Learning-Based Cost Functions for 3D and 4D Multi-Surface Multi-Object Segmentation of Knee MRI: Data from the Osteoarthritis Initiative

Satyananda Kashyap,
Dept. of Electrical and Computer Engineering, The University of Iowa, Iowa City IA 52242, USA

Honghai Zhang,
Dept. of Electrical and Computer Engineering, The University of Iowa, Iowa City IA 52242, USA

Karan Rao, and
University of Iowa Carver College of Medicine, Iowa City, IA 52242, USA

Milan Sonka [Fellow, IEEE]
Dept. of Electrical and Computer Engineering, The University of Iowa, Iowa City IA 52242, USA

Abstract

A fully automated knee MRI segmentation method to study osteoarthritis (OA) was developed using a novel hierarchical set of random forests (RF) classifiers to learn the appearance of cartilage regions and their boundaries. A neighborhood approximation forest is used first to provide contextual feature to the second-level RF classifier that also considers local features and produces location-specific costs for the layered optimal graph image segmentation of multiple objects and surfaces (LOGISMOS) framework. Double echo steady state (DESS) MRIs used in this work originated from the Osteoarthritis Initiative (OAI) study. Trained on 34 MRIs with varying degrees of OA, the performance of the learning-based method tested on 108 MRIs showed significant reduction in segmentation errors ($p<0.05$) compared with the conventional gradient-based and single-stage RF-learned costs. The 3D LOGISMOS was extended to longitudinal-3D (4D) to simultaneously segment multiple follow-up visits of the same patient. As such, data from all time-points of the temporal sequence contribute information to a single optimal solution that utilizes both spatial 3D and temporal contexts. 4D LOGISMOS validation on 108 MRIs from baseline and 12 month follow-up scans of 54 patients showed significant reduction in segmentation errors ($p<0.01$) compared to 3D. Finally, the potential of 4D LOGISMOS was further explored on the same 54 patients using 5 annual follow-up scans demonstrating a significant improvement of measuring cartilage thickness ($p<0.01$) compared to the sequential 3D approach.

Keywords

4D LOGISMOS; knee MRI; knee segmentation; osteoarthritis; random forests classifier; hierarchical random forests; sub-plates analysis; neighborhood approximation forests; just-enough interaction; longitudinal analysis; graph search

I. Introduction

Knee osteoarthritis (OA) is a leading cause of disability in the United States, affecting over 10% percent of the adult population, and accounting for over 600,000 knee replacements in excess of $10 billion per year [1], [2], [3], [4]. Although commonly characterized by joint pain and functional disability, OA is a heterogeneous disease with multiple risk factors and pathways contributing to variable disease symptoms and progression. The pathogenesis of OA is complex, with biomechanical joint stressors, inflammatory mediators, and structural changes to bone, ligaments, joint synovium, and menisci leading to articular cartilage degradation and loss [5], [6]. No FDA-approved drugs are available which can forestall cartilage degeneration. Imaging biomarkers that are sensitive to structural changes are needed for early detection of OA in an effort to provide useful insights to clinical trials investigating disease modifying drugs. Magnetic resonance imaging (MRI) of the knee provides a complete 3D view of the knee anatomy unlike traditional 2D x-rays. Moreover, changes in cartilage morphology are detected earlier in MRIs than x-rays [7]. In a longitudinal study by Raynauld et al. [8], MRI was used to reliably measure structural changes in arthritic knees and detect cartilage volume loss as early as 12 months from baseline. Balamoody et al. [9] demonstrated that similar MRI protocols across different 3T scanners gave comparable morphologic results, further suggesting that MR imaging is a dependable tool for quantitative cartilage morphology analysis.

For analysis of the cartilage structures, accurate segmentation is a crucial first step. In clinical research there is a need for reproducible fully or highly automated segmentation as it offers consistent accuracy and speed over manual segmentation efforts. However, automated knee segmentation is challenging due to the thinning cartilage, appearance of osteophytes, bone marrow and cartilage lesions and surface fibrillation in MRI data. Several of these disease artifacts appear similar to the cartilage further increasing the segmentation complexity (Fig. 8a).

Automated knee segmentation approaches previously reported in the literature were surveyed based on the employed methodology. Active shape and appearance models (ASM/AAM) were widely used in a variety of combinations such as deformable active shape models [10], [11], volumetric and surface based appearance models [12], semi-automated initialization of active shape models [13], and minimum descriptor length based group-wise image registration followed by active appearance model segmentation [14]. In the presence of pathology, active shape/active appearance models need to be flexible enough to properly represent many local changes of shape/appearance, thus requiring large training-set sizes and yielding high-dimensional models. The associated model-training and computational complexity is already substantial in 3D and may become impractical if 4D longitudinal analysis is considered. Multi-atlas based methods were used in several algorithms such as [15] with a non-local patch based fusion, Dam et al. [16] extended their multi-atlas based rigid registration by k-nearest neighbor voxel based classification in [17]. Multi-atlas registration followed by locally weighted voting was reported in [18], multi-atlas based registration followed by outlier detection to segment the tibio-femoral joint in [19]. Since multi-atlas methods optimize a match between the overall template labeling as well as the

individual labeled templates in the set, they may be sensitive to imperfect label templates in the presence of pathology. Both multi-atlas and ASM/AAM approaches may also suffer from converging to a locally rather than globally optimal solutions. Markov random fields were used to construct and optimize local image patches for the region and boundary probabilities from local shape and appearance information in [20]. Graph cuts optimization was applied to the output of a hierarchical two-stage classifier, which was trained to identify the cartilage and bone voxels [21]. As the underlying multi-label graph cuts jointly consider independent classifier outputs for cartilage, bone, and background, label-conflict-resolution may be challenging in regions with multiple labels. In Zhang et al. [22], four different MR acquisitions of the same patients were used to leverage the benefits of different sequence contrasts to model the spatial contextual information using support vector machines in a discrete random field framework. While a conceptually interesting approach, the need for multiple MR acquisitions limits the method's applicability.

Using cost functions designed as gradient-based features with heuristically determined weighting combination based on human expert knowledge (hand-tuned cost functions, Section II) limits the segmentation performance especially in the pathologic cases. Machine learning based methods can be used to identify the boundary properties of the anatomy of interest. Such learning-based cost function design has shown promise previously [23].

A major limitation of learning-based approaches is the time-consuming task of manually segmenting a large number of accurate examples used for training. Several post-processing interactive techniques were reported to ease the corrections, e.g., thin plate splines [24], live wire techniques [25] and active shape models [26] with the major issue being that they correct for the surfaces directly to match the object boundaries and are sometimes infeasible to be computed in close-to-real time.

We present a fully automated LOGISMOS segmentation algorithm that guarantees global optimality with respect to the cost functions provided. A novel hierarchical random forest (RF) classifiers was designed to provide the learned costs. An interactive correction method called just-enough interaction (JEI) [27], [28] was used to substantially reduce the interaction time needed to prepare the training data in comparison with voxel-by-voxel editing approaches. In our method, two variations of the RF classifier were used — a neighborhood approximation forests (NAF) followed by an RF classifier collected on a geometric graph. The novelty of this approach is that the NAF classifier gathers contextual and textural information from a global neighborhood of 3D image patches while the RF classifier collects local feature information along the geometric graph search columns. This combination enables training on contextual and textural features both locally and globally thus improving segmentation accuracy.

OA as a disease progresses slowly in the early stages with the rate of cartilage loss accelerating with advancing disease [29]. Every patient in the OAI was enrolled in the longitudinal study with follow-up MRI scans done at regular clinical visits. The availability of multi-time-point information can be leveraged to provide additional information proving beneficial in patients with progressively worsening OA. Using the spatial and temporal contextual information from the neighboring time-points of the same patient helps reducing

the inter-time-point variability ensuring that the cartilage losses are within physiologically feasible ranges. The LOGISMOS framework was extended to 4D to incorporate this information and simultaneously segment the multiple follow-up scans of the same patient. To the best of our knowledge, this is the first attempt at a 4D knee MRI segmentation algorithm incorporating the additional available information in a LOGISMOS framework.

The work reported here improves several aspects of the LOGISMOS algorithm for cartilage segmentation previously presented in [30]. Our novel hierarchical random forest classifier outperforms the previous method, in which a single RF classifier did not account for the regionally-specific appearance of the surrounding menisci, muscle, bone and other anatomies. As a result, certain intensity profiles appearing in normal-cartilage regions of the knee also occurred in other local regions in a pathological case. This ambiguity resulted in the imperfect training of the classifier and contributed to segmentation inaccuracy. Further, the LOGISMOS algorithm was extended to handle simultaneous multi-3D (4D) segmentation.

The remainder of the paper is organized as follows: Section I-A introduces the LOGISMOS algorithm framework and the 3D graph construction. In Section II our proposed hierarchical RF classifier design is described in detail followed by the 4D LOGISMOS extension in Section III. The experimental design, validation, and discussion follow in Sections IV, V and VI.

A. LOGISMOS Segmentation

LOGISMOS algorithm [30] solves the simultaneous segmentation of multiple objects and surfaces in a graph-based framework with guarantee of global optimality. To ensure global optimality, the LOGISMOS algorithm has to satisfy two important criteria: a) The geometric graph must have non-intersecting column construction and b) the graph edge construction must satisfy the self-closure property (see [31], [32] for proof). The focus of this paper is reporting a new method for cost function design and extending the 3D LOGISMOS to 4D. The workflow originally proposed in [30] is maintained.

The algorithm is initialized by identifying the volume of interest (VOI) which uses an AdaBoost classifier [33] trained on manually identified VOI's. The purpose of VOI detection was to localize smaller regions to reduce computation time. Further the VOI bounds are used for affine fitting of the mean femur and tibia bone shape mesh S_0. A patient specific bone mesh S is computed by a single surface single object LOGISMOS segmentation using S_0. This step is necessary since the final segmentation of the bone and the cartilage surfaces is defined by this prior.

A geometric graph with non-intersecting columns representing each surface (bone and cartilage of femur and tibia) is built by mimicking the behavior of charged particles resulting in electric lines of force (ELF) based columns. The cost functions designed are assigned to the graph nodes of every column. The graph-based multiple surfaces and objects segmentation edge construction and optimization are described in [30].

II. Cost function design

LOGISMOS guarantees global optimality of its solution with respect to the cost functions provided.

Simple hand-tuned cost functions were first designed to capture the bone to cartilage interface, which appeared as a strong dark to bright edge when traversing the image along the graph column from inside of the bone surface outwards to the bright cartilage. Directional 1D derivative operators gave costs based on the edges encountered along the search columns. For the cartilage, a weighted combination of the first and second derivative operators was used. This helps prevent interpreting cartilage inhomogeneities as edges. Although human expert designed cost functions are very effective at capturing the desired features there are limitations. 1) Choosing the correct weighting combination is challenging. 2) Same costs may not work for all parts of the anatomy. 3) The same anatomical objects (e.g., bone, cartilage) appear differently in pathological cases due to the loss of structure and/or the presence of lesions.

RF classifiers [34] were used to provide better and location-specific cost functions. RF classifiers use the concept of bagging, where for each decision tree a random subset of features is chosen thereby reducing inter-decision-tree correlation, which improves accuracy. A single RF classifier was used in [30] where the limitation was that all the information used was localized along the graph columns. Because of this locality, features were unable to capture information of the neighborhood that appeared larger than a few nodes along the column. There are several anatomical features that appear locally like a cartilage boundary but when examined in a global neighborhood may be combined with pathology (e.g., synovial fluid, Fig. 8). Further, the single RF classifier did not account for the regionally-specific appearance of the surrounding menisci, muscle, bone and other anatomies. As a result, certain intensity profiles appearing in a specific normal-cartilage region of the knee also occurred in another local region in a pathological case. This discrepancy resulted in the improper training of the classifier and contributed to segmentation inaccuracy.

