Fully automated segmentation of cartilage from the MR images of knee using a multi-atlas and local structural analysis method

June-Goo Lee, Serter Gumus, and Chan Hong Moon

Department of Radiology, University of Pittsburgh, Pittsburgh, Pennsylvania 15213

C. Kent Kwok

Division of Rheumatology, University of Arizona Arthritis Center, Tucson, Arizona 85716

Kyongtae Ty Bae^{a)}

Department of Radiology, University of Pittsburgh, Pittsburgh, Pennsylvania 15213

(Received 30 August 2013; revised 19 June 2014; accepted for publication 8 August 2014; published 27 August 2014)

Purpose: To develop a fully automated method to segment cartilage from the magnetic resonance (MR) images of knee and to evaluate the performance of the method on a public, open dataset.

Methods: The segmentation scheme consisted of three procedures: multiple-atlas building, applying a locally weighted vote (LWV), and region adjustment. In the atlas building procedure, all training cases were registered to a target image by a nonrigid registration scheme and the best matched atlases selected. A LWV algorithm was applied to merge the information from these atlases and generate the initial segmentation result. Subsequently, for the region adjustment procedure, the statistical information of bone, cartilage, and surrounding regions was computed from the initial segmentation result. The statistical information directed the automated determination of the seed points inside and outside bone regions for the graph-cut based method. Finally, the region adjustment was conducted by the revision of outliers and the inclusion of abnormal bone regions.

Results: A total of 150 knee MR images from a public, open dataset (available at www.ski10.org) were used for the development and evaluation of this approach. The 150 cases were divided into the training set (100 cases) and the test set (50 cases). The cartilages were segmented successfully in all test cases in an average of 40 min computation time. The average dice similarity coefficient was $71.7\% \pm 8.0\%$ for femoral and $72.4\% \pm 6.9\%$ for tibial cartilage.

Conclusions: The authors have developed a fully automated segmentation program for knee cartilage from MR images. The performance of the program based on 50 test cases was highly promising.

© 2014 American Association of Physicists in Medicine. [<http://dx.doi.org/10.1118/1.4893533>]

Key words: osteoarthritis, cartilage, multi-atlas segmentation, knee magnetic resonance imaging

1. INTRODUCTION

Osteoarthritis (OA) is a complex, heterogeneous condition that commonly causes disability in the aging population.^{1,2} The hallmarks of OA pathophysiology are the breakdown of cartilage in joints and the associated changes in underlying bone, which lead to debilitating joint symptoms, deformity, and disability. Recent advances in both magnetic resonance (MR) imaging for serial monitoring of changes over time and quantitative image analysis have increased our understanding of the pathophysiologic changes in articular cartilage and bone in OA. These advances in MR imaging may contribute to the development and evaluation of effective therapeutic strategies that target early structural changes.

Studies have shown that the morphometric assessment of cartilage structure, such as volume, thickness, and area, via MR images provides an accurate and precise measure of OA progression.³ The measurement of these structures requires image processing and segmentation of different tissue regions from the MR image. Segmentation of the structures has been commonly performed using simple manual delineation and semiautomated methods. However, several groups recently proposed “fully automated methods with no

user interaction” for segmenting cartilage and bone from MR images.^{4–8} Folkesson *et al.* proposed k-nearest neighbor framework to perform tissue classification by selecting features such as voxel position, raw and Gaussian smoothed intensities, and intensity derivatives.⁸ Frupp *et al.* presented a segmentation scheme with three-dimensional active shape models.⁷ As well, Yin *et al.* described an optimal multisurface segmentation method that uses intensity distributions of tissue.⁶ Most recently, Lee *et al.* showed a method based on Markov random field optimization of localized region and boundary probabilities acquired employing relevant local shape and appearance information.⁵

These fully automated segmentation methods share a common segmentation hierarchy: larger and easier objects (i.e., bones) are segmented first to aid in the initialization and to provide constraints on the segmentation of more subtle and difficult objects (i.e., cartilages).^{5–8} In order to segment bones, combined global transformations as well as local signal intensity and shape information are learned and extracted from a presegmented training dataset.^{5–7} After the bone segmentation, cartilages are segmented using bone-cartilage interface properties as well as local signal and shape information.^{5–7}

