Fully Automated Segmentation of the Knee using Local Deformation-Model Fitting

Matthias Amberg, Marcel Lüthi, and Thomas Vetter

Computer Science Department, University of Basel, Switzerland
{matthias.amberg, marcel.luethi, thomas.vetter}@unibas.ch

Abstract. In this paper we present a completely automated method for the segmentation of bone and cartilage in MR images of the knee. Our segmentation method is based on fitting a statistical deformation model, which was built from a set of labeled training images of the knee. The resulting deformation field is used to transfer a given segmentation from a reference image onto the target image to be segmented. To enlarge the flexibility of the deformation model in a controlled way, we use the idea of localized fitting, i.e. we fit the deformation model to local regions separately and combine the individual deformation fields.
We applied our procedure to a set of images that were provided for the MICCAI 2010 workshop "Medical Image Analysis for the Clinic". Our results show that our approach yields good results around the interface of tibia and femur. The results also reveal a number of problems that occur with a naive application of this method to the problem of knee segmentation. In particular it turns out that the quality of the segmentation deteriorates near the boundary of the images.

1 Introduction

The problem of cartilage segmentation from MR images of the knee has become increasingly important in recent years. In this paper we present a method for the fully automated segmentation of the knee from MR images. The work we present here is an entry to the MICCAI 2010 workshop *"Medical Image Analysis for the Clinic"*.

The segmentation of anatomical structures from medical images is an ill-posed problem. Usually the image does not provide enough information such that a clear distinction between the different tissue types could be achieved. Therefore prior information has to be incorporated. A commonly used and successful class of priors for segmentation are statistical shape and deformation models [1]. The problem with these model based approaches is, however, their lack of flexibility in case of insufficient training data. To overcome this problem we use a method proposed by Amberg et al. [2]. The main idea is to fit a deformation model locally to different regions in the image and to combine the individual fitting results. As each fit strives to minimize the error only locally, the result is more accurate in this region compared to a global fit. Nevertheless, as each individual fitting result satisfies the shape constraint, the combined result will in each region

correspond to the modeled shape. Thus, the resulting segmentation is resilient to noise, artifacts and missing data in the images. This property, however, also implies that pathological structures in the images, such as complicated lesions, cannot be segmented.

The main contribution of this paper is to present a fully automated workflow for model based segmentation, and to apply it to a number of test images of the knee, which were provided by the workshop organizers. Our procedure works as follows: We single out one of the training images as a reference. To build a statistical deformation model, we establish correspondence between the reference image and each of the training images, using the known labels. For segmentation of an image, we perform a fitting of the deformation model to the image to be segmented. The segmentation of the target image is obtained by transferring the labels given for the reference image with the resulting deformation field. We present segmentation results for 40 given test data-sets and discuss the strength and problems of the proposed method.

2 Segmentation method

Our procedure is divided into two stages: In the training phase, the statistical deformation model is built, using the label information given in the training images. In the segmentation stage, this deformation model is used to find valid deformations between the reference image and the target image to be segmented. Our workflow consists of the following steps:

Training:
- Choose an image of the training data-set as a reference image.
- Rigidly align the training images to this reference.
- Establish correspondence for each training image to the reference image.
- Build a deformation model from the resulting deformation fields.

Segmentation:
- Rigidly align the image with the reference image.
- Perform a (local) deformation model fitting.
- Transfer the reference label using the resulting deformation fields.

In the following, we will describe the steps of this workflow in more detail.

2.1 Rigid Alignment

All the images have to be aligned with the chosen reference image. Having a common alignment for the training and test data ensures that the deformations that are observed during the training phase are representative for the deformations observed in the testing phase. Furthermore, it can also serve as an initial alignment for the subsequent registration step.