In this work, a combination of global and local contextual features was explored. Two RF based classifiers working in hierarchy were used to learn image appearance properties in the cartilage regions (Fig. 4). The first stage used a NAF[1] [35] trained on example image patches. The output probability maps of the NAF were used with other image-based features for training each of the regionally specific second-stage RF classifiers. The second RF classifier collected features along the geometric graph columns. For the second RF classifier, the local regions of the knee were spatially clustered using *k*-means clustering to account for variable local anatomy appearances surrounding the cartilage (Fig. 2). Each of the spatially clustered regions was trained using a different RF classifier in the second stage. Disjoint training sets were used to help build a more realistic RF model based on actual NAF performance on unseen images.

---

[1]Based on code available at: http://www.nmr.mgh.harvard.edu/_enderk/software.html

A. Neighborhood Approximation Forests

NAF uses a random forest framework consisting of a several binary decision trees where each tree independently learns to predict the closest neighborhood. NAF uses the training set to learn which clusters of image patches have the most similar neighborhood structures based on a similarity criterion. The definition of similarity is application specific with the advantage of using a wide variety of distance-based metrics without modifying the core underlying approach. Upon training, NAF approximates the neighbors of an unseen test image based on this similarity criterion.

For each tree in the forest, the training phase starts with the root node and continues to add new branches. Each node branch is trained to learn a set of binary tests which progressively partition the image into subsets with respect to the user-defined distance metric. The node split is optimized such that the user-defined distance function yields the most compact partitioning. For every tree in the forest, a subset of the entire feature set is chosen for training thereby improving generalization and producing an independent prediction.

In this work, the NAF was trained on image patches where the pairwise distance function $\rho(I, J)$ captured the similarity of the image patches according to their segmentation labels. The distance function between training image patches were defined as $\rho(I, J) = \parallel seg(I) - seg(J) \parallel_{l_0}$ where $seg(.)$ is the segmentation label map for the corresponding image patch. The algorithm learned to group image patches that appeared similar to each other based on the segmentation similarity.

The test image was partitioned into several smaller patches. Each patch passed through a trained NAF classifier. The output probability of all patches were combined to produce the final probability map (Fig. 1). This was used as one of the inputs for the second clustered RF classifier.

B. Clustered Random Forest Classifier

The second RF classifier was trained on features collected at every node along the search columns of the geometric graph. Bone mesh surfaces that were corrected using just-enough interaction (see Appendix A) were used for geometric graph construction during training. Positive example labels corresponded to the nearest cartilage mesh intersection along each graph column. The different features collected at each graph node are shown in Table I with volumes of every feature computed individually. The value at each node point along the geometric ELF graph column was linearly interpolated from the closest point in the feature volume.

To handle the variability of cartilage image intensities and the regionally specific surrounding anatomy, a k-means clustering algorithm was applied on the coordinates of femur and tibia mean shape S_0 mesh points. The spatial parcellation of the pre-segmented mesh in 40 clusters each was empirically designed (total 80, Fig. 2). Having too many clusters resulted in a reduced set of graph columns yielding insufficient sets of training examples. Having only a small number of clusters resulted in larger more inhomogeneous sets while negatively affecting the quality of classification. Training used a separate RF

classifier for each of the clusters to learn variable cartilage appearance and consider surrounding anatomies.

The probability response to the features along the search nodes in the testing datasets provided the node costs for LOGISMOS segmentation.

III. 4D Longitudinal Segmentation of Knee MRI

Along with the proposed learning-based cost functions, it is beneficial to leverage the contextual information from multiple time-points to improve the overall segmentation accuracy. Sequential yearly follow-up MRIs of the same knee provide spatial and temporal contextual information which helps to reduce the inter-time-point variability while ensuring physiologically feasible losses of cartilage during osteoarthritis progression. The crucial first step in 4D segmentation is to register the pre-segmented mesh surfaces and the respective images across time-points to establish correspondences temporally between the similar regions of the knee in 4D. A rigid registration was used given that only slight changes in bone shape occur with disease progression. The same configuration is maintained across all time-points in terms of the geometric graph parameters and topology. Enforcing contextual information temporally is done by linking adjacent time-points using inter-time-point edges in the underlying 4D LOGISMOS graph.

A. Establishing Temporal Correspondences

After pre-segmentation of each of the time-points, iterative closest points (ICP) algorithm was used to register the pre-segmented mesh surfaces. With a large translational or rotational movement between the two-time-points, there is a tendency to mismatch the surfaces, i.e., femur matched to the tibia or vice versa if they were the closest in terms of the least squares optimization. To prevent the mismatch, a two step registration was employed. The first step used the femur mesh only for ICP registration. This transform matrix was applied on both the femur and tibia meshes. After the first transform ensured that femoral and tibial surfaces were reasonably aligned, the ICP registration was run again using both the femur and tibia points together as a unified point cloud to further refine the registration. This ensured that vertex to vertex correspondences and thereby column correspondences were established. Fig. 3 illustrates the establishment of correspondence between two time-points. The same two-step transformation matrices were applied to the entire images.

B. Incorporating Inter-Time Point Context

For the 4D segmentation, the datasets at each time-point have their respective 3D multi-surface multi-object constraints enforced as described earlier in [30]. In order to enforce the contextual constraints across adjacent time-points, inter-time-point edges are constructed connecting nodes in every column with the corresponding columns on the registered surfaces. Using the prior information we can set the different minimum and maximum permissible limits for the bones and cartilages respectively so that cartilage thickness changes are within a physiologically feasible range. The edges introduced define inter-time-point min and max (δ_{tmin}, δ_{tmax}) where the inter-time-point maximum allowed change

(δ_{tmax}) was set to 0.6 mm/year based on available clinical literature [29]. The minimum δ_{tmin} was set to zero.

The inter-time-point edges E^t between neighboring time-points t_1, t_2 are constructed for all nodes of the graph:

$$E^t = \{\langle v_i(t_1, n_{cart}, k), v_i(t_2, n_{cart}, k - \delta_{tmin})\rangle\} \cup \{\langle v_i(t_1, n_{cart}, k - \delta_{tmax}), v_i(t_2, n_{cart}, k)\rangle\}$$
$$\cup \{\langle v_i(t_1, n_{bone}, k), v_i(t_2, n_{bone}, k - \delta_{tmin})\rangle\} \cup \{\langle v_i(t_1, n_{bone}, k - \delta_{tmax}), v_i(t_2, n_{bone}, k)\rangle\},$$

where $v_i(t, n, k)$ represents the kth node on the ith column, the inter-time-point edges are constructed between nodes $v_i(t_1, n, k)$, $v_i(t_2, n, k)$ of longitudinally corresponding columns t_1, t_2 for every $n \in \{cart, bone\}$. After the graph edges and nodes have been constructed, the 4D LOGISMOS segmentation problem is solved by computing a minimum s-t cut in a derived edge weighted digraph [41].

IV. Experimental Methods

Knee MRI data used for evaluation originated from the Osteoarthritis Initiative (OAI) database available for public access[2]. All MR acquisitions used double echo steady state (DESS) pulse sequence with in-plane resolution of 0.36×0.36mm and a slice spacing of 0.7mm, resulting in image volumes with $384 \times 384 \times 160$ voxels. 88 datasets at baseline (BL) and 12-months follow-up (12M) scans (176 3D MRIs in total) were used for which manually defined independent standards are available via OAI. BL images of 34 patients were used for training the hierarchical RF classifiers. The 12M data of the same patients were excluded from the testing set given the similarity of appearance resulting in a testing set of 54 MRIs at BL and 12M respectively (total of 108 MRIs).

Fig. 4 shows the training workflow for the hierarchical RF classifiers. For bone surface segmentation, the initially employed gradient-based costs were very robust and remained unchanged. For learning the cartilage-surface costs, the 34 patients were divided into two training sets with 15 and 19 patients each that were used to train the NAF and the clustered RF classifier respectively. The OAI provided independent standards had the cartilage and bone boundaries only in the region of articular cartilage (Fig. 7a). However the LOGISMOS segmentation and further analysis required the entire bone structure. The datasets used for training the clustered RF classifier were first inspected and JEI edited (see Appendix A for details).

All image volumes were first LOGISMOS-segmented using gradient costs. The geometric graphs had 8006 and 8002 graph columns for the femur and tibia objects, respectively. The graph parameters are listed in Table II with the additional inter-time-point constraint set at 0.6 mm for the 4D LOGISMOS.

The NAF features were sampled over 15 datasets with 1521 samples per image patch. Due to the imbalance in the negative to positive labels ratio, a neighborhood around the cartilage

---

[2]https://oai.epi-ucsf.org/datarelease/

was marked as negative examples. A set of 200 trees with 40,000 image patches were used as inputs to the training. The second RF classifier was trained on 19 JEI corrected datasets with every node along the ELF search column collecting 30 features (Table I). A total of 80 RF classifiers (one per cluster) were trained with 800 trees per each forest.

OA is characterized by thickness losses in subregions of the surface. Analyzing the entire cartilage structure may not be sensitive to these localized and regionally smaller changes. Therefore a consensus based nomenclature was proposed by clinicians identifying regions of the cartilage most likely affected by OA [42]. The existing methodologies in the literature require manual initialization for the sub-plate analysis [43], [44]. We developed a fully automated nomenclature-compliant sub-plate analysis algorithm. The algorithm subdivided the segmented surface by automated identification of the trochlear notch anatomy and implicit cutting plane geometry (Appendix B). All of our experimental results were subplate analyzed and reported for the entire femur and tibia cartilage regions, central load-bearing lateral femur/tibia (cLF/cLT), and central load-bearing medial femur/tibia (cMF/cMT).

The experimental results are reported as signed and unsigned errors. The signed errors indicate the measurement bias with the positive errors denoting underestimation and the negative errors denoting overestimation with respect to the independent standard. The unsigned errors indicate the variability of the measurement. The error measurements were collected in regions for which the independent standard was available. The first experiment was to validate the accuracy of the different cost functions in 3D LOGISMOS segmentation. The costs by the proposed hierarchical RF classifier were compared against the gradient costs. To evaluate the benefits of adding the NAF stage, a single RF classifier was trained on the same 19 patient datasets excluding the NAF features while maintaining the same graph parameters as the hierarchical classifier system. The testing workflow is shown in Fig. 5. The second experiment simultaneously 4D segmented the 54 knees using the proposed hierarchical costs and compared against the 108 individually 3D LOGISMOS segmented results using the same segmentation cost functions. It is designed to reveal the benefits of adding contextual information between the adjacent time-points. Given that OA is a slowly progressing disease, the expectation was that there would not be a significant appearance of artifacts or thickness losses over a one year study period. Fig. 6 shows the flowchart highlighting the key differences between the 3D and 4D LOGISMOS pipelines.

To demonstrate the benefits of 4D analysis over a longer study period, the same 54 patient datasets with BL, 12M, 24M, 36M and 48M follow-up scans were used resulting in a total of 237 3D MRIs analyzed for the third experiment (out of 270 due to subjects missing some follow-up visits). Since OAI-provided independent standards were available only for BL and 12 M, a new independent standard was defined in house by an expert analyst using JEI approach (Appendix A). To highlight the improvements, we compared the maximum cartilage thickness errors for 3D and 4D across all time-points and determined statistical significance of their differences.

V. Results

Table III shows surface positioning errors for the bone segmentation with the signed errors being sub-voxel accurate. Since high bone-segmentation accuracy was already achieved using the gradient-based costs, a learning based cost function design was not necessary. Although we saw presence of bone lesions with progression of OA, the gradient-based cost was sufficiently robust. Thus all the different methodologies in both 3D and 4D LOGISMOS used the same gradient-based cost functions for the bone.

