Evaluating two Methods for Geometry Reconstruction from Sparse Surgical Navigation Data

Felix Ambellan[1], Alexander Tack[1], Dave Wilson[2], Carolyn Anglin[3], Hans Lamecker[1,4], Stefan Zachow[1]

[1]Zuse-Institute Berlin, Department for Visual Data Analysis, Berlin, Deutschland
[2]Dalhousie University, Halifax, B3H 4R2, Canada
[3]University of Calgary, Calgary, T2N 1N4, Canada
[4]1000shapes GmbH, Berlin, Deutschland

Kontakt: {ambellan, tack}@zib.de

Abstract

In this study we investigate methods for fitting a Statistical Shape Model (SSM) to intra-operatively acquired point cloud data from a surgical navigation system. We validate the fitted models against the pre-operatively acquired Magnetic Resonance Imaging (MRI) data from the same patients.

We consider a cohort of 10 patients who underwent navigated total knee arthroplasty. As part of the surgical protocol the patients' distal femurs were partially digitized. All patients had an MRI scan two months pre-operatively. The MRI data were manually segmented and the reconstructed bone surfaces used as ground truth against which the fit was compared. Two methods were used to fit the SSM to the data, based on (1) Iterative Closest Points (ICP) and (2) Gaussian Mixture Models (GMM).

For both approaches, the difference between model fit and ground truth surface averaged less than 1.7 mm and excellent correspondence with the distal femoral morphology can be demonstrated.

Keywords: Total Knee Arthroplasty, Sparse Geometry Reconstruction, Statistical Shape Models

1 Problem

Historically, the objective of total knee arthroplasty (TKA) was only to relieve the pain of severe osteoarthritis (OA), whereas functional expectations were low. With younger and more active patients receiving TKA, functional expectations are higher and relief of pain alone is no longer adequate. Despite the success of total knee arthroplasty, the reported overall rate of dissatisfaction is usually in the 20% range with functional satisfaction being even lower. Patient dissatisfaction may be partially related to a mismatch between the pre-operative shape of the distal femur and its shape post-operatively, either due to the shape of the femoral component or its positioning. To improve the post-operative functional satisfaction, patient-specific implant positioning and design could be a suitable strategy.
Statistical Shape Models (SSMs) are able to generate plausible shapes within a learned shape space of a certain anatomical structure through adjusting a sparse deformation model, the so-called shape modes. By adjusting the shape modes an SSM can be fit to given sample points yielding a complete model of the respective anatomy.
The objective of the present study was to evaluate two methods for fitting an SSM of a distal femur to the respective point cloud data collected during routine navigated total knee arthroplasty, and to validate it against MRI data for the same subjects.

2 Material and Methods

We used data of 10 patients (4 males, 6 females) acquired in a previous study [1]. All patients underwent navigated TKA and also had an MRI scan within two months pre-operatively (sagittal PDw turbo spin echo; 1.5 Tesla system by GE Healthcare, Chicago, Illinois, USA; TR/TE 3000/25 ms, slice thickness 2 mm, resolution 0.3125/0.3125 mm). The distal femora including the femoral cartilage were manually segmented by expert users to reconstruct 3D surfaces using the dedicated 3D geometry reconstruction and visualization software AmiraZIBEdition (Zuse-Institute Berlin, Germany) [2].

The Stryker® Precision Total Knee Arthroplasty Navigation System (Stryker® Corporation, Kalamazoo, MI, USA) was used for all TKA cases and anatomical landmarks were digitzed for the anterior cortex (106-241 points), distal medial and lateral femoral condyles (50 points each), posterior medial and lateral femoral condyles (35 points each), lateral and medial epicondyles and the trochlear center.

2.1 SSM adjustment to point cloud data

In order to recover the shape and pose (position and orientation) of the distal femur as accurately as possible, the shape and pose of an SSM of the distal femoral bone was adjusted to optimally match the given point cloud data. Throughout this section we refer to the following surface distance measures (with vector valued components marked in bold face):

Mean surface distance = $\frac{1}{N}\sum_{i=1}^{N} \min_{\boldsymbol{x} \in \boldsymbol{S}} ||\boldsymbol{y}_i - \boldsymbol{x}||_2$, (1)

Root Mean Square (RMS) surface distance = $\sqrt{\frac{1}{N}\sum_{i=1}^{N} (\min_{\boldsymbol{x} \in \boldsymbol{S}} ||\boldsymbol{y}_i - \boldsymbol{x}||_2)^2}$, (2)

Maximum surface distance = $\max_{\boldsymbol{y}_i} \min_{\boldsymbol{x} \in \boldsymbol{S}} ||\boldsymbol{y}_i - \boldsymbol{x}||_2$, (3)

where $\boldsymbol{S}$ is an arbitrary surface and $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_N$ are point coordinates in $\mathbb{R}^3$, representing either vertices of the surface model or point cloud elements, depending on whether a distance between two surfaces or between a surface and a point cloud is measured.