Let $I_R : \Omega \to \mathbb{R}$ denote a reference image and let $I_T : \Omega \to \mathbb{R}$ be a target image with image domain Ω which is large enough to contain both images. To perform the rigid alignment, we translate the image I_T such that its center of

gravity $C[I_T]$, defined by $C[I] := \int_\Omega I(x)x\,dx$ is aligned with the center of gravity $C[I_R]$ of the reference image. Assuming that a standard protocol is used for the acquisition of the images, this already gives a good initial alignment. In a second step, we explicitly optimize for the translation $t \in \mathbb{R}^3$ and rotation $R \in \mathbb{R}^{3\times 3}$, such that the L_2 distance between the images is minimized. More precisely, we solve the optimization problem

$$\arg\min_{R,t} \int_\Omega (I_R(x) - I_T(Rx - t))^2\,dx \tag{1}$$

to obtain the final rigid alignment.

Note that we use the information of the whole image to compute the optimal transformation. This may lead to inaccurate alignments in cases where the angle between the tibia and the femur bone of the target image greatly differs from the one used in the reference. For our segmentation approach this is not a big problem, as long as similar bending angles of the knee are observed in the training and the test data. In other words, as long as the training images are representative for the images to be segmented, such alignment inaccuracies will be accounted for in the model.

2.2 Establishing correspondence

After the images are aligned, we perform a non-rigid registration of the training images. Here, we do not use the intensity images but the label-maps, which represent the different cartilage and bone regions. We compute additional feature images from these label-maps, which we will include in our registration algorithm as additional features to be matched.

Let L denote a label map, in which the individual structures to be segmented are given a non-zero label. We compute the signed distance image $D : \Omega \to \mathbb{R}$ which is defined by

$$D(x) := \begin{cases} \min_{\{x' : L(x') \neq 0\}} \|x - x'\| & \text{if } L(x) = 0 \\ -\min_{\{x' : L(x') = 0\}} \|x - x'\| & \text{otherwise .} \end{cases}$$

and a curvature image $C : \Omega \to \mathbb{R}$ defined by

$$C(x) := \Delta D(x).$$

Our registration-algorithm is an extension of the well known Demons algorithm [3] that takes these feature images into account. The standard Demons algorithm aims to find a deformation field $u : \Omega \to \mathbb{R}^3$ that minimizes the distance

$$\mathcal{D}[F_R, F_T, u] := \int_\Omega \frac{1}{\|\nabla F_R(x)\|^2 + F_R(x) - F_T(x + u(x))} (F_R(x) - F_T(x + u(x)))^2\,dx$$

between a reference image F_R and a target image F_T, subject to some smoothness constraints. We instead minimize the following distance term:

$$\mathcal{D}_E := w_L \mathcal{D}[L_R, L_T, u] + w_D \mathcal{D}[D_R, D_T, u] + w_C \mathcal{D}[C_R, C_T, u]$$

which takes all three feature images into account. Here $w_L, w_D, w_C \in \mathbb{R}$ are weights that determine the influence of the respective feature images. Our experiments showed that choosing $w_L = 5.0$, $w_D = 0.8$, and $w_C = 0.8$ leads to good registration results.

2.3 Statistical Deformation Models

Let $I_1, \ldots, I_n$ be n training images. For each training image I_i we obtain from the registration a deformation field u_i that relates image I_i with the reference image I_R. The idea behind a statistical deformation model is to model the space of deformations as the linear span of these examples. This is usually achieved by applying PCA to the (suitably discretized) example deformations [4]. A statistical deformation model can be written as a generative model defined by the parameter vector $\boldsymbol{\alpha} \in \mathbb{R}^n$:

$$\mathcal{M}[\boldsymbol{\alpha}] : \Omega \to \mathbb{R}^d \qquad x \mapsto \overline{u}(x) + \textstyle\sum_{i=1}^{n} \alpha_i \sqrt{\lambda_i} \phi_i(x) \tag{2}$$

where $\overline{u} : \Omega \to \mathbb{R}^3$ is the sample mean of the deformation fields, $\phi_i : \Omega \to \mathbb{R}^3, i = 1, \ldots, n$ are the principal components and $\lambda_i \in \mathbb{R}$ is the variance captured by the i-th principal component. The quantity $\|\boldsymbol{\alpha}\|^2$ is often taken to measure the likelihood of a given deformation $\mathcal{M}[\boldsymbol{\alpha}]$: The larger $\|\boldsymbol{\alpha}\|^2$, the less likely is the given deformation (see e.g. Lüthi et al. [5] for a precise probabilistic motivation of this term).

