Automatic Atlas-based Three-label Cartilage Segmentation from MR Knee Images

Liang Shan
Department of Computer Science
UNC Chapel Hill
shan@cs.unc.edu

Cecil Charles
Department of Radiology
Duke University
cecil.charles@duke.edu

Marc Niethammer
Department of Computer Science
UNC Chapel Hill
mn@cs.unc.edu

Abstract

This paper proposes a method to build a bone-cartilage atlas of the knee and to use it to automatically segment femoral and tibial cartilage from T1 weighted magnetic resonance (MR) images. Anisotropic spatial regularization is incorporated into a three-label segmentation framework to improve segmentation results for the thin cartilage layers. We jointly use the atlas information and the output of a probabilistic k nearest neighbor classifier within the segmentation method. The resulting cartilage segmentation method is fully automatic. Validation results on 18 knee MR images against manual expert segmentations from a dataset acquired for osteoarthritis research show good performance for the segmentation of femoral and tibial cartilage (mean Dice similarity coefficient of 78.2% and 82.6% respectively).

1. Introduction

Osteoarthritis (OA) is the most common form of joint disease and characterized by cartilage loss. An accurate cartilage segmentation from magnetic resonance (MR) knee images is crucial to study OA. Due to the size of image databases acquired for OA studies, a fully automatic segmentation is needed.

Recently, several automatic methods have been proposed for cartilage segmentation. Folkesson *et al*. [2] proposed a hierarchical classification scheme for cartilage segmentation. Fripp *et al*. [3] used active shape models for bone segmentation in order to extract the bone-cartilage interface followed by tissue classification. A simultaneous segmentation of interacting bone and cartilage was developed by Yin *et al*. [10]. To allow for localized analysis and the suppression of unlikely voxels in a segmentation, introducing a spatial prior is desirable. This can be achieved through an atlas-based analysis method. While such methods have been successfully used in brain imaging, they are typically not used for cartilage segmentation in the knee. Presumably, one of the reasons is that atlas-building in the knee is significantly more challenging (due to the articulated system with different tissue properties) and a highly accurate method is needed to capture the thin cartilage areas. None of the aforementioned methods use a spatial atlas. The work by Glocker *et al*. [4] is an exception, however it assumes that a set of pre-aligned images is already available.

Instead, in this work we discuss a fully automatic atlas-based cartilage segmentation method. The method constructs a bone-cartilage atlas for femur and tibia and uses the resulting atlas as prior information to guide femoral and tibial cartilage segmentation. We demonstrate that an atlas of sufficient quality can indeed be constructed to help increase segmentation robustness, mitigate noise effects, and focus segmentations on desired regions of interest.

Since femoral and tibial cartilage may touch, a joint segmentation using a multi-class segmentation is desirable. We make use of a three-label segmentation approach [8] ensuring distinct labels for touching objects. To customize the method for cartilage segmentation, we incorporate an anisotropic regularization term into the three-label segmentation method to avoid over-regularization of the thin cartilage layers. We demonstrate improved robustness and quality of anisotropic over isotropic regularization for cartilage segmentation.

Section 2 discusses the anisotropic three-label cartilage segmentation approach. Section 3 discusses the probabilis-

tic kNN classification used within the segmentation step. Section 4 describes the knee atlas building method. The overall segmentation method is described in section 5. Experimental results are given in section 6. The paper closes with conclusions and future work.

2. Segmentation of Cartilage

To avoid the possible merging of femoral and tibial cartilage segmentations we make use of a three-label segmentation formulation [8]. The three-label case is a specialization of a multi-label segmentation method [11] which allows for a symmetric formulation with respect to the background segmentation class.

The multi-label segmentation [11] energy defined on an image domain Ω and a labeling space $\mathcal{L} = \{0, ..., L-1\}$ is

$$\begin{aligned} E(u) &= \int_{\mathcal{D}} g \left\|\nabla_{\mathbf{x}} u\right\| + c \left|\nabla_l u\right| \, d\mathbf{x}dl, \\ \mathcal{D} &= \Omega \times \mathcal{L}, \; u(\mathbf{x}, 0) = 0, \; u(\mathbf{x}, L) = 1, \end{aligned} \tag{1}$$