Many of these methods incorporate atlas-based algorithms for segmentation of bones.^{7,9} The use of anatomical atlases for quantitative analysis has shown great promise in brain segmentation¹⁰ and has also been applied to knee segmentation.^{7,9,11} A major disadvantage of atlas-based approaches for segmenting arthritic knee MR images, though, is the large variability of the articular cartilage. Both the potential for bias and lack of correspondence to the target hinder accuracy.^{12,13} This limitation may be overcome, though, by using a multi-atlas segmentation approach,^{14,15} derived from separately registering multiple atlases to the target and resolving voxel-wise label conflicts between the registered atlases through label fusion. While Tamez-Pena *et al.* reported promising results using the multi-atlas based cartilage segmentation algorithm,⁴ they incorporated only a simple label-fusion approach and did not use any further correction methods to address abnormalities, such as osteophytes that form along joint margins. In the label fusion, it is important to find the corresponding position in multiple atlases, which in previous approaches has been accomplished by using the signal intensity of voxels. In our approach, however, we postulated that the added use of a local structural analysis improves cartilage segmentation by exploiting the characteristic thin sheet-like shape of cartilage. To our knowledge, this is the first multi-atlas based approach to segment bone and cartilage simultaneously.

Therefore, the objectives of the study were Eq. (1) to develop a fully automated method for segmentation of cartilage using a hybrid multi-atlas and graph-cut based region adjustment scheme; and Eq. (2) to evaluate the performance of this method. A publicly open database was used both for the development and performance evaluation of the proposed segmentation method.

2. MATERIALS AND METHODS

2.A. MRI data

We downloaded 150 cases of knee MR images from a public open dataset (available at www.ski10.org).¹⁶ The images were acquired at multiple sites on scanners from a variety of vendors. All images were acquired in the sagittal plane with a pixel resolution of $0.4 \times 0.4 \text{ mm}^2$ and a slice thickness of 1 mm. MR field strength was 1.5 T in about 90% of the cases, the remainder was acquired at 3 or 1.0 T. The vast majority of images were scanned by gradient echo based T1-weighted MRI sequence with fat suppression. After the acquisition, all images were segmented interactively by experts at Biomet, Inc. to label and delineate the bone and cartilage of the femur and tibia. The dataset included the original MR images as well as the corresponding segmentation mask that was used as the ground truth segmentation result. We assigned the 150 cases to either the training set (100 cases) or the test set (50 cases).

2.B. Overall procedure

The segmentation scheme consisted of three procedures: multiple-atlas building, applying a locally weighted vote

FIG. 1. A flow-chart showing the overall procedure of the proposed scheme.

(LWV), and region adjustment (Fig. 1). In the atlas building procedure, all training cases were registered to a target image by a nonrigid registration scheme, described in detail below, and the best matched atlases were selected (top panel in Fig. 1). The LWV algorithm was applied to merge the information from these atlases and generate the initial segmentation result (middle panel in Fig. 1). Subsequently, for the region adjustment procedure, the statistical information of bone, cartilage, and surrounding regions was computed from the initial segmentation result. The statistical information directed the automated determination of the seed points inside and outside bone regions for the graph-cut based method. Finally, the region adjustment was conducted by the revision of outliers and the inclusion of abnormal bone regions (bottom panel in Fig. 1).

2.C. Multi-atlas building and selection of *n*-best matched atlases

The segmentation process was initiated by registering MR images of the entire training dataset onto a given test case of MR images. The registration scheme was carried out in two steps. In the first step, an affine registration using Mattes mutual information¹⁷ as a similarity metric was used for a coarse alignment. In the second step, a b-spline deformation was applied to further adjust the outcome of the rigid registration. To implement this multi-atlas building procedure, an open source software, named elastix, was incorporated.¹⁸ We

FIG. 2. Misaligned atlases and correspondence map from local structural analysis. The left-pane shows the segmentation results from MV and the magnified view of ROI. The atlases were misaligned locally around the ROI, producing incorrect segmentation results. On right-pane, 18 images are shown in pairs. The nine images in the green rectangles represent the registration results on a given ROI from atlases. Misalignments are shown in several images. The other nine images in the red rectangles show correspondence maps from structural analysis at the target voxel on the center of cartilage. In these maps, the bright regions represent the center of cartilage. The corresponding points are searched and determined via the proposed local structural analysis.

applied a multiresolution scheme and preprocessing of MR images for faster conversions. The preprocessing included intensity stretching that linearly mapped the intensity range (15%–99.5%) to 1000 discrepant levels, and Gaussian blurring. The final transformation matrix was recorded for the purpose of later transforming the preprocessed anatomical MR and label images back to the target MR image. During the registration process, the final similarity metric values were saved for further sorting. The n -best matched atlases were selected from the sorted list. The optimal number of best atlases will be described in Sec. 3.A.

2.D. LWV with local structural analysis

After the multi-atlas building, we performed the label fusion procedure. To define the problem, consider a target gray-level image represented as the vector, $I \in R^{N \times 1}$, where N is total number of voxels and R represents image space. Let $T \in L^{N \times 1}$, with L representing label space, be the latent representation of the true target segmentation, where $L = \{0, \dots, N_L - 1\}$ is the set of possible labels that can be assigned to a given voxel. N_L is the number of possible labels. Consider a collection of N_A registered atlases with associated intensity values, $A \in R^{N \times N_A}$, and label decisions, $D \in L^{N \times N_A}$. In the multi-atlas based method, the problem is defined as finding T when I , A , and D are known.