2.4 Localized fitting of deformation models

With the statistical deformation model built from the training data, we have now an explicit model of the deformations that relate typical knee images. Given a test image I^* that is aligned to the reference I_R, we can fit a deformation model, and thus find a deformation field $\mathcal{M}[\boldsymbol{\alpha}]$, which relates the reference image with the given target image. To find this deformation we minimize the L_2 distance between the reference and target image. In order to avoid deformations that are unlikely under the given model, we further penalize large values of $\|\boldsymbol{\alpha}\|^2$. This results in the following optimization problem:

$$\boldsymbol{\alpha}^* := \min_{\boldsymbol{\alpha} \in \mathbb{R}^n} \int_{\Omega} (I_R(x) - I^*(\mathcal{M}[\boldsymbol{\alpha}](x)))^2 \, dx + \mu \|\boldsymbol{\alpha}\|^2, \tag{3}$$

where $\mu \in \mathbb{R}$ is a regularization parameter. While conceptually simple, this approach has the problem that the model is often not flexible enough to accurately explain the true deformation for the full image. To alleviate this problem we use an idea inspired by the method of local regression, which is commonly used in statistics (see e.g. Hastie et al. [6]). The intuition behind our method is the following: We fit the model individually to different local regions of the image. As a local region is easier to explain than the full image, the local fits are very accurate for small enough regions. By combining these individual fitting results

we obtain a solution, which is accurate for all the individual regions, but still satisfies the model constraints locally.

We now make this idea more precise. Let $\mathcal{X} = \{x_1, \ldots, x_K\} \subset \Omega$ be a set of points, where each point $x_k \in \mathcal{X}$ marks the center of a region. To define the regions, we use a weighting function w_{x_k}, which assigns every point in the support of w_{x_k} a weight that depends on its distance to x_k. This can for example be achieved with the *Epanechnikov kernel* [6]:

$$w_{x_k}(x) := \kappa_\sigma(x_k, x) = \begin{cases} \frac{3}{4}[1 - (\frac{\|x - x_k\|}{\sigma})^2] & \text{if } \|x - x_k\|/\sigma \leq 1 \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

Here σ is a bandwidth parameter that determines the size of the support of w_{x_k}. For each point x_k we then perform a weighted fitting of the shape model:

$$\boldsymbol{\alpha}^*_{x_k} := \arg\min_{\boldsymbol{\alpha} \in \mathbb{R}^n} \int_\Omega w_{x_k}(x)(I_R(x) - I^*(\mathcal{M}[\boldsymbol{\alpha}](x)))^2 \, dx + \mu \|\boldsymbol{\alpha}\|^2. \tag{5}$$

Note that this procedure yields a full deformation field $u^*_{x_k} := \mathcal{M}[\boldsymbol{\alpha}^*_{x_k}]$ for any point x_k. As each solution is only accurate around the fitting point x_k we need to combine the individual solutions. Each deformation field $u^*_{x_k}$ should determine the final deformation in the region around the point x_k. We obtain a deformation field u^* with this property by defining:

$$u^*(x) := \frac{\sum_{k=1}^K \hat{w}_k(x) u^*_{x_k}(x)}{\sum_{k=1}^K \hat{w}_k(x)}, \; x \in \Omega. \tag{6}$$

As a weight function $\hat{w}_k$ we chose again the Epanechnikov kernel (4), with the same value for the bandwidth parameter as used for the local fitting (5).

2.5 Label transfer

Above procedures yields a deformation field u^* that relates the given test image with the reference. We can use u^* to transfer the labels L_R of the reference image to the target I^*. For this we need to compute the inverse deformation $[u^*]^{-1}$. The labels for the test images are then given by

$$L^*(x) := L_R(x + [u^*]^{-1}(x)), \; x \in \Omega. \tag{7}$$

Rather than computing this inverse mapping directly from the solution u^*, we invert each individual deformation field $u^*_{x_k}$ and combine them similar to Equation (6):

$$[\tilde{u}^*]^{-1}(x) := \frac{\sum_{k=1}^K \hat{w}_k(x) [u^*_{x_k}]^{-1}(x)}{\sum_{k=1}^K \hat{w}_k(x)}, \; x \in \Omega.$$

To compute the inverse deformations fields $[u^*_{x_k}]^{-1}$ we use a fixed point algorithm proposed by Chen et al. [7].[1]

[1] Our registration algorithm does not enforce that the deformation fields are invertible. If they are not, the fixed point scheme yields an approximate inverse. Our experience showed that this approximate inverse is nevertheless rather accurate, provided that the registration results are sufficiently smooth.