Catilage surface positioning errors (compared against independent standard) using the hierarchical classifier, gradient cost, and single-stage RF classifier are reported. Each resulting surface from the above methods were sub-plate analyzed. Table IV shows a significant reduction in signed and unsigned errors for all the femoral sub-plates ($p \ll 0.001$). Significant reduction in unsigned errors over the segmentation using gradient costs was seen for all the tibial sub-plates while the signed errors showed a significant reduction of errors over the gradient-based costs on all sub-plates except the central medial tibia (cMT). Comparison with the single-stage RF classifier saw significant reduction of unsigned errors in the whole tibia and central medial tibia (cMT) plates.

Fig. 7 qualitatively compares the segmentation accuracies between the gradient-based costs and the hierarchical classifier with respect to the independent standard. Both the femur and tibia are shown with their respective bone and cartilage segmentations showing good agreement between learning-based segmentation and the independent standard.

Table V compares signed and unsigned border positioning errors with respect to the independent standard for 3D and 4D LOGISMOS segmentation at BL and 12M follow-up respectively. As expected, the overall segmentation accuracy is almost the same between the two methods except for a few of the plate regions highlighted in bold showing statistically significant differences over 3D ($p < 0.05$).

Table VI shows the performance improvement obtained by the 4D LOGISMOS over 3D for a longer study period of five years (BL——48 M, 5 time-points). To demonstrate improvements in local accuracy, Table VI gives the unsigned errors of cartilage thickness averaged for three quantiles within 90–100% of the largest errors per patient combined for all available time-points for each sub-plate. These unsigned maximum thickness errors were statistically significantly smaller in each of cMF, cMT and cLT for 4D LOGISMOS compared with separate 3D analyses.

Fig. 8 qualitatively shows the improvement of 4D LOGISMOS over 3D. Note the lack of an obvious edge distinguishing the tibia and femur cartilage. Using the spatial and temporal contextual information from all the time-points, the 4D method is able to correctly position the cartilage for the tibia and femur despite the lack of image information locally.

VI. DISCUSSION AND CONCLUSION

A novel method for learning segmentation cost functions using hierarchical RF classifiers was reported and its performance was demonstrated in 3D and 4D LOGISMOS

segmentation of knee MR images. The preparation of the training data was done using our just-enough interaction (JEI) approach, which provided fast and accurate post-processing correction. To minimize the impact of the underlying automated LOGISMOS segmentation on the independent standard used for training and evaluation, a great level of flexibility was embedded in the graph construction so that the expert analyst could freely modify the surfaces as needed and thus create unbiased corrections of the knee surfaces.

Using the single RF classifier improved the segmentation accuracy over the initial gradient-based costs. In addition, using a hierarchy of NAF and RF stages showed further statistically significant improvement in segmentation performance. The previous LOGISMOS-based method [30] used a single-stage RF with the graph constructed in the image-voxel space and a subset of current features contributing to the cost function (Table I) and it placed third on the publicly available SKI10 challenge [45]. The method reported here shows a significant improvement. Further comparison with the various other methods would inevitably be imperfect due to the different testing datasets utilized. In the OAI cohort, our method achieved high accuracy in comparison with the publicly available independent standard. Training of the hierarchical RF classifier and the subsequent 3D/4D LOGISMOS analyses ran on a Linux system with CentOS-7.3, 128 GB RAM. Once trained, the average segmentation times were 34 and 22 minutes per time-point for individual 3D and simultaneous 4D LOGISMOS, respectively.

Leveraging the benefits of the available multi-time-point MRI datasets for the OA patients, we designed a 4D extension of the LOGISMOS segmentation framework. The added temporal inter-time-point contextual constraints helped constrain the changes of cartilage thickness to what was physiologically feasible. This benefit of using information from the previous time-points along with the new cost function design helped notably improve the segmentation accuracy. Further, the advantages of the 4D LOGISMOS were demonstrated on an OAI longitudinal patient dataset with 5 imaging sessions from BL to 48 month. Our sub-plate detection algorithm was used to study the accuracy of segmentation on the load-bearing regions of the knee and provided good performance insight. The robustness of the automated sub-plate division was demonstrated showing consistent measurements between different surfaces used to initialize the automated extraction of the sub-plate load-bearing regions.

The combination of LOGISMOS and learning-based methods has been very effective. Other learning-based techniques such as deep learning may further improve the performance once more ground-truth data become available to determine further benefits of such approaches. 4D LOGISMOS also offers a platform for future clinically-oriented work, such as investigating the rate of cartilage degradation and its impact on OA progression, and identifying cartilage thickness that correlates to disease severity and/or predicts the need for joint replacement. Assessing reproducibility of the reported method is left for future work and the OAI scan-rescan dataset is well-suited for this task. Patellar segmentation was not considered in the work reported here due to the limited validation material available for patellar segmentation development. Although patellar cartilage was not segmented, our early LOGISMOS approach [30] was designed to segment all three bones and cartilages of the knee joint simultaneously.

Acknowledgments

Appendix A

Just Enough Interaction

In this work, JEI was used to prepare data for training the proposed hierarchical RF classifier (Fig. 4) and to assess the benefits of 4D LOGISMOS over a longitudinal study period. JEI is a novel interaction technique that uses the graph based LOGISMOS framework. The user inspects the automated segmentation results and provides segmentation correction points that interact directly with the underlying graph framework. This strategy differs from traditional voxel-by-voxel editing by only requiring limited (i.e., just-enough) interactions to correct the automated LOGISMOS segmentation. The JEI interaction method still guarantees global optimality even with the cost-changes resulting from the JEI corrections.

Following automated LOGISMOS segmentation, the JEI editing was performed on the resulting surface as needed. A custom graphical user interface (GUI) was designed onto which the image volume, residual graph and the ELF geometric graph were loaded. The JEI workflow is shown in Fig. 9. The details of the JEI architecture and GUI are given in [28].

Appendix B

Automated Extraction of Load-Bearing Sub-Regions

Analyzing the whole cartilage structure may not be sensitive to cartilage losses that occur locally. Cartilage sub-regions have been identified as areas which bear the maximum stresses during motion with a consensus based nomenclature by the clinicians on the sub-regions of the femur and tibia that need to be analyzed separately [42]. Several techniques [43], [44] for identifying the sub-plates exist all of which require human interaction to provide an initialization for the sub-region analysis. We have developed a fully automated sub-plate detection algorithm that uses LOGISMOS surface meshes to define individual sub-plates.

A. Trochlear Notch Identification

The first step in the sub-plate extraction is the detection of the trochlear notch. The notch is at the base of a groove along which the patella (knee cap) slides over the femur (Fig. 10). The main anatomic feature exploited is anterior to posterior (AP) curvature of the groove of the femoral bone (Fig. 10b). The notch is at the base of the curvature before the bone ridge rises sharply. An implicit cutting plane was used to isolate the analysis region represented by $\hat{n}.p$ where $\hat{n}$ is the normal direction from the plane and p is a point on the plane. When isolating the groove regions using simple implicit plane cutting, a family of contours were

drawn along this surface. The sharp rise of the bone structure along the contour near the base of the ridge gives a large change in contour positioning value traversing in the AP direction. To increase robustness we find the positions of largest change on all of the closely positioned groove contour lines and average them to find the desired trochlear notch (Fig. 11).

B. Sub-Plate Detection using Implicit Cutting Plane Geometry

Using implicit cutting planes at the trochlear notch separates the posterior region from the anterior and further splits them into the medial and lateral condyle. For each femoral condyle, the load-bearing regions is defined as 60% of the distance in the AP direction from the trochlear notch to the posterior-most point of the respective condyles. Isolating them using cutting planes normal to the AP plane at the 60% region isolates the load-bearing regions of the central medial and lateral femur (Fig. 11b).

The tibia is divided into the medial and lateral compartments using cutting planes positioned at the trochlear notch with its normal perpendicular to the AP plane (Fig. 12). After subdividing the plates into the medial and lateral regions, we isolate the central tibia and the peripheral sub-regions. The central 20% elliptical area of the medial and the lateral plate is computed around the center of mass of each respective plate. The radius of the major and the minor axis is computed as a ratio of bounds of the respective medial/lateral compartments. The major axis radius is computed from the compartment bounds along the AP direction and similarly, the minor axis radius is computed from the ratio of the compartment bounds perpendicular to the AP direction. Furthermore, the remaining regions are isolated using a 45° and 135° cutting planes around the center of mass of their respective compartments to give all four peripheral sub-regions as shown in Fig. 12. The final list of all the extracted sub-plates from the automated algorithm is visually shown in Fig. 13.

C. Robustness of Sub-plate Detection

Automated identification of sub-plate regions uses multi-surface knee joint segmentation as input. Consequently, the sensitivity of such surface segmentation on sub-plate region identification and thus on cartilage thickness measurement must be established. Load-bearing sub-plates were derived from two surface segmentation results (4D LOGISMOS, JEI-defined surfaces) and used to determine sub-plate-specific cartilage thickness from the manually segmented independent standard. In 108 patients, the two sets of sub-plate-specific cartilage thickness measurements were compared using regression analysis, which showed high robustness (low sensitivity) of the sub-plate detection algorithm on the identification of the load-bearing regions (using manually segmented independent standard surface) with respect to the underlying knee-joint segmentation. The sub-plate-specific correlation of the two thickness measurements ranged from $R^2 = 0.98$ to 0.99.

References

1. Zhang Y, Jordan J. Epidemiology of osteoarthritis. Clinics in Geriatric Medicine. 2010; 26(3):355–369. [PubMed: 20699159]
2. Murphy L, Helmick C. The impact of osteoarthritis in the united states: A population-health perspective. American Journal of Nursing. 2012; 112(3):S13–S19. [PubMed: 22373741]

3. Cram P, Lu X, Kates S, Singh J, Li Y, Wolf B. Total knee arthroplasty volume, utilization, and outcomes among medicare beneficiaries, 1991–2010. Journal of American Medical Association. 2012; 308(12):1227–1236.

4. Kremers HM, Larson DR, Crowson CS, Kremers WK, Washington RE, Steiner CA, Jiranek WA, Berry DJ. Prevalence of total hip and knee replacement in the united states. Journal of Bone and Joint Surgery. 2015; 97(17):1386–1397. [PubMed: 26333733]

5. Robinson W, Lepus C, Wang Q, Raghu H, Mao R, Lindstrom T, Sokolove J. Low-grade inflammation as a key mediator of the pathogenesis of osteoarthritis. Nature Reviews Rheumatology. 2016; 12(10):580–592. [PubMed: 27539668]

6. Lane N, Brandt K, Hawker G, Peeva E, Schreyer E, Tsuji W, Hochberg M. Oarsi-fda initiative: defining the disease state of osteoarthritis. Osteoarthritis Cartilage. 2011; 19(5):478–482. [PubMed: 21396464]

7. Sharma L, Eckstein F, Song J, Guermazi A, Prasad P, Kapoor D, Cahue S, Marshall M, Hudelmaier M, Dunlop D. Relationship of meniscal damage, meniscal extrusion, malalignment, and joint laxity to subsequent cartilage loss in osteoarthritic knees. Arthritis and Rheumatism. Jun; 2008 58(6): 1716–26. [PubMed: 18512777]

8. Raynauld J-P, Martel-Pelletier J, Berthiaume M-J, Beaudoin G, Choquette D, Haraoui B, Tannenbaum H, Meyer JM, Beary JF, Cline GA, Pelletier J-P. Long term evaluation of disease progression through the quantitative magnetic resonance imaging of symptomatic knee osteoarthritis patients: correlation with clinical symptoms and radiographic changes. Arthritis research & therapy. 2006; 8(1):R21. [PubMed: 16507119]

9. Balamoody S, Williams TG, Waterton JC, Bowes M, Hodgson R, Taylor CJ, Hutchinson CE. Comparison of 3T MR scanners in regional cartilage-thickness analysis in osteoarthritis: a cross-sectional multicenter, multivendor study. Arthritis Research & Therapy. 2010; 12(5):R202. [PubMed: 21029439]

10. Fripp J, Crozier S, Warfield SK, Ourselin S. Automatic segmentation and quantitative analysis of the articular cartilages from magnetic resonance images of the knee. IEEE Trans. Med. Imag. Jan; 2010 29(1):55–64.