Majority vote (MV) can be used as a simple way to generate label fusion. It gathers voting in the label decisions from the same locations to the target voxel, and the label receiving the maximum vote is selected.

$$T_i(\text{MV}) = \operatorname{argmax}_l \sum_{j \in S} v_{ji}, \quad (1)$$

where S is the selected best-matched atlases, and $v_{ji} = \begin{cases} 1, & \text{if } D_{ji} = l \\ 0, & \text{else} \end{cases}$. The index variables are used throughout: i for the target voxel, i' for iteration over the voxels, and j for iteration over the registered atlases.

Despite its simplicity, MV has been shown to result in a reliable segmentation in a number of applications.^{14,19–21} MV, however, might produce poor results when the voxel-by-voxel correlation could not be achieved due to a misalignment of the registered atlases (see left pane in Fig. 2). In this situation, use of a weighted-voting strategy would improve the segmentation accuracy^{22,23} (see examples in right pane in Fig. 2). In a weighted-voting scheme, the probability of correspondence between an atlas voxel and given target voxel is incorporated to compute the target label. The previous approaches used the intensity similarity metrics for the correspondence. However, this approach is inadequate for knee MR images; the MR signal intensity is often variable between the target and the registered atlas images because of magnetic field inhomogeneity and partial volume effect on the cartilage region. Therefore, we implemented a new local weighted-voting scheme with local structural analysis. In this scheme, cartilage was modeled as a thin plate in 3D and incorporated into a local structural analysis.

A Hessian matrix decomposition method was applied to extract the unique intensity structure at a given local volumetric neighborhood. The Hessian matrix is a 3×3 matrix containing second-order derivatives in 3D space. Eigen decomposition of the Hessian matrix produces three eigenvectors and eigenvalues at each voxel. The eigenvectors of a Hessian matrix represent the orientation angles of the local structure at a given voxel, and the eigenvalues are the magnitudes of the second-order derivatives along the orientation directions determined by the corresponding eigenvectors.²⁴ Let $\lambda_1, \lambda_2, \lambda_3$ be the eigenvalues of $H_\sigma(x, y, z)$. A plate-like local structure can be classified by using the basic logical conditions of these eigenvalues.²⁵ We defined the plate-likeness (p) as follows:²⁶

$$p = \lambda_1 \times \left(1 - \frac{|\lambda_2|}{|\lambda_1|}\right) \times \left(1 - \frac{|\lambda_3|}{|\lambda_1|}\right), \quad (2)$$

where $|\lambda_1| \geq |\lambda_2| \geq |\lambda_3|$. In Hessian-based shape analysis, the 3D Gaussian filtering has dual function: noise reduction

FIG. 3. Example of region adjustment procedure. (a) MR image, (b) initial label by MV, (c) computation mask, (d) seed points, (e) graph-cut results, and (f) graph-cut results on femur and tibia.

and the enhancement of continuity in the local intensity field. In this study, we used 1.0 as σ value to effectively enhance cartilage region with 1–2.5 mm thickness.

The probability of correspondence between an atlas voxel and given target voxel ($\alpha_{ji'i}$) is represented as the following formula:

$$\alpha_{ji'i} = \frac{1}{Z_\alpha} \exp \left\{ -\frac{|p(A_{ji'}) - p(I_i)|^2}{\sigma_c^2} \right\} \quad (3)$$

where $p(A_{ji'})$ is the plate-likeness of i' location in j atlas, $p(I_i)$ is the plate-likeness of i location in target image, σ_c is controlling parameter (experimentally selected as 50), and Z_α is the partition function. Z_α enforces the constraint that

$$\sum_{i' \in N(i)} \alpha_{ji'i} = 1, \quad (4)$$

where $N(i)$ is the set of voxels in the search neighborhood of a given target voxel. Through this constraint, $\alpha_{ji'i}$ can be directly interpreted as the probability that voxel i' on atlas j is the latent corresponding voxel to a given target voxel i . Thus, the LWV can be expressed as the following:

$$T_i(\text{LWV}) = \operatorname{argmax}_l \left\{ \sum_{j \in S} \left(\sum_{i' \in N(i)} \alpha_{ji'i} \times v_{ji'} \right) \right\}, \quad (5)$$

where S is the selected best-matched atlases, and $v_{ji'} = \begin{cases} 1, & \text{if } D_{ji'}=l \\ 0, & \text{else} \end{cases}$.

