Structure-enhanced local phase filtering using L_0 gradient minimization for efficient semiautomated knee magnetic resonance imaging segmentation

Mikhiel Lim
Ilker Hacihaliloglu

Structure-enhanced local phase filtering using L_0 gradient minimization for efficient semiautomated knee magnetic resonance imaging segmentation

Mikhiel Lim and Ilker Hacihaliloglu*

Rutgers, The State University, Department of Biomedical Engineering, 599 Taylor Road, Piscataway, New Jersey 08854, United States

Abstract. The segmentation of bone surfaces from magnetic resonance imaging (MRI) data has applications in the quantitative measurement of knee osteoarthritis, surgery planning for patient-specific total knee arthroplasty, and its subsequent fabrication of artificial implants. However, due to the problems associated with MRI imaging, such as low contrast between bone and surrounding tissues, noise, bias fields, and the partial volume effect, segmentation of bone surfaces continues to be a challenging operation. A framework is presented for the enhancement of knee MRI scans prior to segmentation in order to obtain high contrast bone images. During the first stage, a contrast enhanced relative total variation regularization method is used in order to remove textural noise from the bone structures and surrounding soft tissue interface. This salient bone edge information is further enhanced using a sparse gradient counting method based on L_0 gradient minimization, which globally controls how many nonzero gradients are resulted in order to approximate prominent bone structures in a structure-sparsity-management manner. The last stage of the framework involves incorporation of local phase bone boundary information in order to provide an intensity invariant enhancement of contrast between the bone and surrounding soft tissue. The enhanced images are segmented using a fast random walker algorithm. Validation against expert segmentation was performed on 20 clinical knee MRI volumes and achieved a mean dice similarity coefficient of 0.949. © 2016 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: [10.1117/1.JMI.3.4.044503](https://doi.org/10.1117/1.JMI.3.4.044503)]

Keywords: segmentation; knee osteoarthritis; smoothing; local phase; total variation regularization.

Paper 16050PRR received Mar. 28, 2016; accepted for publication Nov. 4, 2016; published online Dec. 2, 2016.

1 Introduction

Osteoarthritis (OA) is a degenerative disease affecting the joint cartilage and bone. Affecting more than 10% of the population, OA is commonly found in the spine, hip, knee, hands, and feet.¹ Given the increasing age of the population, this disease is becoming more common. OA is characterized by joint stiffness and pain due to degraded cartilage causing bone on bone contact. While the cause of OA is unknown, factors such as genetics, obesity, misalignment of articular surfaces, joint trauma, and metabolic diseases contribute to the pathogenesis of the disease.² Cartilage is composed mostly of an extracellular matrix, rich in collagen type-II and aggrecans. Degradation occurs through loosening of the collagen network or proteoglycan loss, which can be caused by mechanical stress and enzymatic degradation. The regenerative ability of cartilage is limited due to its avascular nature and low replenishment rate of collagen type-II.² Degradation evoked by this disease is observed using magnetic resonance imaging (MRI), a widely available and well accepted noninvasive imaging technique.^{3,4}

As the disease progresses, total knee arthroplasty (TKA) may be necessary to restore function and alleviate pain. In this procedure, the damaged knee is replaced with an artificial knee. Various implant designs are available that take patient factors into account, such as age and weight. Despite these advancements, the implants are still ultimately mass-produced. Similar to a personalized medicine model, an implant specifically

designed for a patient would be an even better fit for their needs. It has recently been suggested that MRI scans of the knee provide results that are comparable to computed tomography scans for patient-specific TKA.^{5–9} Nonetheless, MRI has some problems as well, such as low contrast between bone and surrounding tissue, and noise. The development of algorithms that allow for automatic and semiautomatic segmentation of MRIs has, therefore, been the focus of much research. The current gold standard of segmentation methods is manual segmentation, but this takes several hours per three-dimensional (3-D) volume, affecting cost and efficiency.

Various methods for automated MR bone segmentation, based on intensity and texture information, have been proposed.^{10–16} However, these methods remain sensitive to initialization, and cases where the boundaries between the trabecular bone and other tissues, such as cartilage, are blurred results in segmentation of the bone leaking into other tissues. Model-based methods for the segmentation of bone and cartilage have also been used, which segment an image using statistical appearance models that deform to approximate the shape of the femur and tibia.^{17–20} Atlas-based approaches, which take into account anatomical variations in a population, are also available.^{21,22} A graph cut approach has also been developed and was successful for segmentation of individual slices, but only a few consecutive slices were used to create a 3-D model.²³ An automatic three-label approach has also been proposed to separately segment the femur and tibia, which is useful