11. Fripp J, Crozier S, Warfield SK, Ourselin SS. Automatic segmentation of the bone and extraction of the bone & cartilage interface from magnetic resonance images of the knee. Physics in medicine and biology. Mar; 2007 52(6):1617–1631. [PubMed: 17327652]

12. Williams, TG., Vincent, G., Bowes, M., Cootes, T., Balamoody, S., Hutchinson, C., Waterton, JC., Taylor, CJ. 2010 7th IEEE International Symposium on Biomedical Imaging: From Nano to Macro, ISBI 2010 - Proceedings. IEEE; 2010. Automatic segmentation of bones and inter-image anatomical correspondence by volumetric statistical modelling of knee MRI; p. 432-435.

13. Lynch JA, Zaim S, Zhao J, Peterfy CG, Genant HK. Automating measurement of subtle changes in articular cartilage from MRI of the knee by combining 3D image registration and segmentation. Medical Imaging 2001. Jul.2001 4322:431–439. [Online]. Available: http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=906344.

14. Vincent G, Wolstenholme C, Scott I, Bowes M. Fully Automatic Segmentation of the Knee Joint using Active Appearance Models. Medical Image Analysis for the Clinic: A Grand Challenge. 2010:224–230.

15. Shan L, Zach C, Charles C, Niethammer M. Automatic atlas-based three-label cartilage segmentation from MR knee images. Medical Image Analysis. Oct; 2014 18(7):1233–1246. [PubMed: 25128683]

16. Dam EB, Lillholm M, Marques J, Nielsen M. Automatic Segmentation of High- and Low-Field Knee MRIs Using Knee Image Quantification with Data from the Osteoarthritis Initiative. Journal of Medical Imaging. Apr; 2015 2(2):1–13.

17. Folkesson J, Dam B E, Olsen Fogh O, Pettersen C P, Christiansen C. Segmenting Articular Cartilage Automatically Using a Voxel Classification Approach. IEEE Trans. Med. Imag. 2007; 26(1):106–115.

18. Lee JG, Gumus S, Moon CH, Kwoh CK, Bae KT. Fully automated segmentation of cartilage from the MR images of knee using a multi-atlas and local structural analysis method. Medical Physics. 2014; 41(9)

19. Tamez-Pena JG, Farber J, Gonz??lez PC, Schreyer E, Schneider E, Totterman S. Unsupervised segmentation and quantification of anatomical knee features: Data from the osteoarthritis initiative. IEEE Transactions on Biomedical Engineering. Apr; 2012 59(4):1177–1186. [PubMed: 22318477]

20. Lee S, Park SH, Shim H, Yun ID, Lee SU. Optimization of local shape and appearance probabilities for segmentation of knee cartilage in 3-D MR images. Computer Vision and Image Understanding. Dec; 2011 115(12):1710–1720.

21. Wang, Q., Wu, D., Lu, L., Liu, M., Boyer, KL., Zhou, SK. Semantic Context Forests for Learning-Based Knee Cartilage Segmentation in 3D MR Images. Cham: Springer International Publishing; 2014. p. 105-115.

22. Zhang K, Lu W, Marziliano P. Automatic knee cartilage segmentation from multi-contrast MR images using support vector machine classification with spatial dependencies. Magnetic Resonance Imaging. Dec; 2013 31(10):1731–1743. [PubMed: 23867282]

23. Brejl M, Sonka M. Object localization and border detection criteria design in edge-based image segmentation: Automated learning from examples. IEEE Transactions on Medical Imaging. 2000; 19:973–985. [PubMed: 11131495]

24. Ross JC, Estepar RSJ, Diaz A, Westin C-F, Kikinis R, Silverman EK, Washko GR. Lung extraction, Lobe Segmentation and Hierarchical Region Assessment for Quantitative Analysis on High Resolution CT Images. MICCAI. 2009:690–698. [PubMed: 20426172]

25. Schenk A, Prause G, Peitgen H-O. Efficient Semiautomatic Segmentation of 3D Objects in Medical Images. MICCAI. 2000:186–195.

26. Schwarz T, Heimann T, Tetzlaff R, Rau A-M, Wolf I, Meinzer H-P. Interactive Surface Correction for 3D Shape Based Segmentation. SPIE Symp. Med. Imag. 2008

27. Sun S, Sonka M, Beichel RR. Graph-Based IVUS Segmentation with Efficient Computer-Aided Refinement. IEEE Trans. Med. Imag. Aug.2013 32(8)

28. Kashyap, S., Zhang, H., Sonka, M. An Just Enough Interaction Approach to Knee MRI Segmentation: Data from the Osteoarthritis Initiative; Interactive Medical Image Computing: 3th International Workshop, IMIC 2016, Held in Conjunction with MICCAI 2016; Athens, Greece. October 2016; 2016.

29. Eckstein F, Boudreau RM, Wang Z, Hannon MJ, Wirth W, Cotofana S, Guermazi A, Roemer F, Nevitt M, John MR, Ladel C, Sharma L, Hunter DJ, Kwoh CK. Trajectory of cartilage loss within 4 years of knee replacement a nested casecontrol study from the Osteoarthritis Initiative. Osteoarthritis and Cartilage. Oct; 2014 22(10):1542–1549. [PubMed: 24792212]

30. Yin Y, Zhang X, Williams R, Wu X, Anderson DD, Sonka M. LOGISMOS —Layered Optimal Graph Image Segmentation of Multiple Objects and Surfaces: Cartilage Segmentation in the Knee Joint. IEEE Trans. Med. Imag. 2010; 29(12):2023–2037.

31. Wu, X., Chen, DZ., Li, K., Sonka, M. The Layered Net Surface Problems in Discrete Geometry and Medical Image Segmentation. Berlin, Heidelberg: Springer Berlin Heidelberg; 2005. p. 17-27.

32. Li K, Wu X, Chen DZ, Sonka M. Optimal Surface Segmentation in Volumetric Images – A Graph-Theoretic Approach. IEEE Trans. Pattern Anal. Mach. Intell. 2006; 28(1):119–134. [PubMed: 16402624]

33. Freund Y, Schapire RE. A decision theoretic generalization of on-line learning and an application to boosting. Computer Systems Science. Aug; 1997 57(1):119–139.

34. Breiman L. Random forests. Machine learning. 2001; 45(1):5–32.

35. Konukoglu E, Glocker B, Zikic D, Criminisi A. Neighbourhood approximation using randomized forests. Medical image analysis. 2013; 17(7):790–804. [PubMed: 23725639]

36. Frangi, AF., Niessen, WJ., Vincken, KL., Viergever, MA. International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer; 1998. Multiscale vessel enhancement filtering; p. 130-137.

37. Zhao F, Zhang H, Wahle A, Thomas MT, Stolpen AH, Scholz TD, Sonka M. Congenital aortic disease: 4d magnetic resonance segmentation and quantitative analysis. Medical image analysis. 2009; 13(3):483–493. [PubMed: 19303351]

38. Daugman JG. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. JOSA A. 1985; 2(7):1160–1169.

39. Antonie, M-L., Zaiane, OR., Coman, A. Proceedings of the Second International Conference on Multimedia Data Mining. Springer-Verlag; 2001. Application of data mining techniques for medical image classification; p. 94-101.

40. Viola P, Jones MJ. Robust real-time face detection. International journal of computer vision. 2004; 57(2):137–154.

41. Boykov Y, Kolmogorov V. An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. IEEE Trans. Pattern Anal. Mach. Intell. Sep; 2004 26(9):1124–37. [PubMed: 15742889]

42. Eckstein F, Ateshian G, Burgkart R, Burstein D, Cicuttini F, Dardzinski B, Gray M, Link TM, Majumdar S, Mosher T, Peterfy C, Totterman S, Waterton J, Winalski CS, Felson D. Proposal for a Nomenclature for Magnetic Resonance Imaging Based Measures of Articular Cartilage in Osteoarthritis. Osteoarthritis and cartilage. Oct; 2006 14(10):974–83. [PubMed: 16730462]

43. Wirth W, Eckstein F. A Technique for Regional Analysis of Femorotibial Cartilage Thickness Based on Quantitative Magnetic Resonance Imaging. IEEE Trans. Med. Imag. Jun; 2008 27(6): 737–44.

44. Wirth W, Benichou O, Kwoh CK, Guermazi A, Hunter D, Putz R, Eckstein F. Spatial patterns of cartilage loss in the medial femoral condyle in osteoarthritic knees: Data from the osteoarthritis initiative. Magnetic resonance in medicine. Mar; 2010 63(3):574–81. [PubMed: 20146227]