The Iterative Closest Point (ICP) [3] algorithm is often used to adjust an SSM to sparse points. However, since Gaussian Mixture Models (GMM) outperformed ICP in a recent study [4], we assessed the ability of fitting an SSM to navigation point cloud data for both algorithms. The resulting SSM fit was compared to the ground truth surface reconstructed from MRI data.

To approximate the pathological femoral shapes for the 10 patients with severe OA, we used an SSM consisting of 184 training cases, including male and female subjects with and without varus/valgus knee malalignment. Shape model creation and establishment of correspondence between shapes was done following [5] and thus can be found described there in more details. The explained variance by the number of shape modes is shown in Figure 1. 60 shape modes explain approximately 98.9% of a femur's geometric variation. A majority of subjects exhibit severe OA and no patient of this study is included in the SSM. To test the approximative power of the SSM, prior to fitting to the point cloud data, the fitting was adjusted to match the individual patients' anatomies from the MRI data as closely as possible. For the 10 patients of this study the averaged mean surface distance between SSM reconstruction and ground truth segmentation was 0.54 mm ± 0.2 mm (Table 1). Thus, the SSM was capable of accurately representing the patients' anatomy. An example is shown in Figure 2 illustrating the approximation error for case 3 (mean surface distance of 0.60 mm).

Tab. 1: Mean surface distance (± std. deviation) for each SSM fitted to ground truth segmentation.

Case 1	Case 2	Case 3	Case 4	Case 5	Case 6	Case 7	Case 8	Case 9	Case 10
0.56 ± 0.44	0.50 ± 0.43	0.60 ± 0.47	0.47 ± 0.38	0.53 ± 0.41	0.60 ± 0.47	0.63 ± 0.50	0.42 ± 0.33	0.51 ± 0.41	0.55 ± 0.45

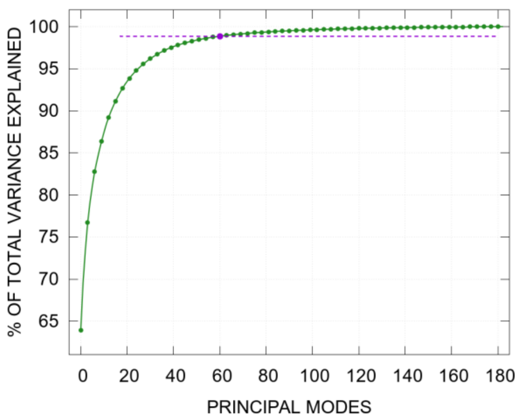

Fig. 2: Percentage of total variance explained with respect to the number of shape modes for the used SSM consisting of 184 training shapes.

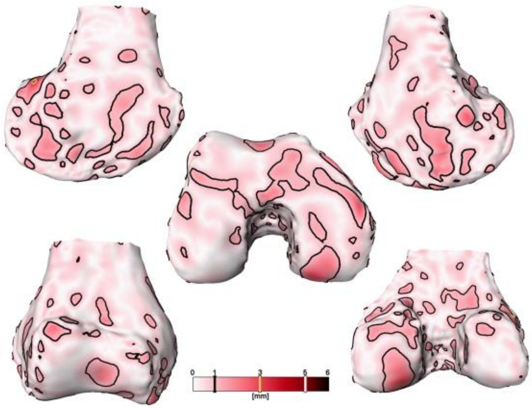

Fig. 2: Surface distance with isolines (1 mm, 3 mm and 5 mm) between the ground truth MRI surface of case 3 and an SSM consisting of 184 training instances fitted to that surface, showing that the patient-specific shape could be well-captured, before attempting to fit to the point cloud data. (The scale to 6 is provided for consistency with Figure 3; in this case, no isolines were greater than 1 mm)

2.1.1 Initial alignment of SSM to point cloud data

Since SSM and navigation point cloud data require an initial alignment, independently of the method used, a suitable transformation had to be found. To cover differences in position, orientation and scale, a similarity transformation with uniform scaling was calculated by minimizing the distance between three digitized anatomical landmarks (lateral and medial epicondyles and the femoral center) and the corresponding landmarks defined on the SSM.