*Address all correspondence to: Ilker Hacihaliloglu, E-mail: ilker.hac@soe.rutgers.edu

for pathological OA cases, where the bones are close together or touching, but fewer than 10 images were validated.²⁴ Phase information of the raw MR image data, in addition to the magnitude (intensity) image information, has also been used to improve segmentation of knee bones, although reported results were not very accurate and the final outputs need to be refined.²⁵

Intensity-based segmentation methods often require high contrast images to be successful. This is accomplished through prefiltering, but sometimes simple methods such as histogram equalization are inadequate. Thus, in this work, we propose an image enhancement framework for the improved segmentation of knee MRI scans. Our proposed method is a two-dimensional (2-D) analysis and consists of three steps: extraction of femur and tibia bone structure from texture noise using relative total variation (RTV) regularization, sharpening of major bone edges by a gradient counting method based on L_0 gradient minimization, and edge enhancement by intensity invariant local phase filtering. After enhancing all slices in a volume, the output filtered images are then segmented using a random-walker based algorithm and stacked to produce 3-D volumes. In order to show the improvement in segmentation accuracy, using the enhanced images, 20 framework-filtered 3-D volumes were validated against 3-D volumes resulting from expert segmentation of unfiltered images, provided in the SKI10 datasets, by calculating the dice similarity coefficient (DSC).

2 Methodology

The data used in this work was from the SKI10 Grand Challenge dataset, which contained 100 knee MR volume images with corresponding expert segmented labels.²⁶ The images from this dataset were obtained by the surgical planning program of Biomet, Inc. from various locations in the United States, 90% of which used scanners with a field strength of 1.5 T, with the rest at 3 T and 1 T. The majority of the images were acquired through T1-weighted sequences, while some used T2-weighted

Fig. 1 Example of unfiltered image. Red arrows indicate blurred area between bone and tissue.

and gradient echo sequences. Fat suppression techniques were also used, and no contrast agents were used. The resulting images were in the sagittal plane and had a pixel spacing of 0.4×0.4 mm and a 1-mm slice distance. The images were processed slice-by-slice by experts in Biomet, Inc., to create the expert segmented labels and were evenly and randomly divided into three training datasets and two test datasets.²⁶ Images from the three training datasets were used in this work.

Figure 1 shows an example image from the dataset. The unfiltered image has low contrast and as can be seen along the anterior side of the femur, the boundary between bone and muscle or fat is difficult to distinguish. Segmentation at this stage would be possible but difficult, and most likely yield inaccurate results. Filtering operations are thus needed to improve contrast and make segmentation more robust. The proposed prefiltering framework consisted of three stages that enhanced the femur and tibia boundaries in order to increase the accuracy of subsequent image segmentation. The flowchart of the proposed method is provided in Fig. 2. The proposed framework was applied to each 2-D slice of the volume.

Before using any filtering, the image intensity values were first normalized to $[0, 1]$. This was done because every image had different intensity ranges, so the filtering processes may affect some images stronger than others. Normalizing

Fig. 2 Framework flowchart.

preserved the distance between any two values and ensured that the prefiltering framework affected all slices equally. Next, the complement, or negative of the image, was taken for most images and made the bone color dark or light. This was purely a visual effect and did not affect any filtering algorithms. The user may find this helpful to see the edge boundary clearer as the segmentation method used requires interaction from the user.

2.1 Total Variation Regularization

Texture in imaging refers to regular or repeated surface patterns. An example found in nature is the fine detail in wood or stone; grains, dust, and speckle are other sources of textural noise in images. Structures, like the edges of important objects, are often “hidden” behind this textural noise, making edge extraction difficult. MRI is not immune to this type of interference, as the signal can be corrupted by the patient or electrical components.

The first stage in this framework was to enhance bone edge structures by removing textural noise through RTV regularization.²⁷ This algorithm consisted of solving, for a user-defined number of iterations, the linear equation:

$$(\mathbf{1} + \lambda L^t) \cdot v_S^{t+1} = v_I, \quad (1)$$

where v_I and v_S are the vector representations of the input textured image and output structure image, respectively. $\mathbf{1}$ is the identity matrix, t is the iteration, and λ is the strength parameter. L is a weighting matrix based on v_S^t and is the RTV regularizer. RTV is calculated pixelwise on pixel q in $R(p)$, a rectangular window centered on pixel p . It is given by the equation:

$$L = \frac{\sum_{q \in R} g_{p,q} \cdot |(\partial S)_q|}{\left| \sum_{q \in R} g_{p,q} \cdot (\partial S)_q \right| + \epsilon}, \quad (2)$$

$$g_{p,q} \propto \exp \left[-\frac{(x_p - x_q)^2 + (y_p - y_q)^2}{2\sigma^2} \right], \quad (3)$$

where ϵ is a small term to prevent division by zero and $g_{p,q}$ is a weighting function based on spatial affinity. $g_{p,q}$ is scaled by σ , a parameter that determines the size of the window $R(p)$ and thus the maximum size of texture elements regularized. The numerator and denominator in L represent windowed total variation and windowed inherent variation, respectively.²⁷ The distinction between the two variation types is the position of the absolute value within the equation; unlike the total variation term, the inherent variation sum takes into account the direction of the gradient as it can be positive or negative.