45. Heimann T, Morrison BJ, Styner MA, Niethammer M, Warfield S. Segmentation of knee images: A grand challenge. Proc. MICCAI Workshop on Medical Image Analysis for the Clinic. 2010:207–214.
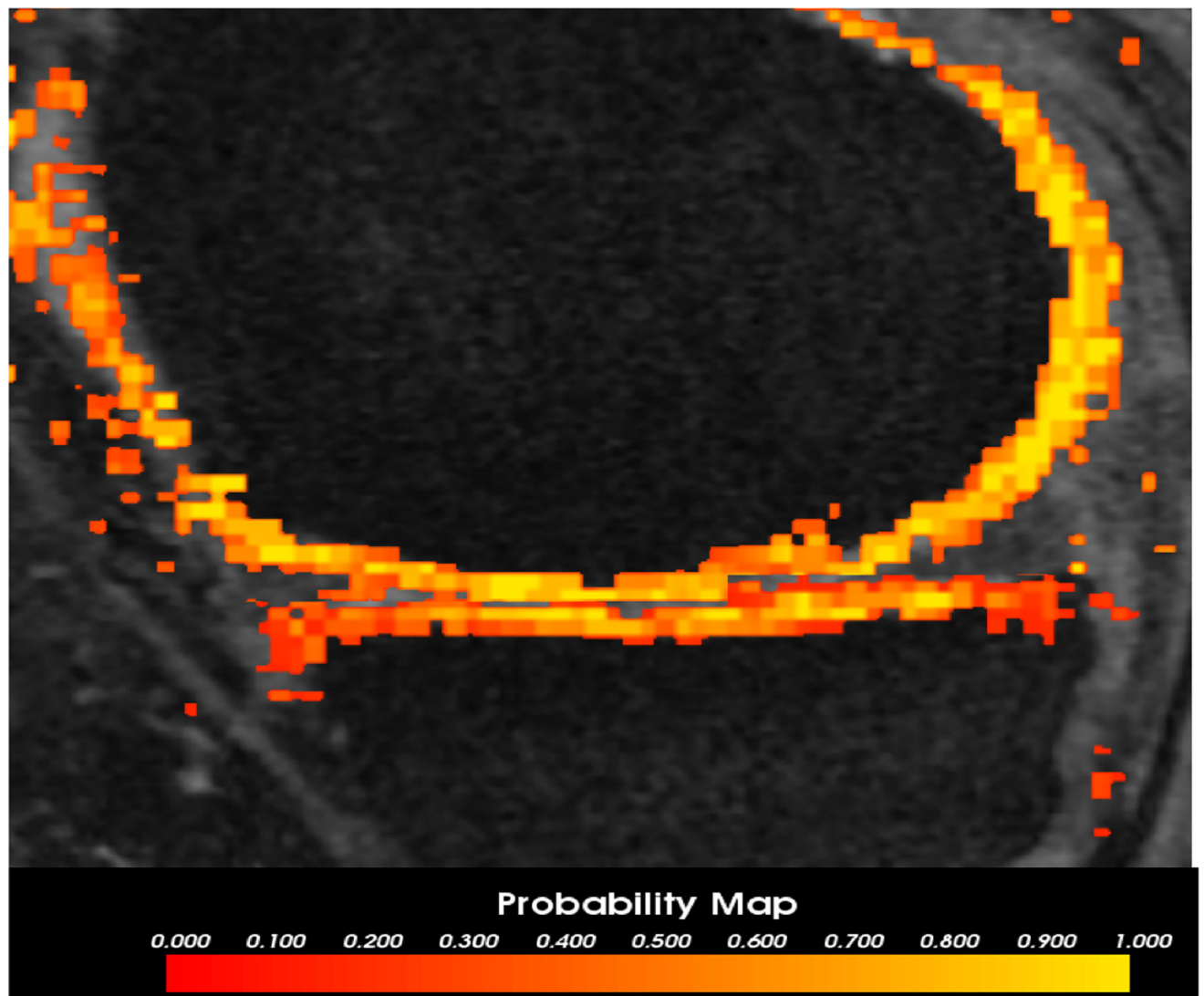
Fig. 1.
The output probability map of the NAF for an image overlaid on the image volume. The color map indicates the probability output values for each 3D voxel belonging to the femur (top) and tibia cartilage. The brighter yellow denotes higher probability while the red indicates regions of lower probability of the voxel being a cartilage.
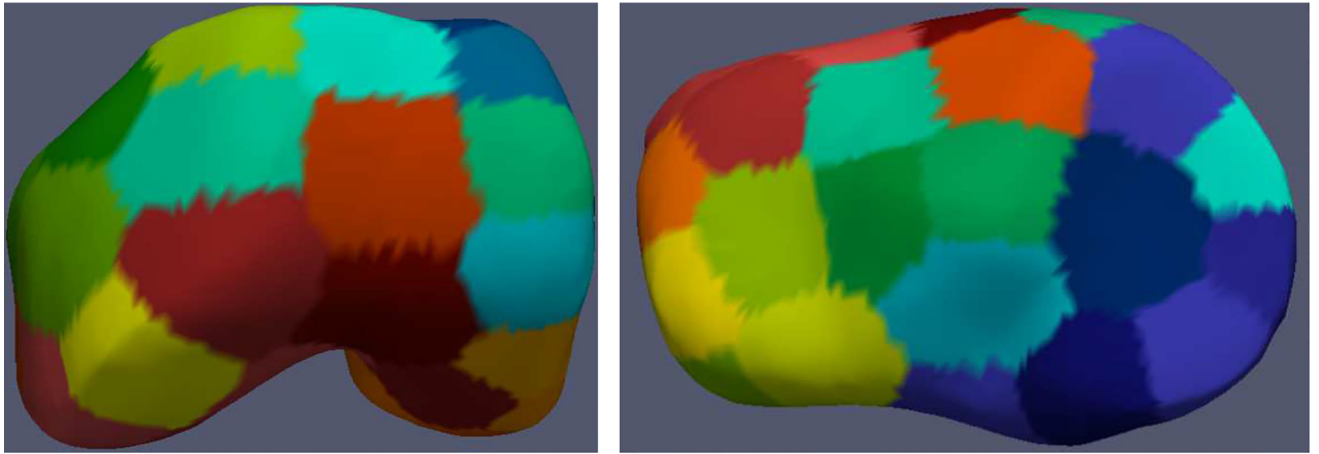
Fig. 2.
Parcellation of the S_0 mesh using k-means clustering of the femur and tibia into 40 clusters each. Each colored region represents a different cluster that is trained using a separate RF classifier to account for the regionally-specific appearance of surrounding menisci, muscle bone and other anatomies.