2.E. Region adjustment by graph-cut

To further improve the segmentation following the LWV, particularly in irregular bone regions like osteophytes, we applied the graph-cut based region adjustment method²⁷ (see Fig. 3). In this approach, compared to our previous study,²⁷ the seed points of the graph-cut algorithm were automatically selected by means of statistical information, including the mean and standard deviation. Those were computed in each bone, cartilage, and surrounding region, segmented from the previous LWV procedure. For computational efficiency, a binary mask that included the boundary regions was generated using morphological operations [see Fig. 3(c)]. In this binary mask, the inner seed points were connected to the source and the outer seed points were linked to the sink to form the graph [see Fig. 3(d)]. The capacity of each connected component was assigned by the following formula:

$$L(p, q) = L(q, p) = \exp \left\{ -\frac{(I_p - I_q)^2}{2\sigma_b^2} \right\}, \quad (6)$$

where $L(p, q)$ is the capacity of two adjacent nodes, p and q . σ_b is the error standard deviation in the bone region.

A max-flow-min-cut algorithm²⁸ was applied on the graph to achieve the globally optimal segmentation. This procedure was repeated twice to adjust bone regions in both the femur and tibia [Figs. 3(e) and 3(f)]. After this adjustment, the cartilage region was updated and segmented on the basis of a connected-component and distance analysis. When a segmented cartilage region was disjoined from the bone region, it was removed.

2.F. Optimization of segmentation scheme by parameter sweeping

The developed segmentation scheme had two key parameters: the number of selected atlases (N_A) used for the multi-atlas building and the kernel radius (KR) for LWV with the local structural analysis step. The interaction of these parameters on the cartilage segmentation performance was evaluated in a parameter sweeping experiment. This information can be in turn used to optimize the segmentation scheme. We determined the dice similarity coefficient (DSC) in our segmentation scheme using a total of 55 combinations of the two parameters. N_A was swept over 11 steps (1, 3, 5, 7, 9, 11, 14, 17, 20, 25, 30), while KR was over 5 steps (1 to 9 with 2 voxel intervals). For example, the size of the kernel corresponded to $7 \times 7 \times 3$ at $KR = 3$, because its size for the slice direction was set smaller (i.e., 3) than that for the in-plane (i.e., 7), given the lower resolution of the slice direction.

Out of 100 training dataset cases, 40 cases were randomly selected as the cross validation group and used for this parameter sweeping experiment. The cartilage segmentation for this group was performed using the remaining 60 training dataset cases. The DSC was computed on each case as a performance measure. A performance map was generated by summarizing the relationship between segmentation performance and parameter combinations.

2.G. Evaluation of segmentation performance

After tuning with the optimal parameter setting, the performance of our proposed segmentation scheme was assessed from the test dataset. Our segmented test dataset was uploaded and submitted to the website from which both the training and original test MR image dataset had been downloaded. The official performance result of our segmentation method was posted on the website. The quantitative evaluation of the

FIG. 4. Regions of interest for the evaluation of femoral (left) and tibial (right) cartilages (presented in Ref. 16).

performance was based on four features computed on the femoral and tibial regions of interest described in Fig. 4.¹⁶ Two features were for the bone analysis: average symmetric surface distance (AvgD) and root-mean-square symmetric surface distance (RMSD). The other two features were for the cartilage analysis: volumetric overlap error (VOE) and volumetric difference (VD) generated by use of the formula:

$$\text{VOE} = 100 \times \left(1 - \frac{|S \cap R|}{|S \cup R|}\right), \quad \text{VE} = 100 \times \left(\frac{|S| - |R|}{|R|}\right),$$

where S is a set of voxels on segmented cartilage region and R is a set of voxels on the reference cartilage region.

3. RESULTS

All the cross validation and test cases were segmented successfully with our proposed method. The running machine specification was Windows 7 64 bit (OS), Intel i7-2600K (CPU) and 16 GB (RAM). The average computation time was about 30 min for the atlas building procedure and 10 min for the label fusion and region adjustment procedures.

3.A. Optimization of segmentation scheme by parameter sweeping

Figure 5 shows the result of the optimization procedure. The maximum DSC value (0.66) was acquired with N_A of 17 and KR of 1 (see flag point in Fig. 5). The DSC initially increased steeply with bigger N_A value, but declined gradually beyond the optimal N_A . With the N_A greater than 3, the DSC decreased with an increase in KR.