Although the data had been scaled, there was still little contrast as most of the intensity values were found to be closer to 0. Therefore, as an intermittent step, histogram equalization was applied next on the images, which resulted in a more uniform distribution of intensity values and enhanced contrast. Histogram equalization was achieved using the contrast-limited adaptive histogram equalization (CLAHE) method.²⁸ The technique was applied after RTV regularization because texture noise had been removed before it could be amplified by histogram equalization (Fig. 10). Along with scaling, this technique ensured that the subsequent filtering operations affected all images similarly. The outcome of this step on an MRI scan can be seen in Fig. 3(b).

2.2 L_0 Gradient Minimization

The result of RTV regularization yielded bone boundaries that were easier to see, but there were still some noise artifacts and variation in the form of gradients present in the image. This occurred by the boundary of bone and soft tissue. Further processing was needed to enhance salient bone edges while minimizing edges created by noise. Thus, in the second stage, the output of RTV regularization and histogram equalization served as the input to an edge enhancement technique that used a sparse gradient counting method based on L_0 gradient minimization.²⁹

This method globally controlled the number of nonzero gradients that resulted in order to approximate prominent structures such as bone edges. Small amplitude gradients can be seen as weak edges and can be due to noise, while high amplitude gradients are strong, structurally relevant edges. Restricting the number of nonzero gradients effectively removed the edges created by noise. Reduction of small amplitude gradients also had the effect of increasing the steepness of transitions, thus enhancing salient edges. In effect, this algorithm was a smoothing operation that preserved edge information. The output image S was calculated by iteratively solving the equation:

$$S = \mathcal{F}^{-1} \left\{ \frac{\mathcal{F}(\delta) + \beta[\mathcal{F}(\partial_x)^* \mathcal{F}(h) + \mathcal{F}(\partial_y)^* \mathcal{F}(v)]}{\mathcal{F}(\delta) + \beta[\mathcal{F}(\partial_x)^* \mathcal{F}(\partial_x) + \mathcal{F}(\partial_y)^* \mathcal{F}(\partial_y)]} \right\}, \quad (4)$$

where δ is the delta function, $\mathcal{F}$ is the Fourier Transform, and $\mathcal{F}(\cdot)^*$ is the complex conjugate. h and v are the gradients in the x - and y -directions for each pixel p and are defined as

$$(h_p, v_p) = \begin{cases} (0,0) & (\partial_x S_p)^2 + (\partial_y S_p)^2 \leq \lambda/\beta, \\ \nabla S_p = (\partial_x S_p, \partial_y S_p) & \text{otherwise} \end{cases}, \quad (5)$$

where ∇ is the gradient operator, λ is a strength parameter, and β is an iteratively growing parameter. In each iteration, the number of nonzero gradients was reduced, and the algorithm stopped when β was greater than a specified $\beta_{\max}$. When applied to the MRI data, the relevant parameters were the smoothing weight λ , and κ , which controlled the increased rate of β . The output of this stage can be seen in Figs. 3(c) and 11.

One of the noted drawbacks of this method was that some noise may not be fully smoothed out if it had high amplitude, which can occur with textural noise. Xu et al.²⁹ suggested the use of filtering that reduces the amplitude of noise-like structures prior to application of L_0 gradient minimization. RTV regularization was, therefore, applied first, as it removed texture noise. Furthermore, this step occurred after histogram equalization since without it, the intensity values would be in a small range, so weak noise edges and salient bone edges would have similar intensities and would have been smoothed equally.

2.3 Local Phase-Based Edge Filtering

The result of L_0 gradient minimization yielded images that provided contrast between bone and surrounding tissues. Segmentation would be possible at this stage. However, since the previous step was essentially a smoothing operation, some degree of blurring occurred at the boundaries. This can be seen by comparing the plots of the pixel intensity values along a line [Figs. 11(i) and 11(ii)], which shows that the gradients of the salient edges have decreased, indicating that the edges are slightly blurred. Local phase-based image enhancement

Fig. 3 Example result of the framework. (a) The original image. (b) The image after noise texture removal and histogram equalization. (c) Effect of L_0 gradient minimization on (b). (d) Effect of local phase based edge filter on (c). (e) Segmentation result that has been cropped to the femur and tibia. (f) The corresponding expert labeled image cropped to the same size as (e). (g) Overlay of (e) and (f); the white indicates areas, where both images overlap perfectly, while the gray are areas of overshoot from either (e) or (f).