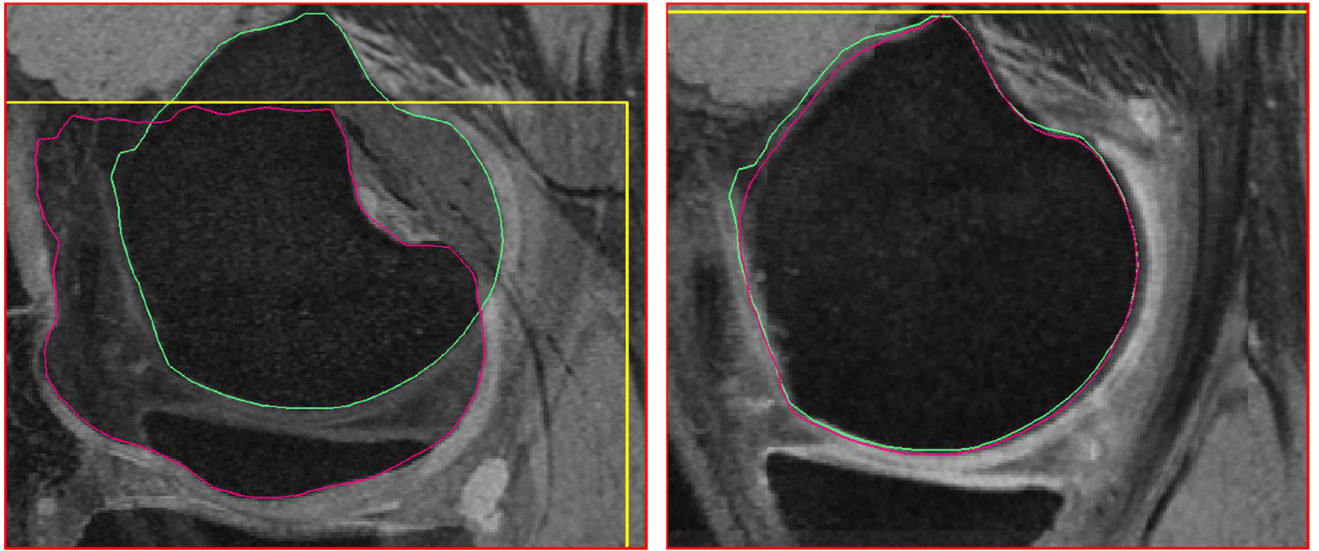
Fig. 3.
Establishing point correspondences between time-points after pre-segmentation. The meshes are registered using ICP to establish column correspondences. The same transformation is applied to the corresponding volumes. (a) Two time-points before registration. The meshes are colored differently and the volume border is highlighted in yellow to indicate its positioning. (b) Meshes registered after ICP
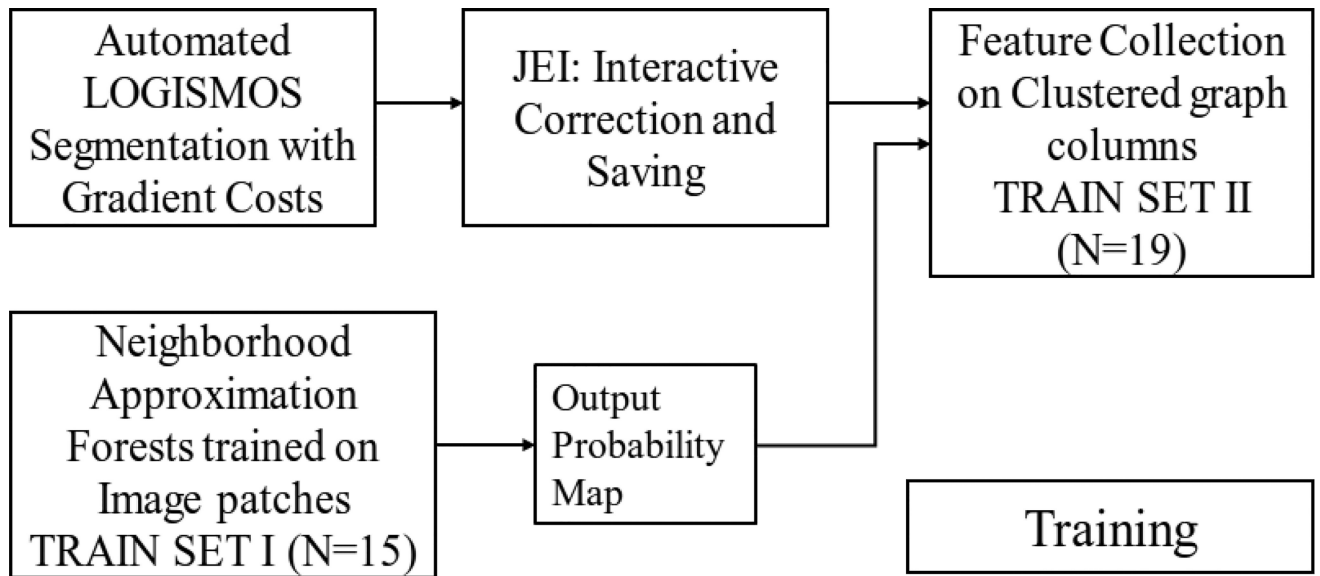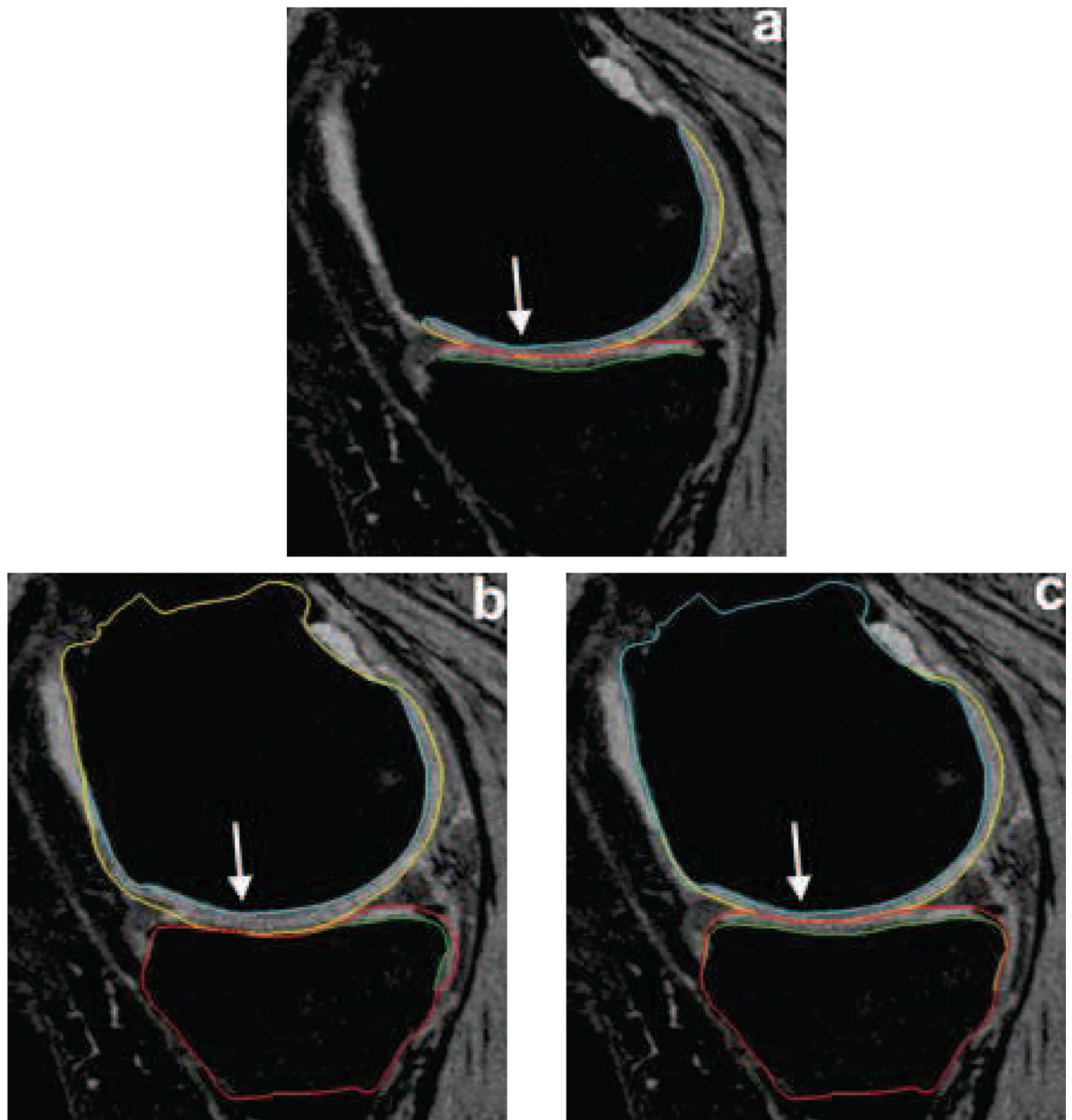
Fig. 4.
Learning-based workflow for training the hierarchical RF classifiers.
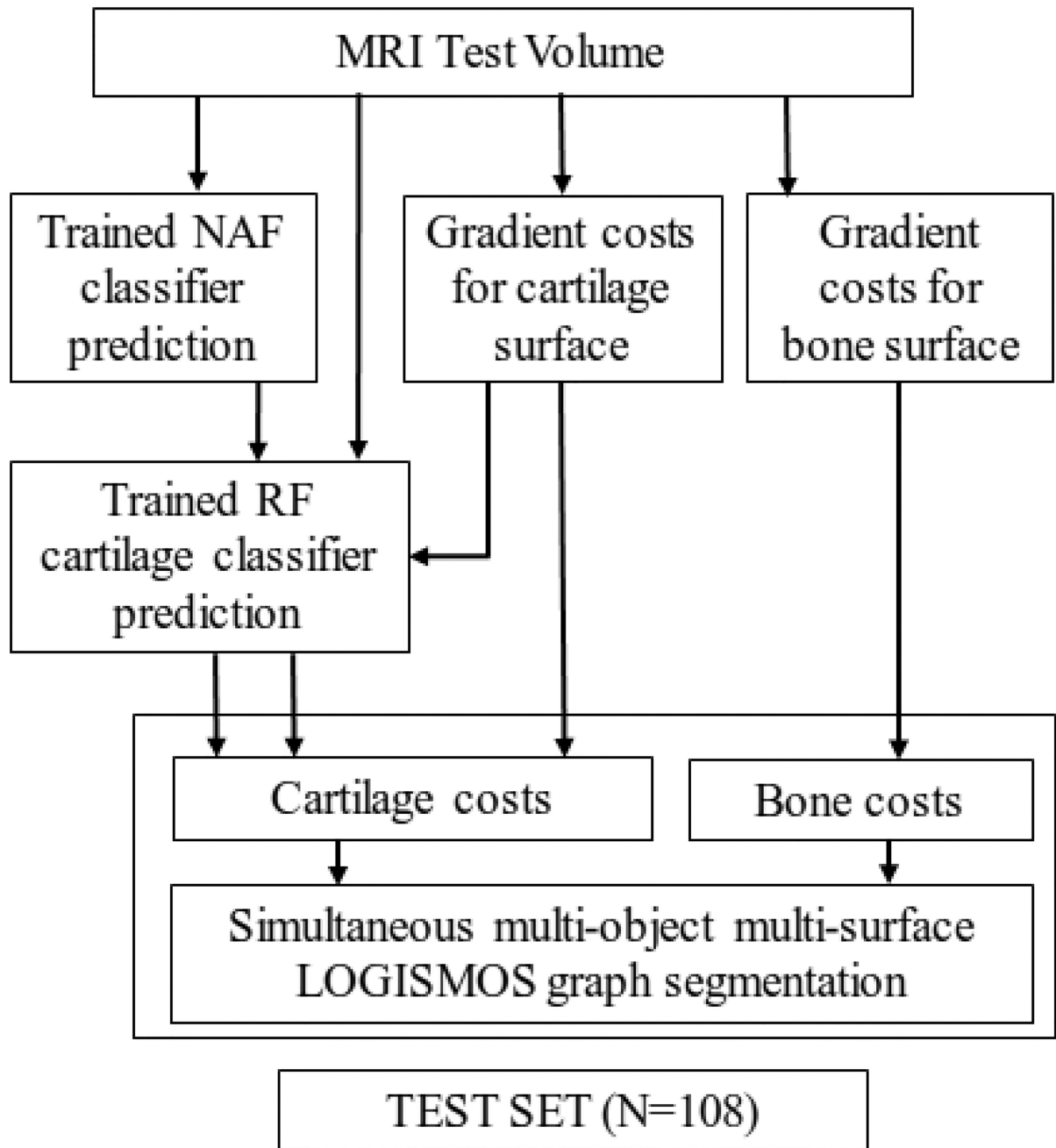
Fig. 5.
Testing workflow to compare the clustered random forest classifiers with the existing methods on a set of 108 patients.
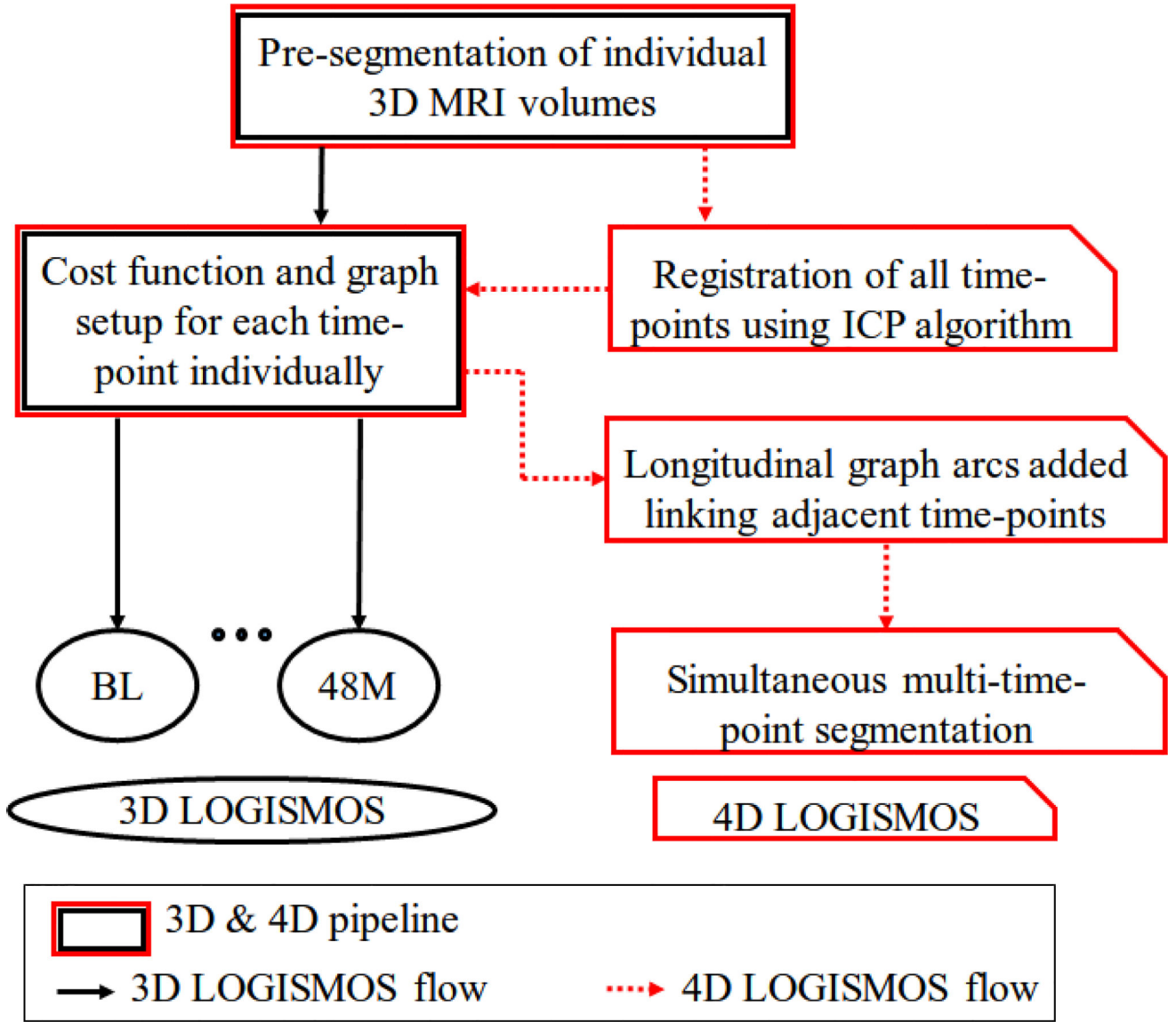
Fig. 6.
Comparison of 3D and 4D LOGISMOS pipelines. The black and red arrows highlight the flow of 3D and 4D LOGISMOS algorithms respectively.