In our study, the optimal kernel radius was selected as unity which corresponds to $3 \times 3 \times 1$ kernel size. With a larger kernel size, the locational relevance is reduced, but the chance to find higher structural affinity increases. Thus, the optimal kernel radius is determined as a trade-off between the locational

FIG. 5. The performance map created during the parameter sweeping experiment. The trends of segmentation performance (presented in DSC) are displayed as pseudocolored images and isocontour plots as the N_A and KR parameters are swept across their full range. The red corresponds to the maximum DSC, while the blue corresponds to the minimum DSC. The optimum settings for the cartilage segmentation are denoted as the red zones on the maps.

relevance and the chance to find higher structural affinity, as supported by a previous study.²⁹ In this study, kernel sizes of $3 \times 3 \times 3$ or $1 \times 1 \times 1$ were proposed as optimal settings, depending on the applications. A kernel size larger than these would degrade the performance. In our dataset, with the z slice thickness larger than twice of the in-plane spacing (1 vs 0.4 mm), use of smaller kernel radius would not be effective or practical. It would be more effective with data with finer isotropic z -resolution, such as the osteoarthritis initiative dataset.

3.B. Evaluation of segmentation performance

The performance of our final segmentation scheme tuned with the optimal parameter setting is summarized in Table I. Also presented is a comparison with the MV-based scheme that used the same multi-atlas building procedure (incorporated nine selected atlases for label fusion) but without the region adjustment procedure. One-tailed paired t -test was performed to compare the performances based on MV and our proposed method. The segmentation error measures from the final scheme were smaller than those from the MV-based

TABLE I. The results of performance evaluation of cartilage segmentation.

50 test cases	Femur bone		Tibia bone		Femur cartilage		Tibia cartilage	
	AvgD (mm)	RMSD (mm)	AvgD (mm)	RMSD (mm)	VOE (%)	VD (%)	VOE (%)	VD (%)
MV	0.88 ± 0.27	1.34 ± 0.41	0.85 ± 0.40	1.29 ± 0.59	39.8 ± 8.9	-3.6 ± 16.8	35.4 ± 8.8	-2.4 ± 14.0
Proposed	0.63 ± 0.17	1.05 ± 0.32	0.53 ± 0.25	0.90 ± 0.48	28.3 ± 8.0	-5.5 ± 14.2	27.6 ± 6.9	-3.7 ± 13.3
p -value	0.000	0.000	0.000	0.000	0.000	0.068	0.000	0.180

FIG. 6. Two examples from the test cases presenting the comparison between MV and the proposed methods. Each row represents one of the two examples showing the original MR (left) and segmented images in the sagittal with MV (middle) and the proposed method (right). The segmented regions are color coded: red for the femur bone, blue for the tibia bone, green for femur cartilage, and yellow for tibia cartilage. Yellow and red cursors point to the wrongly segmented bone and cartilage regions, respectively.

scheme for the bone segmentation (AvgD and RMSD) as well as the cartilage segmentation (VOE). However, VD remained similar ($p > 0.025$). The performance results of both schemes are posted on the website of the open dataset provider.

3.C. Comparison with MV results

The color mapped segmentation images and the original MR images were generated using ITK-SNAP software.³⁰ Figure 6 shows two examples from the test dataset. Each row represents one of the two examples showing the original MR (left) and segmented images in the sagittal with MV (middle) and the proposed method (right). MV failed in irregular bone regions and localization in cartilage regions, marked by yellow and red cursors. These regions were successfully corrected by the proposed method.

3.D. Representative segmentation results

Figure 7 shows three examples from the test dataset. Each row represents one of the three examples showing the original MR and segmented (color-coded) images in the sagittal and coronal planes. Each of the segmented femoral and tibial bones and cartilages are depicted with different colors. Cartilage and bone regions, including small osteophytes in tibia bones, were accurately segmented and delineated.

Although the majority of cases had an excellent segmentation outcome, some cases were poorly segmented. Several examples of incomplete segmentation are illustrated in Fig. 8. Despite our region adjustment procedure, an osteophyte was incompletely segmented [Fig. 8(a)]. We used a computation mask to reduce processing time. When an osteophyte is lo-

cated far away from the main bone region, though, it may not be covered by the computation mask and thus excluded from the bone segmentation. In addition, when the femur and tibia bones are too closely juxtaposed, the cartilage region between the two bones may be incompletely segmented [Fig. 8(b)]. In our scheme, bone cysts are included as a part of bone. However, when a cyst is contiguous with cartilage or adjacent non-bony tissue, it may be excluded [Fig. 8(c)]. Segmentation was incomplete in bone where there was severe MR artifact caused by surgical metal screws in the bone [Fig. 8(d)].