methods have recently gained attention from the medical imaging community due to their intensity invariant feature enhancement properties.^{30,31} Thus, to provide more robust segmentation results for the blurred boundary regions between the bone and soft tissues like cartilage, the final stage of our framework was the enhancement of bone edge information using a local phase-based edge filter. This was accomplished by first converting the image signal into a monogenic signal, the 2-D version of the analytic signal, which is a signal orthogonally decomposed into energetic (local amplitude), structural (local phase), and geometric (local orientation) information.³² The filter then uses the phase information to highlight symmetry in an image. The algorithm did this by applying even-symmetric and odd-symmetric log Gabor filters at different scales, which produce a response at points of symmetry and asymmetry, respectively.³³ The log-Gabor filter transfer function LG is defined as

$$\text{LG}(\omega, \theta) = \exp \left[\frac{-\log(\omega/\kappa)^2}{2 \log(\sigma_\omega)^2} \right] \cdot \exp \left[\frac{-(\theta - \theta_m)^2}{2(\sigma_\theta)^2} \right], \quad (6)$$

where ω and θ are related to the scale and orientation of the filter. κ is the center frequency of the filter, σ_ω is related to the spread of the frequency spectrum in a logarithmic function, and σ_θ defines the angular bandwidth of the filter. θ_m is the specific orientation of the filter. Symmetry in the image was measured by equation:

$$\text{Sym} = \frac{\sum_n |e_n(x, y)| - |o_n(x, y)| - T}{\sum_n A_n(x, y) + \epsilon}, \quad (7)$$

where n is the filter scale, e_n is the even-symmetric filter output, o_n is the odd-symmetric filter output, and T is a compensation term for the noise response. This is normalized by A_n , the magnitude of the filter response, and ϵ , a small constant to prevent division by zero. The algorithm highlighted components that exhibit symmetry, such as edges that create the femur and tibia bones by measuring the filter responses at different orientations; the edges are enhanced when they are orthogonal to the filter orientation.³⁴ By having a number of orientations, sections of curved edges are enhanced as they become orthogonal to the filter orientation. This method is independent of contrast. The outcome of this step on an MR image can be seen in Fig. 3(d).

2.4 Random Walker Image Segmentation

The output of the framework was structure-enhanced filtered images. As such, a variety of image segmentation methods were applicable, like k -means clustering and region growing. Segmentation through these methods was somewhat successful and was inconsistent throughout the slices; a more robust method was needed. The image segmentation technique used in this work was a random walker (RW) algorithm, a graph-based method, and was applied to the enhanced images.³⁵

The algorithm first converted an input image into a graph, where the pixels correspond to nodes connected by edges. The edges between nodes were weighted according to image intensity and gradients, where a high gradient was weighted more. In effect, the weighting approximated image structure. A user places seeds on a pixel to categorize it as foreground or background. The unlabeled pixels release a RW and are categorized depending on which seed the RW has the highest probability of arriving first. For each label, the vector of RW probabilities for each unlabeled pixel x_U was calculated by

$$(L_U + \gamma I_U)x_U = -B^T x_S + \gamma \lambda, \quad (8)$$

where L is the Laplacian matrix of the graph, I is the identity matrix, x is the vector of probabilities of each pixel, λ is an optional vector of prior probabilities weighted by γ , and U and S correspond to unseeded and seeded versions, respectively. The RW is biased from crossing sharp gradients due to the edge weights, so object boundaries are respected and the image is able to be partitioned. This method was interactive, as it required a user to place the foreground (bone) and background seeds. The RW algorithm outputs an outlined binary image. After initial segmentation, the user is able to evaluate the success of segmentation, and if unsatisfied, is able to place more seeds until a desired result is achieved. Due to the interactive nature of this method, the results would vary according to the number of seeds used. Therefore, in order to reduce this variability, the least number of seeds was used in order to segment an image; an image was considered segmented when a closed boundary approximating the bone region is formed.

2.5 Validation

The results of the framework were compared with expert segmentation labels using the DSC, calculated with the equation:

$$\text{DSC} = 2 \cdot \frac{|A \cap B|}{|A| + |B|}, \quad (9)$$

which counts the number of overlapping nonzero pixels in the resulting image A and expert segmented image B against the number of nonzero pixels in each image. While the RW algorithm produced an outlined image of the edges, the penultimate result, a “block” image wherein each pixel has been categorized as background or foreground, was used to create the bone surfaces. This results in a black/white image, where foreground areas are white on a black background. To isolate the bones, postprocessing was performed; large extraneous foreground areas were removed by image filling them to the background color. Similarly, small stray pixels were removed using a morphological opening. The images were cropped to only include the femur and tibia in order to reduce the amount of postprocessing required to remove stray pixels that could affect DSC calculation. The overall DSC of a 3-D volume was calculated by taking the average of the DSC of every slice. The filtered volumes were viewed using 3-D Slicer.^{36,37} In order to determine improvement due this framework, volumes were created using unfiltered images and the output from each step of the framework. DSCs were calculated for each volume and compared using statistical analysis.