Fig. 7.
Segmentation accuracy in a representative subject. (a) Independent standard. (b) Gradient-costs LOGISMOS segmentation. (c) Learned-costs LOGISMOS segmentation. Region marked by the arrow shows clear improvement in the segmentation quality when using the learned costs.
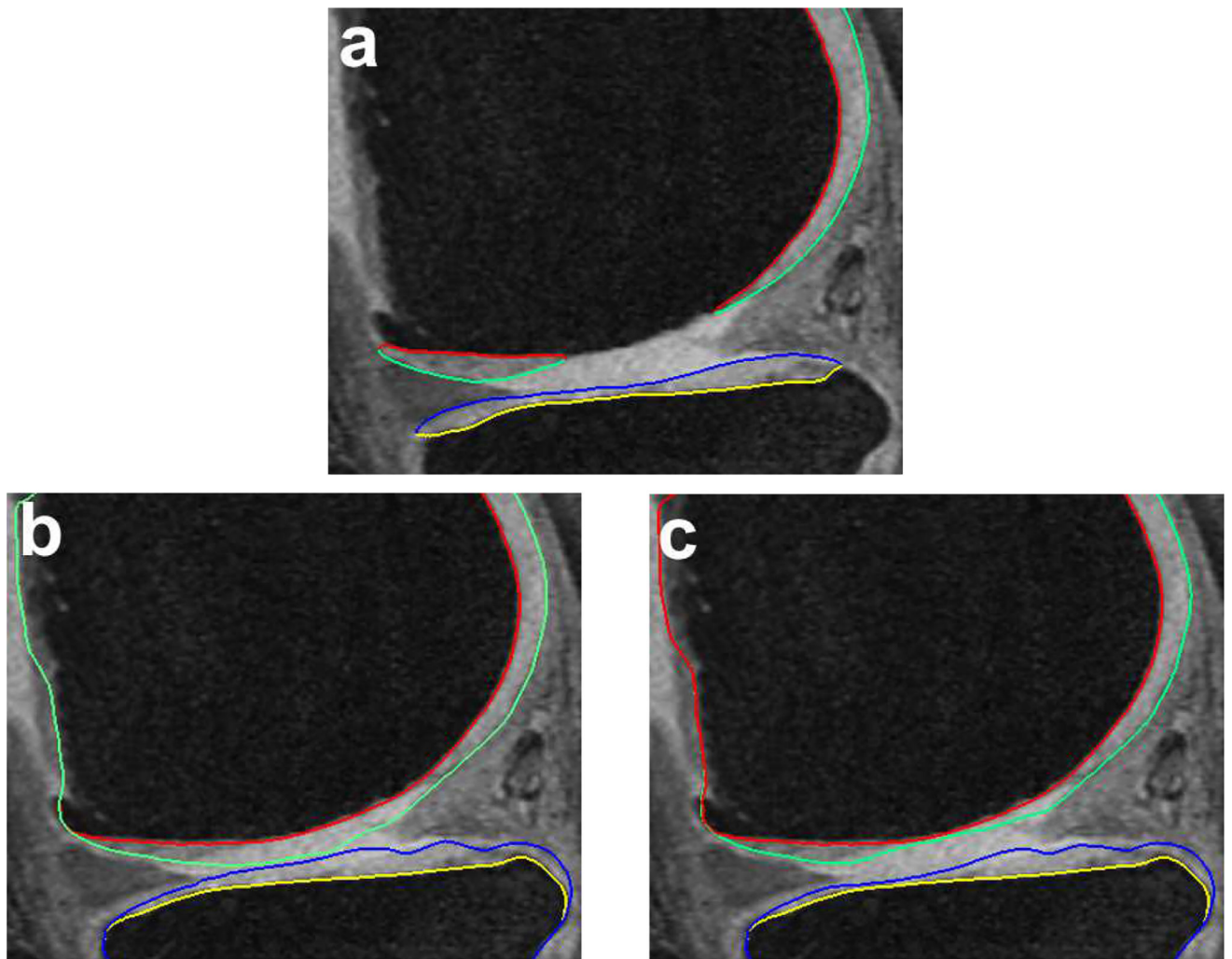
Fig. 8.
Qualitative improvement of 4D segmentation versus 3D for a patient with severe osteoarthritis. Parts of the region between the femur and tibia exhibit synovial fluid leakage, the appearance and texture of which mimic that of a cartilage. (a) Independent standard, (b) 3D segmentation, (c) 4D segmentation.
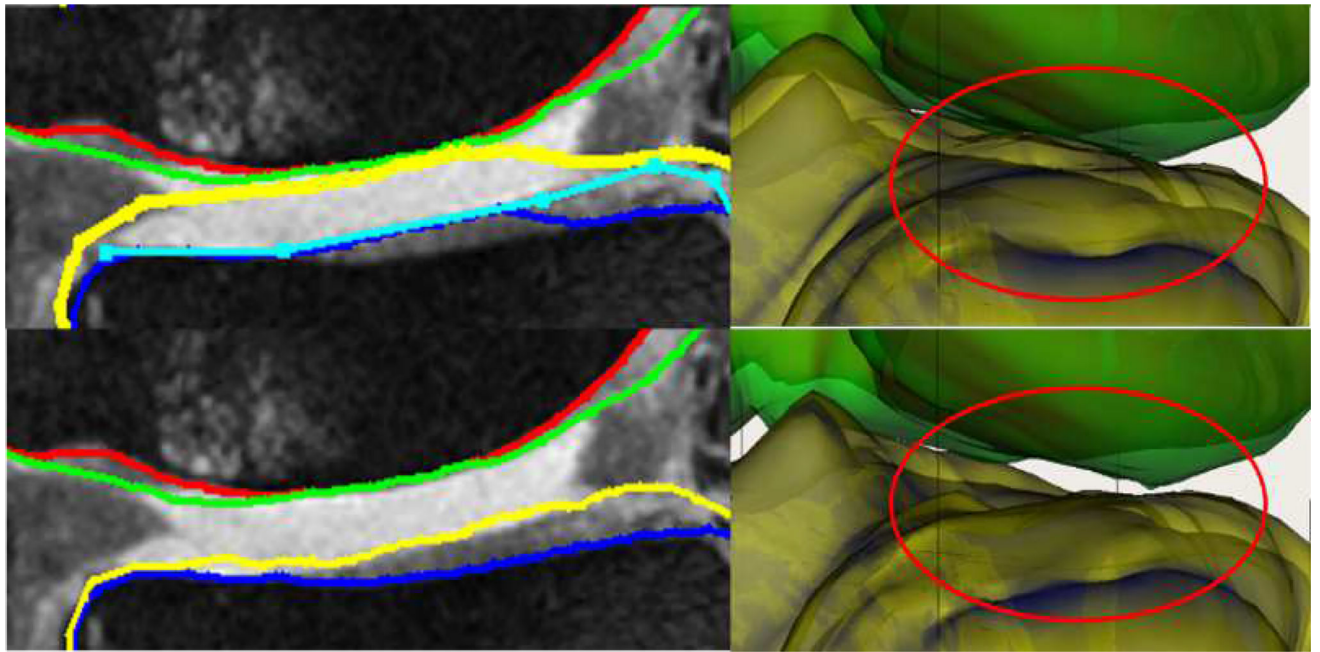
Fig. 9.
A graphical depiction of the JEI workflow showing the interactive correction steps. A single slice that is identified with improper segmentation of the tibia cartilage (yellow) is shown. The user provides the correcting set of boundary points (cyan) which identifies the 3D neighborhood of graph columns on which the costs are modified locally. The max-flow was recomputed in 3D resulting in the corrected surfaces within milliseconds. The resulting correction from JEI on a single 2D slice can be appreciated in the corresponding 3D model representation of the surface.
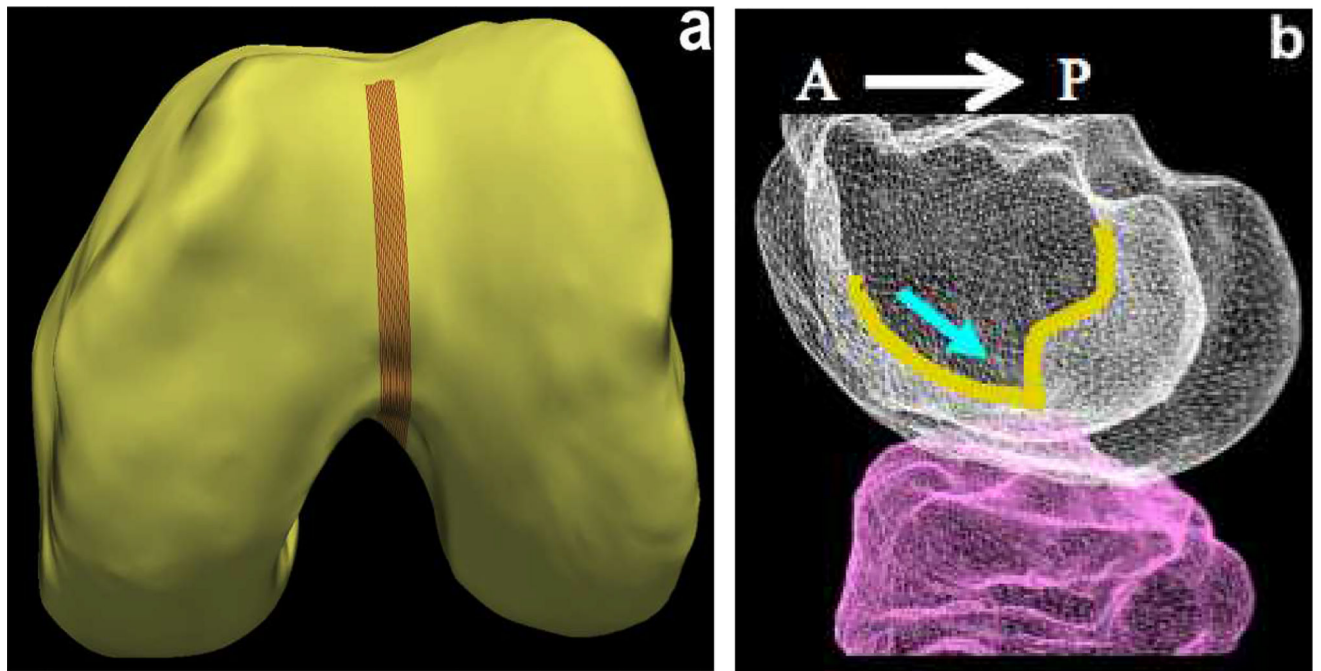
Fig. 10.
Trochlear notch. (a) LOGISMOS-segmented bone mesh with the trochlear groove region highlighted. (b) Bone mesh highlighting the trochlear groove (overlayed in yellow) in the anterior to posterior direction. The blue arrow indicates the trochlear notch on the groove at the base of the groove curvature before the bone ridge structure rises sharply.
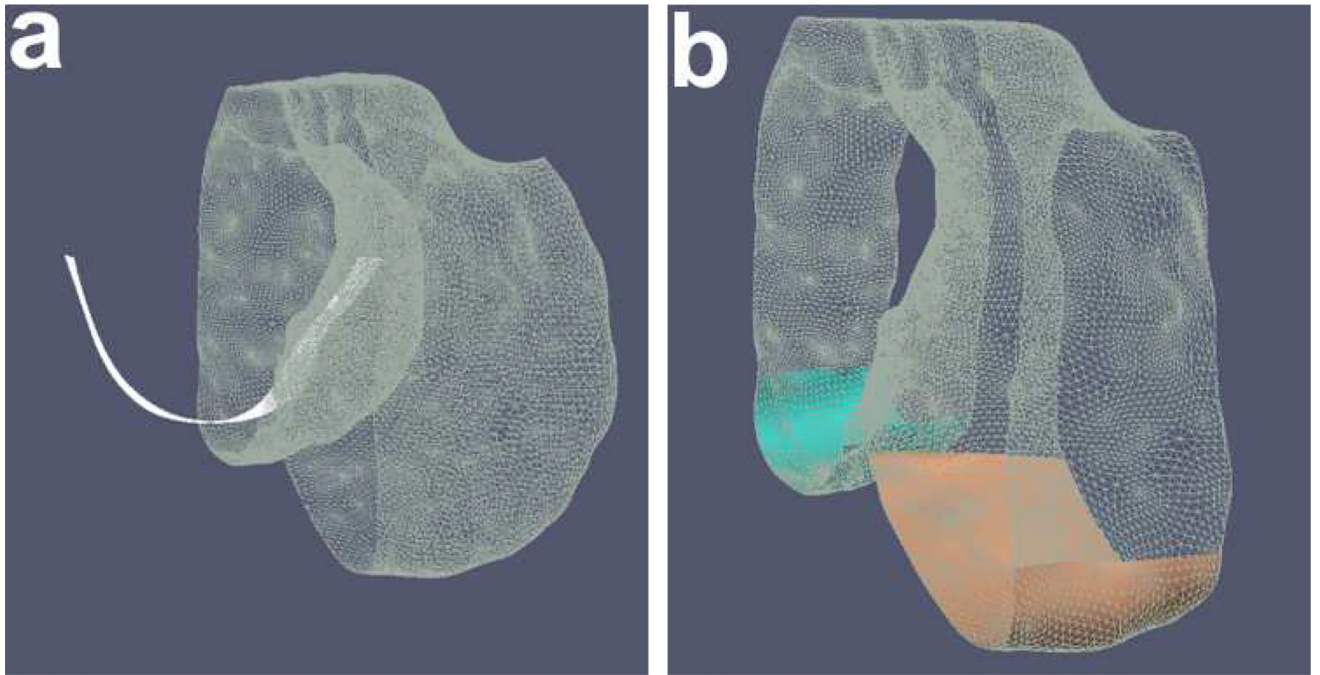
Fig. 11.
Trochlear notch. (a) Use of AP curvature of the groove on the femoral bone. After isolating the groove region, the trochlear notch is identified by taking an average gradient on a family of closely positioned contour lines along this surface. (b) The load-bearing regions of the femur identified by isolating 60 % of the distance from the trochlear notch to posterior most in the AP direction on each condyle respectively.
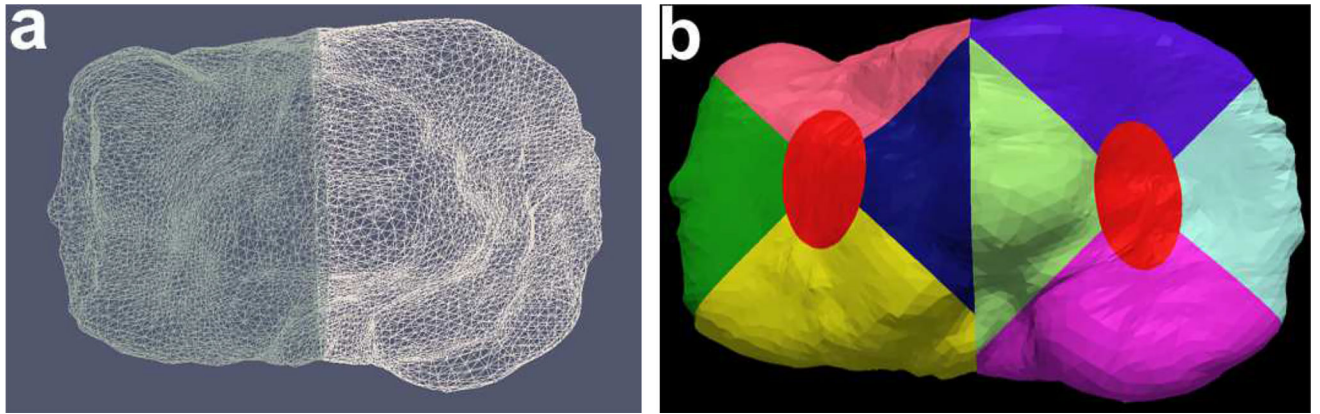
Fig. 12.
(a) Isolated medial and lateral tibia, (b) All the tibial sub-plates.
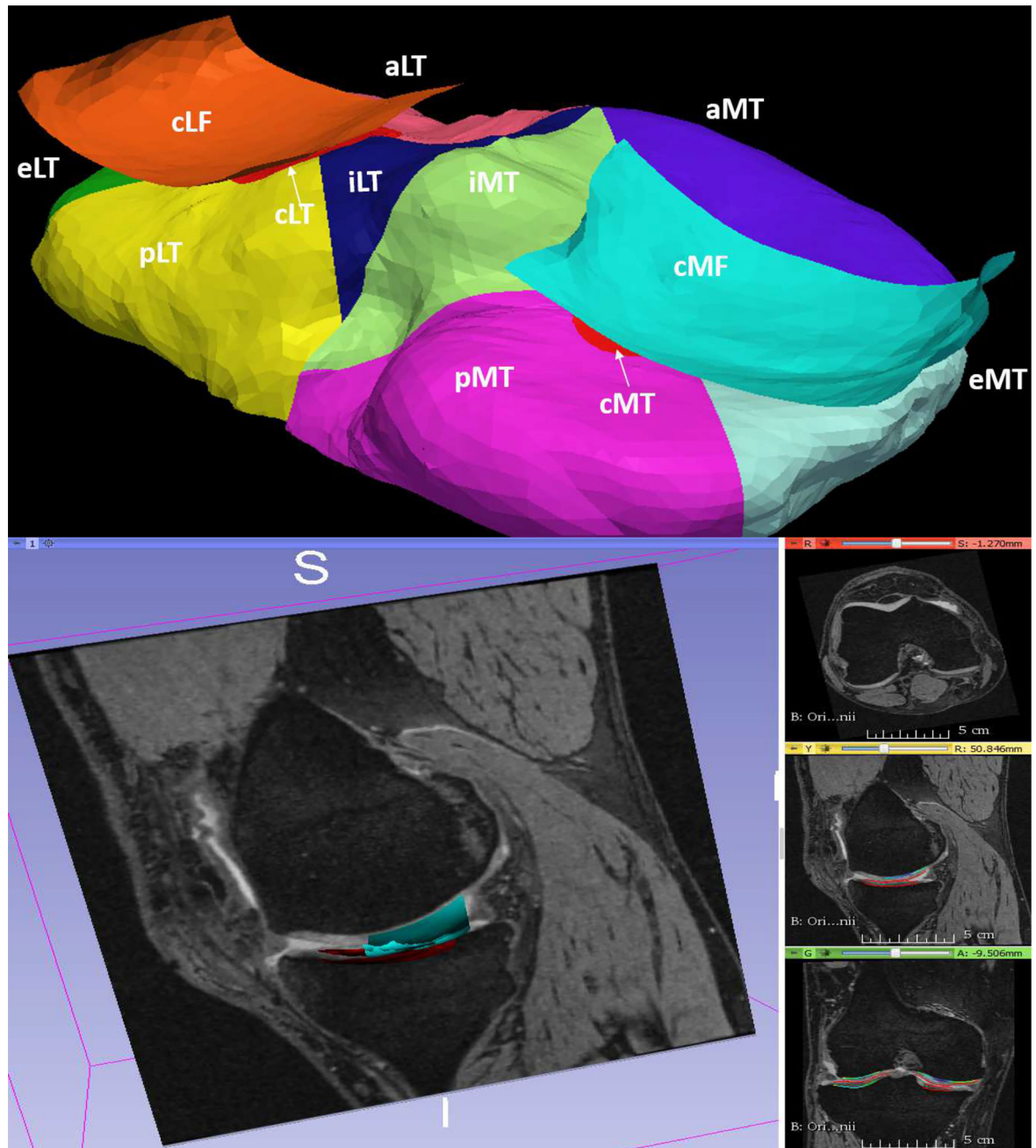
Fig. 13.
Automated sub-plate division of the cartilage. The 60% central lateral (cLF) and central medial femur (cMF) are shown. Each Medial (MT) and lateral tibia (LT) regions are subdivided as: central (cLT/cMT), interior (iLT/iMT), exterior (eLT/eMT), anterior (aLT/aMT) and posterior (pLT/pMT) regions respectively.