2.6 Tricks of the trade

During our experiments we noticed that a number of slight modification in above workflow makes the approach more robust. The first modification concerns the image domain over which we optimize Problems (1) and (5). Instead of optimizing over the full image domain Ω, better results can be obtained by solving the optimization problem only over a region $\Omega_M \subset \Omega$. This region is chosen slightly larger than the one labeled as bone in the reference image. This has the effect that the result is not influenced by image information far away from our region of interest. The second trick is to adapt the image intensities of the individual images. Even when the images are acquired according to the same protocol, the intensities slightly vary among different patients. Before performing the rigid alignment, we therefore align the image intensities using histogram matching. After the non-rigid registration has been performed, it even becomes possible to compute simple intensity statistics from the training images $I_1, \ldots, I_n$. It turns out that the fitting result is already greatly improved if we use an average image $I_{\text{AVG}}(x) := \frac{1}{n} \sum_{i=1}^{n} I_i(x + u_i^*(x))$, where u_i^* denotes the deformation field obtained from the registration step, instead of the reference image I_R in Equation (5). For a more detailed treatment and numerical results for these modifications, we refer the reader to the recent Master's thesis by Amberg [8].

3 Results

3.1 Experimental Setup

The main focus of our participation was to evaluate the performance and robustness of our model building and automated segmentation pipeline given a novel application scenario. Therefore, we did not focus on extensive parameter tuning. All training data was utilized for building the statistical model and no cross validation was conducted. Similarly, as a reference image, we chose an image that seemed to be a representative example for the data set, but we did not employ cross validation or similar techniques to choose it optimally.

For all our experiments, we used the same fixed set of parameters, which are summarized in Table 1. The implementation is done in ITK [9], where we extended the standard registration framework provided by this toolkit.

3.2 Results

We applied our algorithm to the test images given by the workshop organizers. With our current implementation, each segmentation takes approximately 10 minutes on a standard PC (8 cores, 2.3 GHz, 24 GB RAM). The running time increases linearly with the number of fitting points, as for each point the full deformation model is fitted. Each such fit takes around 30 seconds with our implementation.

We submitted 40 segmentation results to the challenge organizers for evaluation. Table 2 lists the different evaluation results. The cartilage results, evaluated

	Parameter	Value
Registration	Number of resolution levels	5
	Number of iterations per level	128
	Smoothness parameter	1
Model fitting	Number of model components	50
	Number of fitting points	29
	Point distribution	Uniformly at random on shape surface
	Regularization parameter μ	100
	Bandwidth parameter σ	30
	Optimizer	lbfgs

Table 1: Parameter settings used for all the experiments.

in the provided regions of interest, are considerable better than the bone segmentations, which suffer from boundary problems. Figure 1 shows sample slices of different segmentation results compared to the reference segmentation. The method successfully identifies the individual bones and cartilage regions and yields a rather accurate segmentation around the interface between tibia and femur. The most obvious problem is the bad segmentation near the upper and lower border of the image. This is on one hand due to the choice of the reference. It can happen that an image shows a larger region of the bone than the one depicted on the reference image. This implies that at these places no valid correspondence can be found, and therefore the deformations at these places are arbitrary. Further artifacts are introduced by the boundary conditions of our registration algorithm. To alleviate this problem, the images would have to be standardized to show the same region. Furthermore, a margin would have to be introduced around the images, such that boundary effects do not play any role. A second main problem is that the deformation does not adapt to all the fine structures shown in the image. This is a typical problem with statistical shape and deformation models. We claimed that this problem can be solved using localized fits. While this is indeed the case, it requires careful tuning of the parameters, in order to find the best trade-off between accuracy and robustness.