3 Results

The proposed framework was implemented in MATLAB (The Mathworks Inc., Natick, Massachusetts) and 20 subject knees were used. Since there are four volumes for each subject knee corresponding to each framework stage, a total of 80 volumes were produced. An example result on a slice can be seen in Fig. 3 which details the outcomes of the framework steps. In RTV regularization, a strength parameter of $\lambda_{ts} = 0.0005$ or 0.001, depending on initial contrast of the image, a scaling parameter $\sigma = 1$, and four iterations produced images that had homogeneous pixel intensity in the bone area while preserving contrast between the bone and background. The images during L_0 gradient minimization used a smoothing strength $\lambda_L = 0.02$; a rate $\kappa = 1.4$ was found to provide a good balance between processing time and adequate gradient minimization. The log Gabor filters used an angular bandwidth $\sigma_\theta = 25$ deg, a center frequency $\kappa = 0.03$ pixels, and a frequency spectrum spread $\sigma_\omega = 0.45$. Phase-based edge filtering used two filter scales and six orientations; it was also successful in further

Table 1 DSC results on 3-D volumes from each step of the framework.

Dataset	Image #	Unfilt	RTV	L_0	Phase-filtered	DSC nonzeros
Train-A	4	0.882	0.894	0.908	0.912	0.924
	18	0.864	0.883	0.914	0.905	0.943
	26	0.856	0.859	0.899	0.905	0.918
Train-B	1	0.933	0.929	0.952	0.962	0.962
	2	0.887	0.892	0.906	0.909	0.956
	5	0.911	0.917	0.952	0.947	0.959
	6	0.908	0.916	0.936	0.936	0.945
	10	0.919	0.908	0.911	0.910	0.931
	12	0.932	0.925	0.929	0.957	0.957
	13	0.916	0.914	0.935	0.943	0.954
	15	0.896	0.883	0.896	0.905	0.930
	18	0.896	0.881	0.915	0.918	0.951
	19	0.933	0.936	0.952	0.957	0.957
Train-C	20	0.952	0.954	0.960	0.966	0.976
	6	0.945	0.935	0.949	0.955	0.955
	10	0.939	0.940	0.955	0.968	0.968
	12	0.913	0.894	0.928	0.933	0.945
	14	0.922	0.928	0.942	0.938	0.938
	15	0.900	0.898	0.898	0.909	0.939
Average:	16	0.929	0.919	0.952	0.951	0.963
		0.912	0.910	0.929	0.934	0.949

enhancing the bone edges. All the parameters identified in this section were not changed throughout the validation studies. The L_0 operation was the most time consuming part of the framework, which takes about 3 min to filter all the slices in a 3-D volume. Each volume contained about 110 slices, which corresponded to about 1.6 s/slice. In contrast, RW segmentation was faster, taking less than a second to segment a slice although user involvement in placing seeds and postprocessing to isolate the femur and tibia takes the bulk of the time in creating a 3-D volume. The amount of time to fully segment and process a volume dependent on image quality, stage of the framework, and user comfort with the RW algorithm. In general, it was found that unfiltered images took around 30 min for segmentation and 10 min for postprocessing, while after the framework, segmentation took 15 to 20 min and postprocessing took 7 min.

The overall DSC of a 3-D volume was determined by taking the average DSC of every slice. The DSC for volumes created from unfiltered images and from each stage of the framework can be seen in Table 1. The 20 framework-filtered 3-D volumes yielded an average DSC of 0.934 ± 0.023 when slices where DSC was 0 were included. Zero similarity occurred when the expert segmentation slice was composed of cartilage only, while the framework detected the cartilage as an edge. When not counting those slices, mean DSC was 0.949 ± 0.015 . The image enhancement framework results have comparable DSCs to previously proposed methods.^{15,16,23,24} In order to determine if there is an improvement in segmentation accuracy due to the framework, repeated measures ANOVA was performed to compare the mean DSCs of the unfiltered, RTV, L_0 , and phase-filtered groups, and resulted in a statistically significant ($P < 0.05$) difference in means between unfiltered and framework-filtered volumes. The only group pairings that did

not have a statistical significant difference in means were the unfiltered-RTV group pairing and the L_0 -phase filtered group pairing. When also considering the volumes that did not count zero similarity slices, there was a statistically significant ($P < 0.05$) difference in mean DSCs when compared to the framework-filtered volumes.