4. DISCUSSION

In this study, we present a multi-atlas based approach for cartilage segmentation in knee MR images. The underlying hypothesis of our study was that a multi-atlas approach with local structural analysis could effectively segment not only the bone region, but also the cartilage region. To test this hypothesis, we introduced a new label-fusion method with local structural analysis that determined corresponding voxels from multiple atlases to a test target image voxel. This is a rather simple segmentation scheme that is based on the application of weighted convolution around target voxels from registered atlases and the selection of the label with maximum convoluted value. Despite its simplicity, the multi-atlas based segmentation scheme exhibited a satisfactory performance in which the mean average distance of bone surface was less than 0.63 mm and DSC values of cartilages were around 0.72, including severely advanced osteoarthritic cases. Although our performance appears worse than previously published studies with DSC values ranging 0.84–0.89,^{4,7,31}

FIG. 7. Examples from test cases showing segmentation results. Each row represents one of the three examples showing the original MR and segmented images in the sagittal and coronal planes. The segmented regions are color coded: red for the femur bone, blue for the tibia bone, green for femur cartilage, and yellow for tibia cartilage.

those studies included only a small number of cases, from 14 to 20, representing primarily normal subjects. Furthermore, in the previous studies, knee MR images were acquired at higher field and resolution (3 T, 0.35 mm in-plane and 0.7 mm in slice thickness), whereas most of the MR images in our study were at lower field and resolution (1.5 T, 0.4 mm in-plane and 1 mm in slice thickness). We are currently in the process of applying our method to a different set of MR images that were acquired with higher resolution from a well-defined patient cohort.

For the quantitative evaluation of the performance, it is recommended that the dataset be divided into three independent groups: training, cross validation, and test sets. The cross validation set is usually used to optimize operational parameters.

However, this was not practiced in previous studies, likely because of the small size of the dataset. Rather, the performance of their segmentation methods was assessed by using a leave-one-out method.^{4,7} In our study, with a large number of training cases, we were able to use 40 cases as the cross validation set. We believe this approach strengthens and improves the rigor for the development and performance test of our proposed scheme.

Our segmentation scheme provides an important advantage; the ability to inspect the internal operation and optimize operational parameters that would significantly contribute to the improved performance of image analysis. The inspection of the internal operations in a segmentation scheme allows an understanding of the roles and interactions of the internal

FIG. 8. Cases of incomplete segmentation: (a) an osteophyte (arrow) segmented incompletely and was excluded from the bone region; (b) a cartilage region (arrow) was excluded where the femur and tibia bones were closely juxtaposed; (c) a bone cyst contiguous with cartilage region was segmented as cartilage not bone; and (d) incomplete bone segmentation because of a MRI artifact related to surgical metal screws in the bone.

parameters on the final segmentation performance. The segmentation procedure utilizes numerous internal parameters in processing stages from multi-atlas building to region adjustment. Ideally, all segmentation parameters should be optimized for best performance. However, because the computational burden grows exponentially as parameter numbers increase, many parameters are frequently fixed at empirically determined values, and only the parameters related to the newly devised label fusion are tuned using training datasets.

Our atlas selection method following the registration was based on the similarity measures, which has been proven effective in previous studies.^{20,32} Alternatively, some investigators have proposed to reduce computation time using locally different number of atlases.³³ One interesting property of our multi-atlas building method was its ability to select the desired number of atlases from fast rigid registration results. According to the parameter optimization result (Fig. 5), the optimal parameter set was found at 17 atlases and a $3 \times 3 \times 1$ kernel size. When the number of atlases exceeds the optimal number, the performance deteriorates. The same trend was also observed in a previous study using the atlas selection method.³⁴ Thus, our scheme that allowed the selection of the subset of registered atlases is an important feature ensuring the scalability and adaptiveness of the overall method. The rigid registration is relatively fast to compute, so the inclusion of a large diverse dataset remains easy. This inclusion of a specific dataset, especially one that includes severe OA cases, can get more adapted segmentation results on the severe cases.

Most of the previous multi-atlas approaches used majority voting for label fusion and intensity to find correspondence to the target. In these approaches, any misalignment from the registration scheme degrades the segmentation outcome. This degradation effect is particularly significant in the segmentation of cartilage, which is more sensitive and vulnerable because of its thin plate shape and variable signal intensities. We applied a Hessian-based shape analysis instead of using intensity values and devised a new LWV with this analysis to adjust for misalignment in atlases and intensity variation.

Our study has also limitations. First, there were two methods that showed better performance on the same dataset.^{35,36} They incorporated the statistical shape models to segment cartilage with shorter computation time (6–15 vs 40 min). However, these methods based on active-shape models require initial models. Our current method was implemented without a statistical shape model to be more flexible and scalable. We are investigating a new approach such as a parallel or cloud computing scheme that may allow us to substantially reduce computation time. Second, the patellar cartilage was not segmented from MR images because the open dataset at this point does not provide any training or test sets with the reference segmentation of patellar cartilage. When the reference for the patellar cartilage and bone becomes available, our scheme can be readily extended to the segmentation of these regions. Third, the high quality of the reference training dataset is essential for the development of a robust segmentation program. Our institutional expert radiologists disagreed with some of the reference segmentations of the train-

ing dataset available from the open dataset. Establishment of a large, well-controlled, well-accepted reference training dataset is an extremely important, but challenging task. It is also critical that the training dataset contains a broad range of OA severity so that is sufficient for the development of a segmentation program appropriate for processing a test dataset with varying OA severity.