4 Conclusion

We presented segmentation results of our statistical shape model based segmentation pipeline. The complete process, consisting of initial alignment, establishing correspondence, statistical model building, fitting and label transfer was completely automated. The results show that with our model based approach automatic segmentation of the knee and cartilage becomes feasible. Yet, to reach the accuracy of an human expert, the algorithm would have to be better tailored for the given problem. For example the values of the parameters should be optimally chosen for the given data-sets. Furthermore, additional measures, such as using a full intensity model similar to Active Appearance Models [10], or employing a label fusion approach [11] may increase the segmentation accuracy.

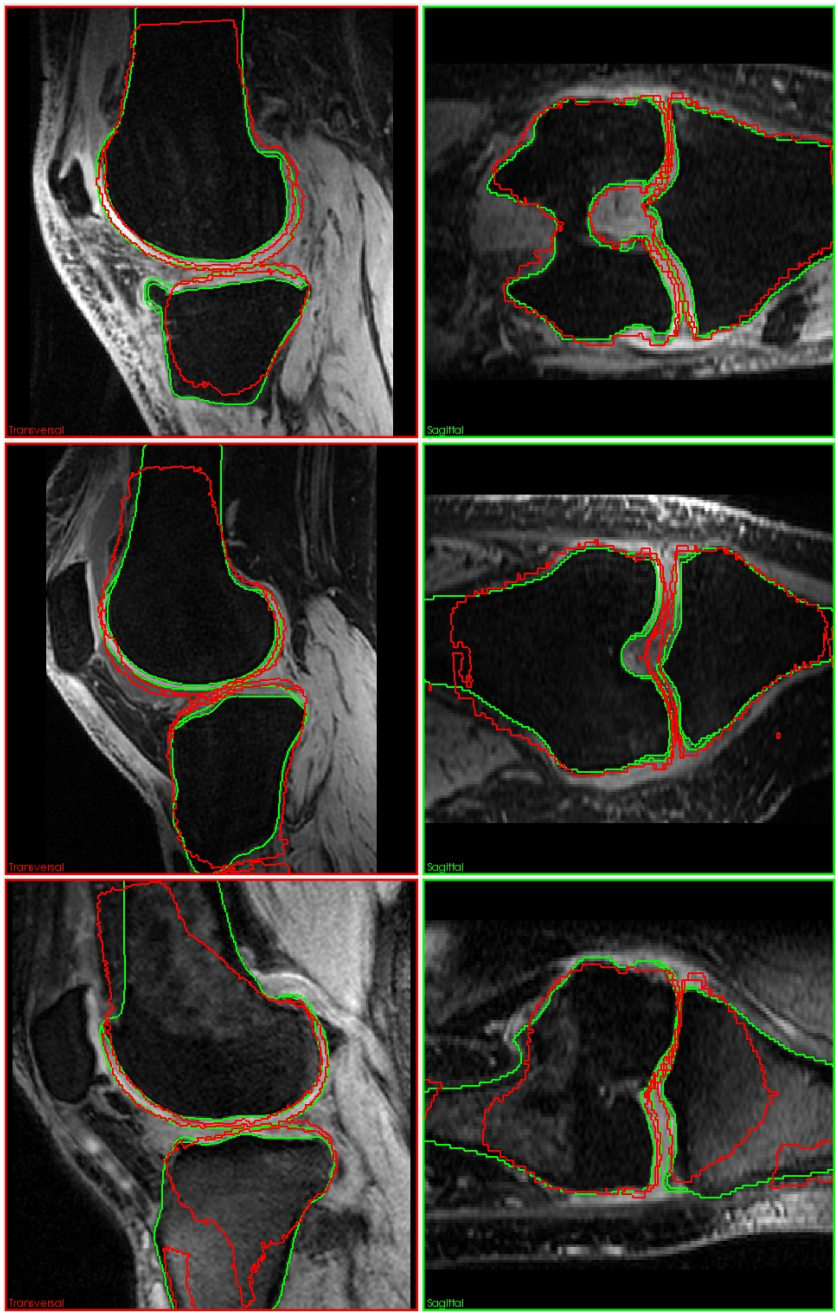

Fig. 1: Different views on selected test cases 8 (top), 16 (center), and 31 (bottom). The outline of the reference segmentation is displayed in green, the outline of the automatic method described in this paper in red.