The segmentation results from the RW algorithm after post-processing can be seen in Fig. 3(e). The success of the framework can be qualitatively observed by overlapping the framework segmentation result with expert segmentation [Fig. 3(g)]. A comparison of RW segmentation on filtered and unfiltered slices, when the same number and location of seeds are used, can be seen in Fig. 4, which demonstrates that the prefiltering framework made segmentation more accurate. Figure 5 shows the segmentation results for multiple slices of the same knee and that the slices have good similarity with expert segmentation. Figure 6 shows the segmentation result from unfiltered images and from each stage of the framework, and demonstrates the increase

Fig. 4 Comparison of RW segmentation on filtered and unfiltered images using the same number and location of background and foreground seeds. (a) Original image, (b) extracted bone edges from framework filtered image, and (c) extracted bone edges from unfiltered image.

Fig. 5 Framework result against expert segmentation on different slices of the same knee. (a)–(d) Original input images. (i) Framework segmentation. (ii) Expert segmentation. (iii) Overlay of (i) and (ii); white are areas of perfect overlap while gray are areas of overshoot from either.

Fig. 6 Segmentation result from unfiltered images and from each stage of the framework. Results come from using the fewest number of seeds to create a closed boundary approximating the bone area. (a) Original Image, (b) segmentation result from unfiltered image, (c) segmentation result after RTV, (d) segmentation result after L_0 , (e) segmentation result after phase filtering, and (f) expert segmentation.

in accuracy due to the framework. This is supported by Fig. 7, which shows the calculated DSC per slice in a 3-D volume. This is an indication of the overall performance of segmentation with the framework and is compared to segmentation without filtering. Since it appeared that there was more consistent DSC due to the framework, Levene's Test on slice-by-slice DSC was used, due to the data not following a normal distribution, and determined that there is no difference in variance between the filtered and unfiltered versions of a volume. Figure 7 also reveals a trend where the framework is not accurate in the beginning and end slices but works well in the middle slices. Since outliers greatly affect variance, when these data points, often the marginal slices, were removed, Levene's Test indicated that there was a statistically significant ($P < 0.05$) difference in variance between unfiltered and framework-filtered versions of a volume. This was true for all but 2 of the 20 volumes. This observation is expanded upon by Fig. 8, which shows a volume's DSC trend for each stage in the framework. Figure 9 shows the reconstruction of a filtered 3-D volume.

4 Discussion

4.1 Prefiltering Framework Evaluation

Overall, the three-stage prefiltering framework was successful in providing high contrast images for segmentation. As shown in Fig. 10, the output from RTV regularization yielded preservation of main bone structural edges, while pixel values in the bone area were homogenized as intensity variations were regularized. Compared to the original image, the bone edges are already much clearer. Performing histogram equalization after RTV regularization also helped increase contrast as noise was removed before it could be amplified (Fig. 10). However, further filtering was needed since there were still some slight pixel gradients that can be detected as an edge.

L_0 gradient minimization was the step that contributed most to the success of the prefiltering framework. This can be seen in Table 1, where DSC increased from 0.91 to 0.93 at the L_0 stage. The repeated measures ANOVA indicated that there was no

Fig. 7 Calculated DSC for every slice in a 3-D volume. Framework-filtered DSCs (blue) are more consistent than unfiltered DSCs (red). The framework provides accurate results in the middle slices but is inaccurate in the beginning and end slices. (a) Slice-by-slice DSC for volume TrainB6 and (b) slice-by-slice DSC for volume TrainC16.

Fig. 8 Slice-by-slice DSC according to framework stage. L_0 and phase-filtered groups provide more consistent results than RTV and unfiltered groups.

Fig. 9 Framework results—3-D model. Letters represent the directions of the RAS anatomical coordinate system.

Fig. 10 Histogram equalization effect with and without RTV Regularization. (a) Original input image, (b) histogram equalization on (a), (c) histogram equalization on RTV regularized image; contrast is enhanced while texture noise has been suppressed.