where u is a level function whose discontinuity set defines labels; $\nabla_{\mathbf{x}} u$ is the spatial gradient of u, $\nabla_{\mathbf{x}} u = (\partial u/\partial x, \partial u/\partial y, \partial u/\partial y)^T$ and $\nabla_l u$ is the gradient in label direction, $\nabla_l u = \partial u/\partial l$; g controls the regularization and c defines the labeling cost. This is a convex formulation and yields a global optimal solution. For cartilage segmentation, we use three labels (background, femoral and tibial cartilage). By positioning the background label ($l = 1$) between the labels for femoral ($l = 0$) and tibial ($l = 2$) cartilage, the segmentation problem becomes symmetric with respect to the cartilage labels [8]. The labeling cost c for each label l in $\{0, 1, 2\}$ are defined by log-likelihoods for each label:

$$c(\mathbf{x}, l) = -log(P(l|\mathbf{f}(\mathbf{x}))) = -log\left(\frac{p(\mathbf{f}(\mathbf{x})|l) \cdot P(l)}{p(\mathbf{f}(\mathbf{x}))}\right), \tag{2}$$

where $\mathbf{f}(\mathbf{x})$ denotes a feature vector (specified in section 3) at $\mathbf{x}$. The likelihoods $p(\mathbf{f}(\mathbf{x})|l)$ are obtained from a probabilistic k nearest neighbor (kNN) classifier (section 3) and the priors $P(l)$ from a probabilistic atlas (section 4).

To avoid a shrinking of cartilage segmentations (due to the thin shape of femoral and tibial cartilage) while allowing for spatial regularity, we replace the isotropic regularization term, g in (1), by an anisotropic one

$$\begin{aligned} E(u) &= \int_{\mathcal{D}} \left\|\mathbf{G}\nabla_{\mathbf{x}} u\right\| + c\left|\nabla_l u\right| \, d\mathbf{x}dl, \\ \mathcal{D} &= \Omega \times \mathcal{L}, \; u(\mathbf{x}, 0) = 0, \; u(\mathbf{x}, L) = 1, \end{aligned} \tag{3}$$

where $\mathbf{G}$ is a positive-definite matrix determining the amount of regularization. This avoids over-regularization at the boundaries of the cartilage layers and improves segmentations. Figure 1 illustrates the problem with isotropic regularization which tends to shrink the segmentation boundary by cutting thin objects short and the benefit of anisotropic regularization. We choose $\mathbf{G}$ similar to [7] as

$$\mathbf{G} = g\left[\mathbf{I} + (\alpha - 1)\mathbf{n}\mathbf{n}^T\right], \alpha \in [0, 1], \tag{4}$$

where $\mathbf{I}$ is the identity matrix and $\mathbf{n}$ is a unit vector indicating the direction of less regularization (the normal direction to the cartilage surface); α determines the degree of anisotropy ($\alpha = 1$ is equivalent to isotropic). See Fig. 1 (d) for an illustration of isotropic versus anisotropic regularization. The computation of the normal direction $\mathbf{n}$ is described in section 6. The numerical solution method for (3) is discussed in Appendix A.

3. Probabilistic kNN Classification

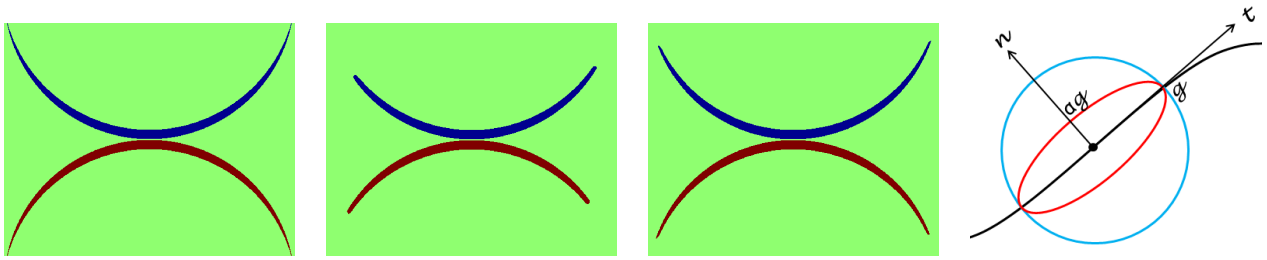
Data likelihoods for femoral and tibial cartilage, $p(\mathbf{f}(\mathbf{x})|l)$, of (2) are estimated by probabilistic kNN classification [1]. Compared to [2] this allows to integrate the kNN classification results into the overall three-label segmentation framework, ensuring spatial regularity. Compared to [2] we choose a reduced set of 15 features: intensities on three scales, first-order derivatives in three directions on three scales and second-order derivatives in axial direction on three scales. The three different scales are obtained by convolving with Gaussian kernels of $\sigma = 0.3\ mm, 0.6\ mm$ and $1.0\ mm$. All features are normalized to be centered at 0 and to have unit standard deviation. We use a one-versus-other classification strategy and expert segmentations of femoral and tibial cartilage to build the kNN classifier. Specifically, let f denote the femoral cartilage class, t the tibial cartilage and b the background class. The training samples of class f and class t are the voxels labeled as femoral and tibial cartilage respectively. The training samples of class b are the voxels surrounding the femoral and tibial cartilage within a specified distance. The outputs of the probabilistic kNN classifier given a query voxel $\mathbf{x}$ with its feature vector $\mathbf{f}(\mathbf{x})$ are:

$$\begin{aligned} p(\mathbf{f}(\mathbf{x})|l = 0) &= n_f(\mathbf{f}(\mathbf{x}))/k, \\ p(\mathbf{f}(\mathbf{x})|l = 2) &= n_t(\mathbf{f}(\mathbf{x}))/k, \\ p(\mathbf{f}(\mathbf{x})|l = 1) &= n_b(\mathbf{f}(\mathbf{x}))/k. \end{aligned} \tag{5}$$

Here n_f, n_t, n_b denote the number of votes for the femur, tibia, and background class respectively; k is the number of nearest neighbors of concern. Since kNN is sensitive to the number of training samples, we scale the outputs according to the training class sizes to balance the three classes.

4. Atlas Building

A probabilistic atlas provides the spatially-dependent prior, $P(l)$, for the segmentation method. This allows to restrict the segmentation to regions of interest, helps minimize noise influences and improves segmentation robustness.



(a) (b) (c) (d)

Figure 1. Synthetic example. (a) original image to be segmented; (b/c) three-label segmentation results with isotropic and anisotropic regularization respectively. Anisotropic regularization avoids over-regularization at the tips of the synthetic shape. (d) Difference between isotropic and anisotropic regularization. The black curve indicates an edge in an image. The regularization is illustrated at a pixel (the dot). The blue circle indicates the isotropic case where regularization is enforced equally in every direction. The red ellipse shows the anisotropic situation where less regularization is applied in the normal direction and more in the tangent direction.

Bone segmentations are obtained through a three-label segmentation [8] with background positioned in the middle. Once the segmentations for femur and tibia have been obtained, we apply a segmentation-based alignment method proposed in [9] to bring the knee images to a reference position. Transformation models between bone and surrounding tissue are typically distinct. We choose to affinely align femurs and tibiae separately across subjects. Two separate space transformations are combined into one globally valid transformation, enforcing local transformations for bones exactly, where the ambient space (to the bones) is deformed elastically.

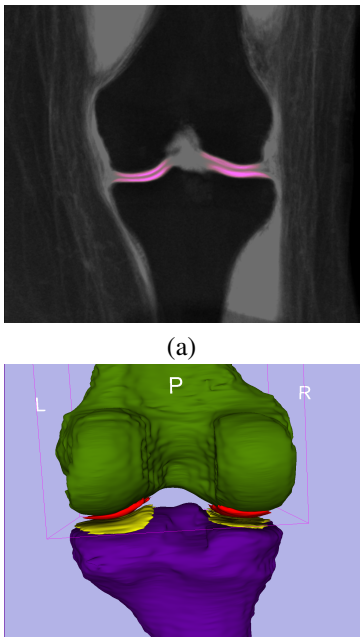
Given a set of aligned images, we compute a knee atlas by local averaging. The atlas contains an average of T1 weighted images, averages of femur and tibia segmentations as well as averages for femoral and tibial cartilages. The segmentation averages directly yield local label probabilities.

5. Overall Segmentation Method

Given a new query image, we first perform the bone segmentation [8] followed by the segmentation-based alignment step [9] to align the atlas to the query image in order to obtain local priors for femoral and tibial cartilage. Then we extract features from the region covered by the propagated atlas of femoral and tibial cartilage and import these feature vectors into the probabilistic kNN classifier. The soft classification yields the data likelihoods of being femoral and tibial cartilage. Once the data likelihoods and atlas priors are obtained we perform the cartilage segmentation using the three-label segmentation method of section 2.