The biggest problem that we faced is that our model degrades considerably at the image borders, therefore also decreasing the quality of our overall bone segmentation. While it is possible to introduce an artificial margin to reduce this effect, our model based approach is not ideal for cases where the structure to be segmented is not completely represented in the image. This scenario does not only lead to numerical problems, but also choosing a representative reference image becomes difficult.

Participating in this challenge provided us with a valuable opportunity to evaluate our method on novel data. While detail improvements are needed, the underlying approach of local deformation model fitting has shown to be promising for image segmentation.

Acknowledgements

This work has been supported by the CO-ME/NCCR research network of the Swiss National Science Foundation (http://co-me.ch).

References

1. Heimann, T., Meinzer, H.: Statistical shape models for 3D medical image segmentation: A review. Medical Image Analysis (2009)
2. Amberg, M., Lüthi, M., Vetter, T.: Local regression based statistical model fitting. In: Proceedings of the 32nd DAGM Symposium on Pattern Recognition. (2010)
3. Thirion, J.: Image matching as a diffusion process: an analogy with maxwell's demons. Medical Image Analysis **2**(3) (1998) 243260
4. Rückert, D., Frangi, A.F., Schnabel, J.A.: Automatic construction of 3-D statistical deformation models of the brain using nonrigid registration. Medical Imaging, IEEE Transactions on **22**(8) (2003) 1014–1025
5. Lüthi, M., Albrecht, T., Vetter, T.: Probabilistic Modeling and Visualization of the Flexibility in Morphable Models. Mathematics of Surfaces XIII (2009) 251–264
6. Hastie, T., Tibshirani, R., Friedman, J.: The Elements of Statistical Learning. Springer Series in Statistics. Springer New York Inc., New York, NY, USA (2001)
7. Chen, M., Lu, W., Chen, Q., Ruchala, K.J., Olivera, G.H.: A simple fixed-point approach to invert a deformation field. Medical Physics **35** (2008) 81
8. Amberg, M.: Automatic tooth segmentation using local model fitting. Master's thesis, Computer Science Department, University of Basel (2010)
9. Ibanez, L., Schroeder, W., Ng, L., Cates, J.: The ITK Software Guide. Kitware, Inc. (2005)
10. Cootes, T., Edwards, G., Taylor, C.: Active appearance models. Computer VisionECCV98 (1998) 484
11. Rohlfing, T., Brandt, R., Menzel, R., Maurer, C., et al.: Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains. NeuroImage **21**(4) (2004) 1428–1442