2.1.2 Adjusting the SSM to point cloud data via an ICP approach

After initial alignment, an iterative optimization was carried out to adjust the SSM. For each point of the acquired cloud a closest vertex in the SSM is detected in every step. The closest vertices are collected in $\boldsymbol{v}$ and their 3D distance vectors to the point cloud in $\Delta\boldsymbol{v}$. Vertices of the SSM without matching point cloud data are not considered within the SSM fitting procedure. To minimize the distance (in mm) between the SSM and the point cloud, the following problem was solved by adjusting SSM-modes $\boldsymbol{p}_k$ through the real coefficients b_k and a rigid transformation with uniform scale T in an alternating fashion

$$\min_{b,T}\left(||(\boldsymbol{v}+\Delta\boldsymbol{v}) - T(\overline{\boldsymbol{v}} + \textstyle\sum_k b_k\boldsymbol{p}_k)||^2 + \gamma \cdot \sum_k \frac{{b_k}^2}{\lambda_k}\right), \tag{4}$$

with the mean shape's vertices $\bar{\boldsymbol{v}}$ and the eigenvalues λ_k of the SSM as well as a factor controlling regularization $\gamma \in \mathbb{R}_0^+$. This regularization penalizes shapes that are far away from the mean shape and thus favors configurations that are in a normal range of anatomical variation.

In a compromise between accuracy, robustness and with respect to the sparsity of the navigation data, of the 183 SSM-modes only the first third of the most significant modes were adjusted (approx. 98.9% variability). Between the SSM and the point cloud data the RMS surface distance was calculated, serving as a stopping criterion. The choice of regularization is important for the fitting accuracy. Studying the sensitivity with respect to this parameter showed that $\gamma < 0.05$ yields an anatomically incorrect fit. Choosing $\gamma > 2$ regularizes too strong towards the mean shape and hardly adjusts the SSM at all. Tests showed that regularization with $\gamma = 0.5$ improved the mean results slightly by approx. 0.05 mm compared to parameters between 0.1 and 1.

2.1.2 Adjusting the SSM to point cloud data via GMM

In a recent study a GMM-based algorithm outperformed the ICP algorithm for fitting an SSM to a sparse set of points [4]. That algorithm adjusted the modes of the SSM with a probabilistic approach considering anisotropic covariances, which are oriented according to the surface normal, however, without simultaneously optimizing the transformation. The expected amount of anisotropy is chosen with the parameter η, usually in the range of 2-8 (c.f. [4]). In this study, we used $\eta = 2$.

We deployed the aforementioned method to adjust the shape modes of our SSM. To optimize both, transformation and shape modes, in the GMM-based method the initial alignment and the adjustment of shape modes were executed in an iterative manner. Acknowledging the sparsity of the point cloud, we brought in a strong regularization towards the mean shape. Due to implementation details, this was achieved through scaling of the SSM-eigenvalues by a factor of 0.025 before applying the GMM-based fitting procedure in order to match the SSM to the sparse data. The GMM approach uses regularization similar to the ICP approach and this scaling is similar to a regularization with factor $\gamma = 40$ in the aforementioned ICP setup. However, the remaining objective functions of ICP and GMM cannot be compared directly. Again, we evaluated the sensitivity of fitting accuracy to the choice of regularization. Reasonable regularization was found for γ in the interval from 20 to 100. The chosen factor $\gamma = 40$ improved the mean results slightly by 0.02 mm compared to parameters in this interval.

2.2 Validation

The final SSM fits were further analysed in AmiraZIBEdition. The quality of the generated models was evaluated visually (Figure 3) and quantitatively with the help of mean surface distance, RMS surface distance, and maximum surface distance between the fitted SSM and the ground truth surface. Additionally, to evaluate the robustness of the proposed methods w.r.t. sparsity, we performed a series of tests on randomly reduced landmark sets. We always used the trochlear center and both epicondyles. The remaining landmarks were reduced to 100%, 75%, 50%, 25%, 15%, 5% and 0% of their original amount. Reduction was done in different regions (anterior cortex, distal medial and lateral femoral condyles, posterior medial and lateral femoral condyles) independently and three mentioned landmarks which are crucial for the fitting process were not considered for reduction and thus were included in every landmark set used for testing. To take randomness into account we performed five tests for every reduction level. The average number of landmarks per reduction level is given in Table 2.