6. Experimental Results

We test the proposed approach on a set of 18 MR images (T1 weighted SPGR images acquired coronally at a resolution of $1.00 \times 0.31 \times 0.31 mm^3$; as well as matching sagittal T1 and T2* weighted images) from different



(a)

(b)

Figure 2. Atlas. (a) shows a slice of the probabilistic atlas (pink) of femoral and tibial cartilage overlaid onto the atlas of the T1 weighted image (coronal view). Brightness denotes probability. (b) is a 3-dimensional rendering of the thresholded atlas of femur(green), tibia (purple), femoral (red) and tibial cartilage (yellow).

subjects. Expert cartilage segmentations are available coronally for all images. The femoral cartilage segmentation is drawn only on the weight-bearing part while the tibial cartilage segmentation covers the entire region. Therefore, we expect partial femoral cartilage segmentations and full tibial cartilage segmentations. The 18 images have different Kellgren-Lawrence grades (KLG) [6]: 12 images have KLG $= 0$, 3 have KLG $= 2$, and 3 have KLG $= 3$.

Expert segmentations are provided in coronal space. We therefore perform also all segmentations in coronal space.

Table 1. Improvement of mean DSC (standard deviation) by including the probabilistic atlas for the leave-one-out experiment with 18 datasets (only 17 are tested, the reference image for atlas building not tested). An isotropic regularization with $g = 0.5$ is used.

	Without atlas	With atlas
Femoral cartilage	69.2%(6.7%)	77.3%(5.2%)
Tibial cartilage	77.3%(5.2%)	82.2%(3.8%)

Both T1 and T2* weighted images are needed to estimate the bone likelihoods [8]. Since T2* weighted images are not available coronally for our data, we register the coronal and sagittal T1 weighted images affinely (within subject, acquired in the same session) to create "virtual" coronal T2* weighted images. We choose one KLG $= 0$ image as a reference to align all other images. An atlas is built for each of the 17 test images (except for the one chosen as the reference) in a leave-one-out fashion. Classification results are also computed in a leave-one-out fashion. Each of the 17 data sets (except for the reference image in the atlas building step) is tested using the remaining 17 data sets as training samples and k for the kNN classifier is chosen to be 30 through all our experiments. The reference image is used in training but not tested. In our experiment, the normal direction $\mathbf{n}$ in Fig. 1 (d) is computed by taking the gradient of the diffusion smoothed three-label bone segmentation result in-between the joint area. We find $\alpha = 0.2$, which controls the degree of anisotropy, produces good results and use this setting in all anisotropic experiments.

Figure 2 shows an atlas constructed from 17 images which provides the priors for the segmentation method. Image features and averaged labelmaps are crisp, which indicates a good quality of knee image alignment.

Figure 3 illustrates the beneficial behavior of our three-label segmentation method compared to a binary segmentation which treats femoral and tibial cartilage as one object. While the three-label method is able to keep femoral and tibial cartilage separated (due to the joint estimation of the segmentation) the binary segmentation approach cannot guarantee this separation.

Figure 4 demonstrates the benefit of incorporating the probabilistic atlas. Prior knowledge of the cartilage location can reduce misclassified regions. Table 1 shows the significant improvement in segmentation quality as measured by the Dice similarity coefficient (DSC) by including the probabilistic atlas.

Figure 5 shows the advantage of anisotropic regularization. Isotropic regularization has a tendency to cut long and thin objects short as shown in Fig. 5 (a) at the medial femoral cartilage. Anisotropic regularization, on the other hand, avoids this problem (see Fig. 5 (b)) resulting in a better segmentation of the medial femoral cartilage.