Img	Femur bone			Tibia bone			Fem. cartilage			Tibial cartilage			Total
	AvgD [mm]	RMSD [mm]	Scr	AvgD [mm]	RMSD [mm]	Scr	VOE [%]	VD [%]	Scr	VOE [%]	VD [%]	Scr	Score
1	1.23	1.94	34	1.25	2.02	17	53.5	-24.2	38	78.4	-53.8	21	27.7
2	1.32	2.35	25	2.53	4.24	0	58.9	-42.0	28	89.7	-65.8	17	17.7
3	1.33	1.95	31	1.14	2.09	19	47.5	-20.8	46	82.3	-37.8	20	29.2
4	0.95	1.50	49	0.96	1.50	37	62.6	-19.8	42	89.9	-60.9	17	36.5
5	1.17	1.68	40	1.20	1.90	21	46.5	-23.9	41	83.6	-50.2	19	30.5
6	1.39	2.44	22	2.03	4.19	0	55.6	-12.5	58	71.9	-53.1	24	25.8
7	1.40	2.15	26	1.08	1.65	30	56.8	-33.5	29	84.5	-41.1	19	26.2
8	1.52	2.59	16	1.30	1.98	16	56.2	-42.2	29	80.2	-61.7	21	20.5
9	1.10	1.69	42	2.25	3.93	0	54.2	-38.1	30	88.8	-54.5	18	22.4
10	1.54	2.50	17	1.47	2.85	0	59.2	-36.5	28	75.1	-56.8	23	16.9
11	1.30	2.14	29	1.11	1.80	26	41.8	-17.7	54	64.3	-39.8	27	33.9
12	1.21	2.04	33	0.94	1.59	36	52.3	-13.7	57	86.0	-32.0	19	36.1
13	1.13	1.78	40	1.87	2.61	0	58.5	5.6	69	71.9	-32.9	24	33.0
14	1.14	1.97	36	1.47	2.57	0	56.3	-12.0	58	81.2	-52.3	20	28.8
15	0.97	1.60	47	1.04	1.71	30	51.7	-8.1	67	73.5	-40.7	23	41.8
16	2.30	3.58	0	1.88	3.81	0	72.8	-4.6	65	92.5	-43.0	16	20.4
17	1.64	2.80	9	0.97	1.76	32	55.6	-5.3	70	78.6	-58.5	21	33.2
18	1.34	2.82	17	1.50	2.66	0	44.9	-25.1	39	64.7	-34.8	26	20.7
19	0.96	1.61	47	1.22	2.05	18	46.3	-12.2	62	72.7	-46.9	23	37.5
20	1.87	3.20	0	1.82	3.06	0	59.1	-25.9	33	71.5	-52.6	24	14.1
21	1.35	2.45	23	3.57	6.61	0	55.6	-22.1	41	80.1	-53.5	21	21.0
22	2.26	3.80	0	1.91	3.67	0	58.2	-23.0	38	80.5	-52.6	21	14.7
23	1.35	2.06	29	1.39	2.02	12	47.1	-22.0	44	88.5	-36.5	18	25.7
24	1.26	2.40	26	2.14	5.31	0	46.1	-13.8	59	73.0	-37.0	23	27.1
25	1.39	2.37	23	1.53	2.51	0	48.9	-33.2	32	88.5	-39.5	18	18.1
26	1.55	2.46	17	0.97	1.42	39	49.2	-31.4	32	82.8	-53.0	20	26.9
27	1.10	1.99	37	2.41	4.42	0	54.8	-16.3	51	89.6	-63.3	17	26.4
28	2.29	3.71	0	2.19	3.38	0	62.1	-6.4	66	88.7	-63.5	18	20.9
29	1.56	2.79	11	2.55	4.32	0	52.8	-35.8	31	92.7	-55.8	16	14.5
30	0.96	1.79	44	0.95	1.35	41	44.6	-23.2	43	72.5	-29.5	23	37.7
31	1.70	2.77	8	2.32	3.65	0	47.7	-24.8	39	88.6	-44.2	18	16.1
32	0.90	1.48	51	1.15	1.92	22	42.1	-2.6	80	70.5	-36.5	24	44.4
33	1.02	2.01	39	1.33	3.29	5	54.5	-10.8	61	69.6	-51.8	25	32.4
34	1.26	1.80	36	1.31	2.07	14	55.4	21.0	43	78.6	-34.4	21	28.5
35	1.54	2.49	17	2.18	3.22	0	65.9	-14.5	50	86.7	-59.9	18	21.4
36	1.53	3.06	8	2.02	4.29	0	44.8	-29.8	34	83.4	-52.0	20	15.3
37	1.64	2.42	15	1.91	2.90	0	51.7	-16.3	52	84.0	-34.2	19	21.7
38	1.32	2.48	23	1.03	2.09	23	45.8	-23.5	42	78.7	-47.0	21	27.3
39	2.09	2.96	2	5.35	7.73	0	65.5	-15.9	48	86.2	-41.8	18	17.1
40	1.24	2.18	30	1.42	2.38	4	46.6	-16.7	54	85.1	-20.3	33	30.2
Avg	1.40	2.35	25	1.72	2.96	11	53.3	-19.3	47	80.7	-46.9	21	26.0
	±0.36	±0.59	±15	±0.83	±1.41	±14	±7.0	±12.8	±14	±7.7	±11.0	±3	±7.9

Table 2: Results of the comparison metrics and scores for all 40 test cases. AvgD and RMSD are the average and RMS surface distance, respectively, VOE is the volumetric overlap error and VD indicates the volumetric difference.