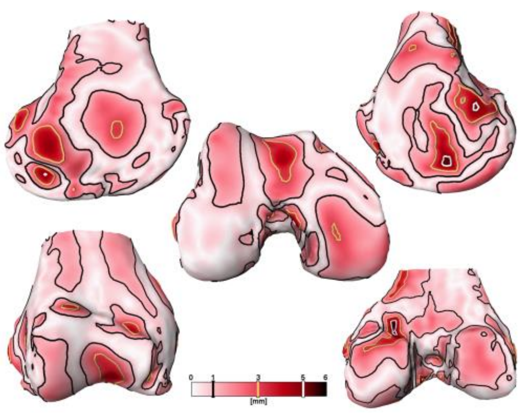

Fig. 3: Surface distance with isolines (1 mm, 3 mm and 5 mm) between the ground truth surface and the ICP-based fitted SSM to the point cloud data for case 3. (Mean surface distance = 1.20 mm)

Tab. 2: Average number of landmarks used for sparse fitting.

100%	75%	50%	25%	15%	5%	0%
352 ± 29	262 ± 22	176 ± 14	87 ± 7	53 ± 5	16 ± 2	3 ± 0

3 Results

Using the SSM generated femurs that were fit to the navigation data and then compared to the MRI ground truth, the ICP-based approach performed slightly better on average but not significantly (significance level: 0.01) better than the GMM approach with respect to the mean and RMS surface distance (Figure 4). Case 3 had a mean surface distance of 1.20 mm (ICP) and 1.37 mm (GMM), which is close to the mean surface distance averaged over all 10 cases (1.19 mm for ICP and 1.38 mm for GMM). Since case 3 in this way represents a typical result, its surface distance is shown in Figure 3.

Both methods achieve good results in the experiments with a reduced number of landmarks (Figure 5). We performed pairwise t-tests between the results using all landmarks (100%) and the results using a reduced number of landmarks (75%, 50%, 25%, 15%, 5% and 0%) for ICP and GMM, respectively. We found significant differences between 100% and 5% (ICP, GMM: $p < 0.001$) and between 100% and 0% (ICP, GMM: $p < 0.001$).

For the 10 patients the average computing time for reconstructing the anatomy was approx. 20 seconds (ICP) and ca. 25 seconds (GMM) on a standard personal computer (Intel Xeon E5-2650 v2; 8 cores; 2.60 GHz).

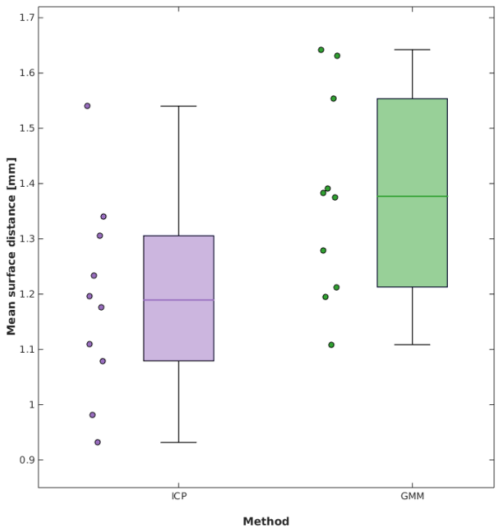

Fig. 4: Mean surface distance between the SSM fit to the point clouds and the ground truth surfaces. The plot showing mean value ± std. devivation with Whisker bars at min and max value.