5. CONCLUSION

We developed, optimized, and evaluated a new, fully automated cartilage segmentation scheme from knee MR images by utilizing the multi-atlas and the local structural analysis. The quantitative evaluation was performed on 50 independent test cases and showed a promising result.

ACKNOWLEDGMENTS

The work was supported partly by the Radiological Society of North America (RSNA) Research Seed Grant No. RSD1320 and NIH/NIAM Grant No. HHSN268201000021.

^{a)} Author to whom correspondence should be addressed. Electronic mail: baek@upmc.edu

¹ D. T. Felson, "An update on the pathogenesis and epidemiology of osteoarthritis," *Radiol. Clin. North Am.* **42**(1), 1–9 (2004).

² P. Sarzi-Puttini et al., "Osteoarthritis: An overview of the disease and its treatment strategies," *Semin. Arthritis Rheum.* **35**(1 Suppl. 1), 1–10 (2005).

³ F. Eckstein, F. Cicuttini, J.-P. Raynaud, J. C. Waterton, and C. Peterfy, "Magnetic resonance imaging (MRI) of articular cartilage in knee osteoarthritis (OA): Morphological assessment," *Osteoarthritis Cartil.* **14**, 46–75 (2006).

⁴ J. G. Tamez-Peña et al., "Unsupervised segmentation and quantification of anatomical knee features: Data from the osteoarthritis initiative," *IEEE Trans. Biomed. Eng.* **59**(4), 1177–1186 (2012).

⁵ S. Lee, S. H. Park, H. Shim, I. D. Yun, and S. U. Lee, "Optimization of local shape and appearance probabilities for segmentation of knee cartilage in 3-D MR images," *Comput. Vis. Image Underst.* **115**(12), 1710–1720 (2011).

⁶ 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 Trans. Med. Imaging* **29**(12), 2023–2037 (2010).

⁷ J. Frapp, S. Crozier, S. K. Warfield, and S. Ourselin, "Automatic segmentation and quantitative analysis of the articular cartilages from magnetic resonance images of the knee," *IEEE Trans. Med. Imaging* **29**(1), 55–64 (2010).

⁸ J. Folkesson, E. B. Dam, O. F. Olsen, P. C. Pettersen, and C. Christiansen, "Segmenting articular cartilage automatically using a voxel classification approach," *IEEE Trans. Med. Imaging* **26**(1), 106–115 (2007).

⁹ T. G. Williams et al., "Anatomically corresponded regional analysis of cartilage in asymptomatic and osteoarthritic knees by statistical shape modelling of the bone," *IEEE Trans. Med. Imaging* **29**(8), 1541–1559 (2010).

¹⁰ D. L. Collins, C. J. Holmes, T. M. Peters, and A. C. Evans, "Automatic 3-D model-based neuroanatomical segmentation," *Hum. Brain Mapp.* **3**(3), 190–208 (1995).

¹¹ J. Carballido-Gamio and S. Majumdar, "Atlas-based knee cartilage assessment," *Magn. Reson. Med.* **66**(2), 575–581 (2011).

¹² J. Ashburner and K. J. Friston, "Unified segmentation," *Neuroimage* **26**(3), 839–851 (2005).

¹³ X. Han and B. Fischl, "Atlas renormalization for improved brain MR image segmentation across scanner platforms," *IEEE Trans. Med. Imaging* **26**(4), 479–486 (2007).

¹⁴ R. A. Heckemann, J. V. Hajnal, P. Aljabar, D. Rueckert, and A. Hammers, "Automatic anatomical brain MRI segmentation combining label propagation and decision fusion," *Neuroimage* **33**(1), 115–126 (2006).

¹⁵ T. Rohlfing, D. B. Russakoff, and C. R. Maurer, Jr., "Performance-based classifier combination in atlas-based image segmentation using