statistically significant difference in means between the unfiltered and RTV groups, but there was a difference between the RTV and L_0 groups. Similarly, there was no statistically significant difference in means after phase filtering, so the L_0 operation had the biggest impact in the framework. This is further supported by Fig. 8, where the DSC trend at the L_0 stage follows the final stage results more closely while the previous stages perform irregularly. The output image from this stage had greatly reduced small amplitude details, mainly minor variations in the skin and bone. The images had smooth and homogeneous

areas, with major edges, such as the bone boundary still preserved [Fig. 3(c)]. This is supported by Fig. 11, which shows the pixel intensity values along a line to demonstrate the L_0 minimization effect. This algorithm was also the most time consuming part of the framework. The parameter κ controlled processing time as it was a multiplier of β , so having a high κ made β reach $\beta_{\max}$ faster, thus ending the algorithm faster. Having a low κ would, therefore, make the image smoother, but processing time would increase greatly. Since multiple slices were processed at a time for efficiency, filtering time had to be

reasonably low. The images were found to be sensitive to the smoothing strength parameter λ ; this is because the algorithm depended on calculating the gradients of the pixels. As seen in Eq. (5), with a higher λ , the minimum gradient amplitude that becomes smoothed increases. Without histogram equalization, the intensity range of the image would be small and λ varied greatly, from 0.003 to 0.015, in order to provide adequate results. Due to histogram equalization, a single λ value, $\lambda = 0.02$, worked for all slices and almost all volumes and had the advantage of allowing the slices to be filtered at once. A small amount of blurring occurred in L_0 minimization, as evidenced by lower amplitude salient gradients [Figs. 11(i) and 11(ii)]. To make segmentation more robust, these edges had to be reinforced.

The local phase-based edge filter helped strengthen and highlight the bone edge information. It was essential that the input image for this filter had a reduced total number of gradients. As can be seen in Figs. 3(c) and 3(d), there is a gradient in middle of the tibia that was detected as an edge by the filter. Without gradient minimization, more edges would have been detected, making segmentation difficult. Two scales and six orientations were used during phase filtering. Since features are enhanced when they are orthogonal to the filter orientation, having this number of orientations allowed sections of curved surfaces to eventually become orthogonal to the filter. While Table 1 shows that phase-filtering resulted in a higher mean DSC than L_0 minimization, the repeated measures ANOVA revealed that there is no statistically significant difference between the two. Nonetheless, although phase filtering does not have an effect on the 3-D volumes as a whole, its overall effect is keeping RW segmentation restrained within the bone boundary. It also helps on a case-by-case basis to prevent leakage on specific slices and does not have high computational cost (Fig. 12). This filter was the most robust part of the framework as it detected edges very well but its success depends on the outcome of previous steps. The optimization of RTV, L_0 , and phase filtering parameters are a subject of future work.

4.2 Image Segmentation

The RW-based algorithm was a reliable method of segmentation due to its interactive nature in that all slices were able to be segmented. Figure 13 shows the DSC distribution of segmented volumes according to framework stage. Unfiltered and RTV volumes have similar performance, although the mean DSC of RTV volumes is actually lower. This is because the RTV operation is a

Fig. 12 Reinforcement of weak edge by phase filtering. (a) Image after L_0 operation; red arrow shows weak edge. (b) Phase filtered image; reinforced edge could prevent leakage during segmentation.

Fig. 13 Boxplot of DSC distribution of volumes according to framework step. Unfiltered and RTV volumes are similar, as are L_0 and phase filtered volumes. There is a narrower distribution at the L_0 stage.

Fig. 11 Effect of L_0 gradient minimization. (a) Image prior to L_0 minimization. (b) After L_0 minimization. (i)–(ii) Corresponding pixel intensity values along the line.

smoothing operation; since it is applied on unfiltered images, there is a degradation effect on the edges. There is an increase in median DSC at the L_0 stage and a narrower distribution. Levene's Test did not result in a statistically significant difference in variance for any groups because outliers have great influence. These outliers are usually the marginal slices, and when removed, Levene's Test show that there is a statistically significant ($P < 0.05$) difference in variance between unfiltered and framework-filtered versions of a volume (Fig. 7).

Marginal slices having low DSC was a trend observed in all volumes and is a drawback of the framework. In the middle slices, where there is high DSC, segmentation is consistent and easier. Since RW segmentation is interactive, to maintain fairness among the framework stages, the fewest number of seeds was used to create a segmented image, with an image considered segmented when a closed boundary approximating the bone region was formed. These results are seen in Fig. 6. Unfiltered and RTV slices required around three foreground seeds and four or more background seeds to "guide" the segmentation to an acceptable result. In contrast, L_0 and phase filtered slices required only one to two background and foreground seeds for segmentation, and had consistent and more accurate results. With less sensitive seed placement, a user can segment images more objectively and at a faster pace. Thus, the framework is able to produce reliable images less prone to user error for RW segmentation.

