 (2014)
17. Taubmann, O., Wasza, J., Forman, C., Fischer, P., Wetzl, J., Maier, A., Hornegger, J.: Prediction of respiration-induced internal 3-d deformation fields from dense external 3-d surface motion. 28th Trans Comput Assist Radiol Surg (CARS) pp. 33–34 (2014)
18. Tibshirani, R.: Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society. Series B (Methodological) pp. 267–288 (1996)
19. del Toro, O.A.J., Goksel, O., Menze, B., et al.: Visceral–visual concept extraction challenge in radiology: Isbi 2014 challenge organization. Proc VISCERAL Challenge at ISBI pp. 6–15 (2014)
20. Wasza, J., Fischer, P., Leutheuser, H., Oefner, T., Bert, C., Maier, A., Hornegger, J.: Real-time respiratory motion analysis using 4-d shape priors. IEEE Transactions on Biomedical Engineering **63**(3), 485–495 (2016)
21. Zhong, X., Strobel, N., Sanders, J., Kowarschik, M., Fahrig, R., Maier, A.: Generation of Personalized Computational Phantoms Using Only Patient Metadata. In: IEEE (ed.) 2017 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC) (2017)
22. Zou, H., Hastie, T.: Regularization and variable selection via the elastic net. Journal of the Royal Statistical Society: Series B (Statistical Methodology) **67**(2), 301–320 (2005)