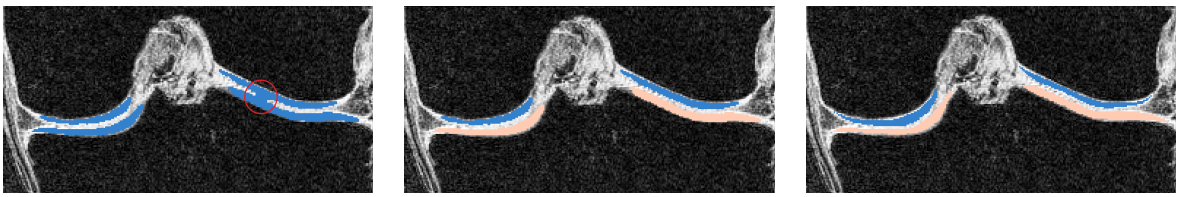
Besides avoiding unrealistic segmentation results, anisotropic regularization is also less sensitive to parameter settings than isotropic regularization. This is illustrated in Fig. 6 (a) and (b). Note that the anisotropic regularizer is parametrized in such a way that its regularization is reduced in the normal direction, but equal to the isotropic regularization in the plane orthogonal to the normal and the results are therefore comparable (see Fig. 1 (d)). The faster drop-off in the isotropic case indicates a stronger dependency on the parameter settings for isotropic regularization. An anisotropic regularization does not only yield an improvement in DSC, but also increases sensitivity notably with little sacrifice in specificity (see Fig. 6 (c) and (d)).

Different methods are not directly comparable as testing is performed on different datasets. However, Table 2 gives a rough idea of the good performance of our method in comparison to other approaches. Our method achieves the highest DSC together with the lowest standard deviation for tibial cartilage. It fares slightly worse for the femoral cartilage. Note however, that our experiments are based only on the weight-bearing region of the femoral cartilage and hence will be much more sensitive to slight missegmentations. Our method is designed for the data from the Pfizer longitudinal study for which T1 and T2* weighted images are available (the datasets of the osteoarthritis initiative are multi-spectral also). Hence, we use both images to obtain bone segmentations for robust alignment. Unfortunately, this precludes testing our method on the dataset of the cartilage segmentation challenge (SKI10) [5], since it only includes T1 weighted images.

7. Conclusion and Future Work

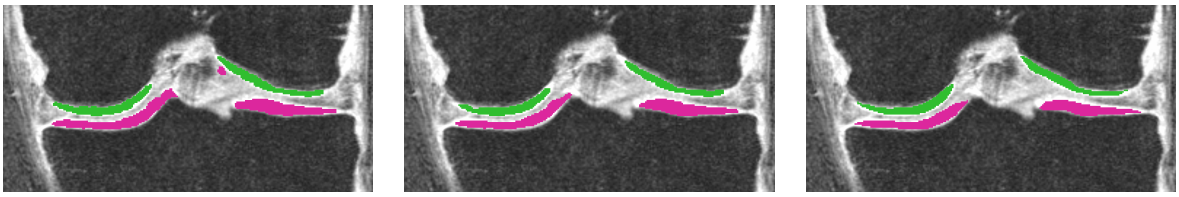
We propose an automatic atlas-based three-label cartilage segmentation approach. The method is based on a knee atlas which provides spatial priors for the segmentation method. Data likelihoods are obtained through a probabilistic kNN classifier. By introducing anisotropic regularization into the three-label segmentation framework, we improve overall segmentation accuracy. Validation of the proposed method shows good performance (a mean DSC of 78.2% for femoral cartilage and 82.6% for tibial cartilage) and demonstrates the feasibility of atlas-based analysis for cartilage segmentations.

A number of improvements over the current approach are conceivable. While kNN is a sensible classification choice, a more advanced classifier could potentially improve the classification accuracy. The normal direction $\mathbf{n}$ is purely geometrically determined for anisotropic regularization. An image-based anisotropic regularizer will be more suitable for images with cartilage lesions. We will also develop a new bone segmentation pipeline which requires only T1 weighted images to test our method on SKI10 [5] dataset. Most crucially, our current test is performed on a limited number of images with different stages of OA. We will test our method on a large set of images and evaluate it for dif-



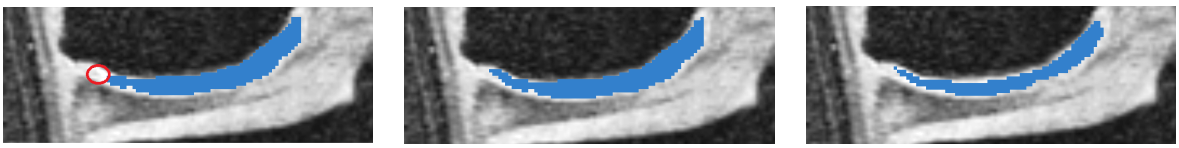
(a) (b) (c)

Figure 3. Example comparing binary and three-label segmentation methods. (a) is the binary segmentation result. (b) is the three-label segmentation result in which femoral and tibial cartilage have distinct labels. (d) is the expert segmentation. In (a), as the red circle indicates, the lateral (right) femoral cartilage and tibial cartilage are segmented as one object and the joint boundary is not well captured. The three-label segmentation (b) keeps the femoral and tibial cartilage separate and is therefore superior to binary segmentation.