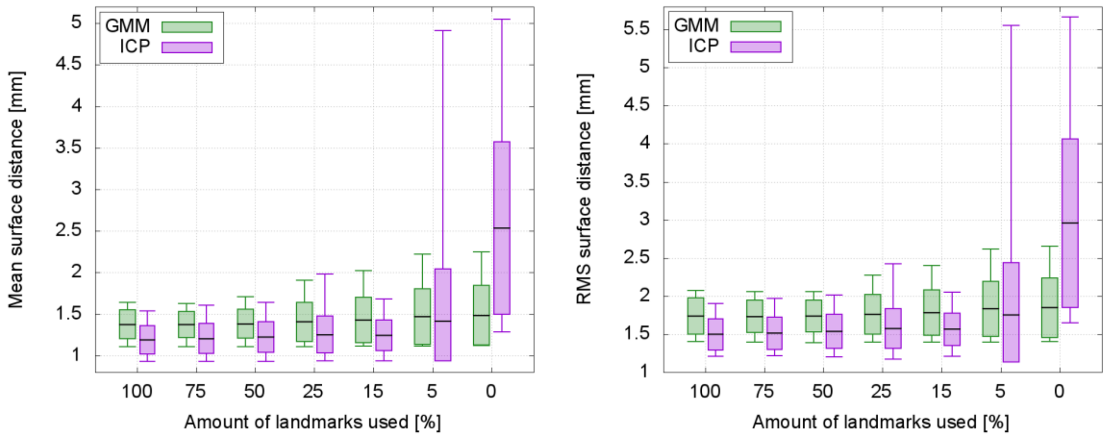

Fig. 5: Averaged Mean (left) and RMS surface distances (right) between the SSM fit to the point clouds and the ground truth surfaces for the two methods. For each reduction level (apart form 100% and 0%) five runs were performed. Boxplots showing mean value ± standard deviation with Whisker bars at minimal and maximal value.

4 Discussion

This study evaluated the accuracy of two methods for fitting an SSM to digitized surgical navigation data. As measured with the mean surface distance, RMS surface distance and maximum surface distance the errors between the two methods are similar, although the ICP-based approach performed slightly better on average than the GMM approach. We hypothesise that the reason for ICP perfoming slightly better is the challenging and sparse distribution of navigation landmarks. This should be subject to investigation in future studies. The two methods did not only achieve acceptable error measures, given that postoperative prosthesis vs. preoperative bone distances of up to 6 mm have been reported even in well-functioning patients [6], but the maximum errors occurred in less clinically-relevant areas such as the intercondylar notch and the superomedial edge of the articular surface. Beyond that, the error within the results stayed acceptable even for a reduced number of landmarks, showing the robustness of using an SSM for surface reconstruction.

One of the major limitations of this work is the accuracy of the MRI segmentation. Although we used this as our ground truth, there can be relatively large errors associated with using this technique. However, MRI is an excellent image modality for the assessment and segmentation of cartilage. It was attempted to only use areas of clinical interest to perform our comparison (e.g. the femoral condyles and trochlea); but due to difficulty in predictably determining the transition between condyle and osteophyte it was decided to use the entire distal femur despite the relatively large errors this resulting thereof.

Both methods of SSM fitting are fast enough to allow for intraoperative feedback with respect to computation time, although the implementations were not optimized for run time yet. Please note, that for a given point cloud the SSM fitting is deterministic for both approaches and uniquely determined.

As TKA evolves, a patient-specific approach will be demanded by patients. Patient-specific implant design and positioning could be a strategy to improve function post-operatively, whereby surgical navigation allows the surgeon to place the components with high precision in any position.

5 Conclusion

The results of this study show that even with the relatively sparse dataset available from routine navigated TKA, the SSM can provide a reasonably accurate approximation of the distal femur. These models can be used retrospectively to compare native anatomy with implant positioning, providing valuable insight into patient function and satisfaction. This could potentially allow for optimization of implant selection and position for a given patient and potentially improve patient satisfaction and function.

Acknowledgements

Parts of this work were also funded by the German Federal Ministry of Education and Research (BMBF), grant no. 01EC1406E (TOKMIS) and 01EC1408B (Overload-PrevOP). The authors would like to thank Robert Joachimsky (ZIB) and Agnieszka Putyra (ZIB) for assisting with the manual segmentations on which this work is based on.

References

[1] Wilson, Anglin, Ambellan, Grewe, Tack, Lamecker, Dunbar, Zachow, *Validation of three-dimensional models of the distal femur created from surgical navigation point cloud data for intraoperative and postoperative analysis of total knee arthroplasty*, International Journal of Computer Assisted Radiology and Surgery, (2017 – accepted for publication)

[2] AmiraZIBEdition, *https://amira.zib.de*, (2017)

[3] Besl, *A method for registration of 3D shapes*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 14 239-254 (1992)

[4] Bernard, Salamanca, Thunberg, Tack, Jentsch, Lamecker, Zachow, Hertel, Goncalves, Gemmar, *Shape-aware Surface Reconstruction from Sparse 3D Point-Clouds*, Medical Image Analysis, 38 77-89 (2017)

[5] Lamecker, *Variational and statistical shape modeling for 3D geometry reconstruction*, PhD Thesis, Freie Universität Berlin, Berlin (2008)

[6] Mahfouz, Abdel Fatah, Merkl, Mitchell, *Automatic and manual methodology for three-dimensional measurements of distal femoral gender differences and femoral component placement*, Journal of Knee Surgery, 22(4) 294-304 (2009)