TABLE I

List of features for the second RF classifier.

Index	Description
1–9	3 eigenvalues of Hessian matrices on intensity image at σ = 0.5, 1.0, 2.0 mm [36]
10–15	Gaussian gradient on intensity and NAF probability volumes at σ = 0.36, 0.7, 1.4 mm [37]
16–18	Intensity, Gaussian smoothed intensity, and NAF probability volumes
19–20	Laplacian derivative of intensity volume at σ = 0.36, 0.7 mm [37]
21	Gabor texture feature [38]
22–25	Intensity statistics: mean, variance, skewness and kurtosis of a 2 mm^3 region centered around each graph node [39]
26–28	Haar features (1.5mm kernel) along horizontal, vertical & diagonal directions [40]
29–30	1D directional gradient along the search column direction on NAF probability and intensity volume

TABLE II

Parameters used for graph construction.

	Inter-surface max (mm)	Inter-object max (mm)	Smoothness (mm)	Column size (mm)
Learned cost	6	18	0.6	18.15
Gradient cost	4	12	0.4	12.2

*
Minimum inter-surface, inter-object and inter-time-point separations are zero.

TABLE III

Bone segmentation errors (in mm) using the gradient-based cost functions.

	Signed	Unsigned
Femur	−0.19±0.07	0.40±0.06
Tibia	−0.13±0.09	0.39±0.09

TABLE IV

Cartilage surface positioning errors (in mm) achieved by hierarchical classifier, gradient cost and single-stage RF classifier using 3D LOGISMOS algorithm. Paired t-test comparisons are reported for (NAF+RF vs. Gradient) and (NAF+RF vs. RF only)

n=108	NAF+RF	Gradient	*p*-value	RF only	*p*-value
Femur signed	**−0.01±0.18**	−0.31±0.24	≪ 0.001	−0.10±0.17	≪ 0.001
Femur unsigned	**0.55±0.11**	0.69±0.13	≪ 0.001	0.56±0.10	≪ 0.001
cMF signed	**−0.04±0.29**	−0.38±0.58	≪ 0.001	−0.11±0.27	≪ 0.001
cMF unsigned	**0.52±0.16**	0.78±0.35	≪ 0.001	0.55±0.17	≪ 0.001
cLF signed	**−0.26±0.24**	−0.52±0.35	≪ 0.001	−0.36±0.20	≪ 0.001
cLF unsigned	**0.42±0.12**	0.65±0.20	≪ 0.001	0.47±0.11	≪ 0.001
Tibia signed	**0.06±0.17**	−0.11±0.35	≪ 0.001	0.11±0.22	≪ 0.001
Tibia unsigned	**0.60±0.14**	0.79±0.20	≪ 0.001	0.62±0.18	≪ 0.001
cMT signed	−0.15±0.31	−0.25±0.77	0.193	**−0.09±0.34**	0.003
cMT unsigned	**0.52±0.20**	0.92±0.41	≪ 0.001	0.58±0.22	≪ 0.001
cLT signed	**−0.03±0.32**	0.36±1.12	≪ 0.001	−0.01±0.32	0.16
cLT unsigned	**0.46±0.17**	0.79±0.99	≪ 0.001	0.47±0.18	0.31

*
Bold indicates statistically significant reduction in positioning errors.

TABLE V

Cartilage surface positioning errors (in mm) of 4D versus 3D LOGISMOS.

	n=54 Baseline			n=54 12 months		
	4D	3D	*p*-value	4D	3D	*p*-value
Femur signed	**0.01±0.18**	0.01±0.19	0.027	−0.02±0.17	−0.04±0.17	0.062
Femur unsigned	**0.53±0.11**	0.54±0.11	≪0.001	0.55±0.11	0.55±0.11	0.601
cMF signed	**−0.01±0.25**	−0.03±0.27	0.015	**0.02±0.27**	−0.04±0.30	0.016
cMF unsigned	0.51±0.16	0.52±0.17	0.092	0.52±0.15	0.53±0.16	0.614
cLF signed	−0.26±0.22	−0.25±0.23	0.108	−0.30±0.21	−0.30±0.23	0.853
cLF unsigned	0.42±0.10	0.42±0.10	0.790	0.43±0.12	0.44±0.13	0.318
Tibia signed	0.08±0.17	**0.07±0.16**	0.039	0.07±0.19	0.06±0.18	0.464
Tibia unsigned	0.59±0.14	0.60±0.14	0.140	0.60±0.16	0.60±0.14	0.310
cMT signed	−0.13±0.28	−0.14±0.29	0.125	−0.14±0.29	−0.15±0.31	0.619
cMT unsigned	**0.50±0.18**	0.51±0.20	0.039	0.52±0.17	0.53±0.17	0.341
cLT signed	0.00±0.30	0.00±0.31	0.338	−0.06±0.31	−0.06±0.30	0.629
cLT unsigned	**0.45±0.17**	0.46±0.18	≪0.001	0.45±0.16	0.44±0.16	0.256

*
Bold indicates statistically significant reduction in positioning errors.

TABLE VI

Longitudinal max thickness error values of 4D versus 3D LOGISMOS.

	2% Max Average Errors (98–100%)			5% Max Average Errors (95–100%)			10% Max Average Errors (90–100%)		
	4D	3D	*p*-value	4D	3D	*p*-value	4D	3D	*p*-value
cMF unsigned	**0.76±0.60**	1.11±0.63	≪0.001	**0.57±0.43**	0.87±0.41	≪0.001	**0.43±0.32**	0.68±0.29	≪0.001
cLF unsigned	1.46±2.36	1.67±1.82	0.610	1.07±1.73	1.16±1.02	0.74	0.74±0.95	0.84±0.55	0.528
cMT unsigned	**1.28±0.47**	1.78±1.36	0.008	**1.09±0.42**	1.48±1.00	0.004	**0.92±0.38**	1.19±0.63	≪0.001
cLT unsigned	**1.29±0.47**	1.40±0.49	≪0.001	**1.09±0.40**	1.17±0.41	≪0.001	**0.90±0.33**	0.98±0.34	≪0.001

*
Bold indicates statistically significant reduction in positioning errors.