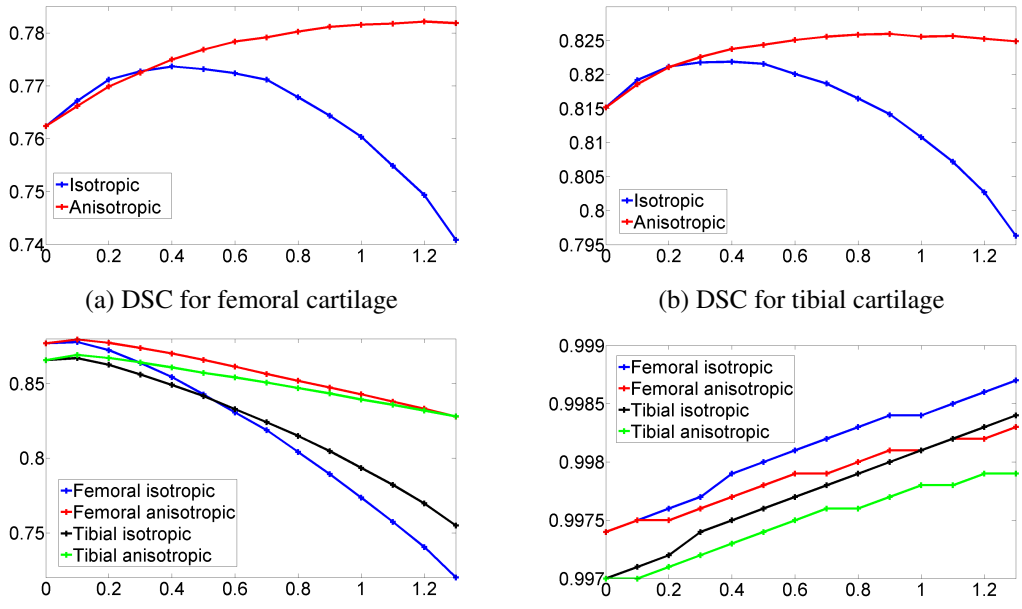
(a) Without atlas (b) With atlas (c) Expert segmentation

Figure 4. Segmentation results with and without the atlas. All are coronal views and all use an isotropic regularization with $g = 0.5$. Only joint region is shown. (a) is the segmentation without the atlas. (b) is the segmentation with the atlas. (c) is the expert segmentation. The atlas eliminates the spurious region.



(a) Isotropic (b) Anisotropic (c) Expert segmentation

Figure 5. Improvement by anisotropic regularization. (a) uses isotropic regularization and misses circled region. (b) uses anisotropic regularization and captures the missing region in (a). (c) is the expert segmentation.



(a) DSC for femoral cartilage (b) DSC for tibial cartilage

(c) Sensitivity (d) Specificity

Figure 6. Change of mean DSC, sensitivity and specificity of femoral and tibial cartilage with isotropic and anisotropic regularization over the amount of regularization g (abscissa) from our leave-one-out experiment with 18 cases (only 17 tested, the reference image for atlas building not tested). The parameter α is set to be 0.2 for all anisotropic tests.

Table 2. Comparison of different methods for automatic cartilage segmentation according to mean (standard deviation) of validation results reported in the respective manuscripts. Best result is in bold.