- expectation-maximization parameter estimation," *IEEE Trans. Med. Imaging* **23**(8), 983–994 (2004).
- ¹⁶T. Heimann, B. Morrison, M. Styner, M. Niethammer, and S. Warfield, "Segmentation of knee images: A grand challenge," in *Proceedings of MICCAI Workshop on Medical Image Analysis for the Clinic* (CreateSpace, Seattle, 2010), pp. 207–214.
 - ¹⁷D. Mattes, D. R. Haynor, H. Vesselle, T. K. Lewellyn, and W. Eubank, "Nonrigid multimodality image registration," *Proc. SPIE* **4322**, 1609–1620 (2001).
 - ¹⁸S. Klein, M. Staring, K. Murphy, M. A. Viergever, and J. P. W. Pluim, "Elastix: A toolbox for intensity-based medical image registration," *IEEE Trans. Med. Imaging* **29**(1), 196–205 (2010).
 - ¹⁹P. Aljabar, R. A. Heckemann, A. Hammers, J. V. Hajnal, and D. Rueckert, "Multi-atlas based segmentation of brain images: Atlas selection and its effect on accuracy," *Neuroimage* **46**(3), 726–738 (2009).
 - ²⁰T. Rohlfing, R. Brandt, R. Menzel, C. R. Maurer, Jr., and C. R. Maurer, "Evaluation of atlas selection strategies for atlas-based image segmentation with application to confocal microscopy images of bee brains," *Neuroimage* **21**(4), 1428–1442 (2004).
 - ²¹T. Rohlfing and C. R. Maurer, Jr., "Shape-based averaging," *IEEE Trans. Image Process.* **16**(1), 153–161 (2007).
 - ²²X. Artaechevarria, A. Munoz-Barrutia, and C. Ortiz-de-Solorzano, "Combination strategies in multi-atlas image segmentation: Application to brain MR data," *IEEE Trans. Med. Imaging* **28**(8), 1266–1277 (2009).
 - ²³M. R. Sabuncu, B. T. T. Yeo, K. Van Leemput, B. Fischl, and P. Golland, "A generative model for image segmentation based on label fusion," *IEEE Trans. Med. Imaging* **29**(10), 1714–1729 (2010).
 - ²⁴Q. Li, S. Sone, and K. Doi, "Selective enhancement filters for nodules, vessels, and airway walls in two- and three-dimensional CT scans," *Med. Phys.* **30**(8), 2040–2051 (2003).
 - ²⁵Y. Sato *et al.*, "Tissue classification based on 3D local intensity structures for volume rendering," *IEEE Trans. Vis. Comput. Graph.* **6**(2), 160–180 (2000).
 - ²⁶J.-G. Lee, J. Hyo Kim, S. Hyung Kim, H. Sun Park, and B. Ihn Choi, "A straightforward approach to computer-aided polyp detection using a polyp-specific volumetric feature in CT colonography," *Comput. Biol. Med.* **41**(9), 790–801 (2011).
 - ²⁷H. Shim, S. Chang, C. Tao, J. Wang, C. K. Kwok, and K. T. Bae, "Efficient and reproducible segmentation of knee cartilage from high-resolution MR imaging using a semi-automated graph-cuts algorithm method," *Radiology* **251**(2), 548–556 (2009).
 - ²⁸Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Trans. Pattern Anal. Mach. Intell.* **26**(9), 1124–1137 (2004).
 - ²⁹A. J. Asman and B. A. Landman, "Non-local statistical label fusion for multi-atlas segmentation," *Med. Image Anal.* **17**(2), 194–208 (2013).
 - ³⁰P. A. Yushkevich *et al.*, "User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability," *Neuroimage* **31**(3), 1116–1128 (2006).
 - ³¹P. Dodin, J.-P. Pelletier, J. Martel-Pelletier, and F. Abram, "Automatic human knee cartilage segmentation from 3D magnetic resonance images," *IEEE Trans. Biomed. Eng.* **57**(11), 2699–2711 (2010).
 - ³²S. Klein, U. A. van der Heide, I. M. Lips, M. van Vulpen, M. Staring, and J. P. W. Pluim, "Automatic segmentation of the prostate in 3D MR images by atlas matching using localized mutual information," *Med. Phys.* **35**(4), 1407–1417 (2008).
 - ³³E. M. van Rikxoort *et al.*, "Adaptive local multi-atlas segmentation: Application to the heart and the caudate nucleus," *Med. Image Anal.* **14**(1), 39–49 (2010).
 - ³⁴P. Aljabar, R. Heckemann, A. Hammers, J. Hajnal, and D. Rueckert, "Classifier selection strategies for label fusion using large atlas databases," *Proc. Med. Image Comput. Comput. Interv.* **4791**, 523–531 (2007).
 - ³⁵G. Vincent, C. Wolstenholme, I. Scott, and M. Bowes, "Fully automatic segmentation of the knee joint using active appearance models," in *Proceedings of MICCAI Workshop on Medical Image Analysis for the Clinic* (CreateSpace, Seattle, 2010), pp. 224–230.
 - ³⁶H. Seim, D. Kainmueller, H. Lamecker, M. Bindernagel, J. Malinowski, and S. Zachow, "Model-based auto-segmentation of knee bones and cartilage in MRI data," in *Proceedings of MICCAI Workshop on Medical Image Analysis for the Clinic* (CreateSpace, Seattle, 2010), pp. 215–223.