	Cartilage	DSC	Sensitivity	Specificity
Folkesson *et al*. [2]	Femoral cartilage	77%(8.0%)	80.3%(11.6%)	99.91%(0.03%)
	Tibial cartilage	81%(6.0%)	**86.8%(7.7%)**	99.96%(0.01%)
Glocker *et al*. [4]	Patellar cartilage	**84%(6%)**	**94.06%**	99.92%
Fripp *et al*. [3]	Femoral cartilage	**84.8%(7.6%)**	83.7%(16.2%)	99.9%(0.0%)
	Tibial cartilage	82.6%(8.3%)	82.9%(20.7%)	99.9%(0.0%)
	Patella cartilage	83.3%(13.5%)	82.1%(13.5%)	100.0%(0.0%)
Yin *et al*. [10]	Femoral cartilage	84%(4%)	80%(7%)	**100%(0%)**
	Tibial cartilage	80%(4%)	75%(8%)	**100%(0%)**
	Patellar cartilage	80%(4%)	76%(8%)	**100%(0%)**
Our method	Femoral cartilage	78.2%(5.2%)	**84.3%(6.9%)**	99.8%(0.06%)
	Tibial cartilage	**82.6%(3.8%)**	83.9%(4.0%)	99.8%(0.06%)

ferent stages of OA separately.

A. Numerical solution

This section discusses an iterative scheme to optimize (3). We introduce two dual variables $\mathbf{p}$ and q and rewrite (3) as

$$E(u,\mathbf{p},q) = \int_{\mathcal{D}} \langle \mathbf{p}, \mathbf{G}\nabla_{\mathbf{x}} u\rangle + q\nabla_l u \, d\mathbf{x}dl, \quad \text{subject to } \|\mathbf{p}\| \leq 1,\ |q| \leq c, \tag{6}$$

in which $\langle\cdot,\cdot\rangle$ represents inner products. Minimizing (3) with respect to u is equivalent to minimizing (6) with respect to u and maximizing it with respect to $\mathbf{p}$ and q. The gradient descent/ascent update scheme of (6) is

$$\mathbf{p}_t = -\mathbf{G}\nabla_x u,\ \|\mathbf{p}\| \leq 1, \tag{7}$$

$$q_t = -\nabla_l u,\ |q| \leq c, \tag{8}$$

$$u_t = -\operatorname{div}_{\mathbf{x}}(\mathbf{G}\mathbf{p}) - \nabla_l q. \tag{9}$$

B. Acknowledgements

The authors thank Pfizer Inc. for providing the data from the Pfizer Longitudinal Study (PLS-A9001140).

References

[1] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification (second edition)*. Wiley-Interscience, 2001.

[2] J. Folkesson, E. B. Dam, O. F. Olsen, P. C. Pettersen, and C. Christiansen. Segmenting articular cartilage automatically using a voxel classification approach. *IEEE Transactions on Medical Imaging*, 26(1):106–115, 2007.

[3] J. Fripp, S. Crozier, S. K. Warfield, and S. Qurselin. Automatic segmentation and quantitative analysis of the articular cartilages from magnetic resonance images of the knee. *IEEE Transaction on Medical Imaging*, 29(1):21–27, 2010.

[4] B. Glocker, N. Komodakis, N. Paragios, C. Glaser, G. Tziritas, and N. Navab. Primal/dual linear programming and statistical atlases for cartilage segmentation. *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2007*, LNCS 4792:536–543, 2007.

[5] T. Heimann, B. Morrison, M. Styner, M. Niethammer, and S. Warfield. Segmentation of knee images: A grand challenge. *Proc. MICCAI Workshop on Medical Image Analysis for the Clinic*, pages 207–214, 2010.

[6] J. Kellgren and J. Lawrence. Radiological assessment of osteoarthritis. *Annals of Rheumatic Diseases*, 16(4):494–502, 1957.

[7] C. Reinbacher, T. Pock, C. Bauer, and H. Bischof. Variational segmentation of elongated volumetric structures. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2010.

[8] L. Shan, C. Zach, and M. Niethammer. Automatic three-label bone segmentation from knee MR images. *IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, pages 1325–1328, 2010.

[9] L. Shan, C. Zach, M. Styner, C. Charles, and M. Niethammer. Automatic bone segmentation and alignment from MR knee images. *Medical Imaging 2010: Image Processing. Proceedings of the SPIE*, 7623:76231K–76231K–8, 2010.

[10] Y. Yin, X. Zhang, R. Williams, X. Wu, D. D. Anderson, and M. Sonka. Logismos-layered optimal graph image segmentation of multiple objects and surfaces: cartilage segmentation in the knee joint. *IEEE Transaction on Medical Imaging*, 29(12):2023–2037, 2010.

[11] C. Zach, M. Niethammer, and J. M. Frahm. Continuous maximal flows and Wulff shapes: Application to MRFs. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1911–1918, 2009.