An advantage of an interactive RW segmentation is that a boundary can be forcibly made with the foreground and background seeds (Fig. 14). This is useful if there are problematic slices where the boundaries are not easily defined. The user can derive where the bone edges are based on context, such as where the edges are expected to be. The user can also look through the slices to see how the shape evolves and determine if it becomes a part of the tibia or femur. Therefore, there is a learning curve to this segmentation method, wherein the user gets better at approximating bone boundary locations after some practice. The total time for segmentation is reduced as well with some training. Although forcibly making boundaries increases accuracy, segmentation can become time consuming and in extreme cases, akin to manual segmentation, since the user is

Fig. 14 Forced creation of a bone edge during RW segmentation. (a) Femur boundary is blurred at the posterior superior end. (b) The user can approximate the correct location of the femur boundary using anatomical knowledge, and place red (foreground) and green (background) seeds to create that edge. Number of seeds exaggerated for visual clarity.

essentially tracing the bone outline to make sure that the result is satisfactory. The goal of the framework was to produce images that are less prone to user error during segmentation. These problematic slices occurred in only 5 to 10 slices out of all 20 volumes.

The entire procedure of filtering, segmentation, and creation of a 3-D model from the framework takes about 15 to 20 min. This is dependent on the user's comfort with the segmentation method. The framework's filtering time is about 3 min/volume; most of the time is devoted to segmentation and postprocessing. In contrast, the process to create a 3-D volume takes over 30 min using unfiltered images. This is an improvement over manual segmentation, which can take several minutes per slice, yet has similar results. About 20 volumes have been segmented using the framework thus far; a more detailed validation on more scans is needed for this method to be clinically acceptable.

4.3 Drawbacks

The main drawback of the framework is that it does not work well in the beginning and end slices. One reason is that the same RTV and L_0 parameters are used for all slices. Although the parameters are good for when the tibia and femur are large, the structures are small in the marginal slices so they look indistinguishable from soft tissue. The other reason is that the framework does not differentiate between cartilage and bone. The marginal slices are composed of mostly cartilage, so the framework will detect the shape created by it and the user will have no way of knowing whether or not it is actually bone unless they have seen the ground truth label. It is not much of a problem in the middle slices because the images are composed of more bone than cartilage. Figure 5 shows the framework segmentation results on different slices of the same knee. The slices in Figs. 5(b) and 5(c) overlap with the expert segmentation almost perfectly. However, segmentation is less accurate by the beginning and end slices [Figs. 5(a) and 5(d)]. The drawback is also reflected by the slice-by-slice DSC trend of a 3-D volume. Marginal slices have low similarity, then DSC increases quickly in the middle slices (Fig. 7). Slices with zero similarity occur when the slice is composed of cartilage only. The repeated measures ANOVA showed that there is a statistical difference in the mean DSC when the zero similarity slices are not included. Furthermore, Levene's Test shows a statistical difference in variance between filtered and unfiltered versions of a volume only when the middle slices are used. Thus, the marginal slices statistically affect the mean DSC and variance of a 3-D volume. Nonetheless, the mean DSC even due to this drawback is >0.90 , which is still very high and indicates good similarity with expert segmentation and the overall success of the framework (Table 1). Figures 7 and 13 also show that framework filtered images have more consistent results than unfiltered images as there is less variation in DSC per slice.

A potential solution to this drawback is by automating filtering parameters based on bone structure information, or structure-adaptive parameter selection, which may preserve bone boundaries in sensitive areas. Another possibility is extending the framework from 2-D to 3-D analysis, which could provide improvements by taking into account shape evolution of the bones to limit drastic differences in DSC. Another drawback is the subjective nature of RW segmentation. Since the seeds are placed by the user, segmentation is stopped when the resulting shape, according to the user, is a good approximation of the

femur or tibia. Although the framework reduces user input in segmentation, the end result is ultimately dependent on the user's opinion. A more objective RW segmentation can be achieved by automating seed placement, a subject of future work.

5 Conclusion and Future Work

The proposed three-step prefiltering framework was successful in providing enhanced, high contrast knee MR scans for segmentation. It first improved contrast between bones and soft tissues by removal of texture like noise from structural edges such as bone, then sharpened bone edge information by L_0 gradient minimization. Last, bone edge information was enhanced by a local phase-based feature extraction. The framework was applied to every slice of a volume and validated on 20 3-D volumes. It resulted in a high mean DSC with a statistically significant improvement in DSC from unfiltered images, indicating that the framework segmented images closely resemble expert labeled images. The 3-D models have been created through RW-based segmentation using the filtered images and have many potential applications, such as providing anatomical information for surgery planning in TKA, and as a basis for a patient-specific 3-D printed artificial knees. This framework can also be applied to other bones, such as the hip for surgery planning in femoroacetabular impingement. The framework is not computationally intensive as evidenced by its low processing time. Its main drawback is low accuracy in marginal edges, which may be avoided by adapting filtering parameters depending on the bone morphology. Future work includes optimization and automation of filtering parameters, automating seed placement for more objective validation, and application of this framework on more datasets and on pathological cases. Since the focus of this work is on the prefiltering process, using a different segmentation algorithm in the future may provide better or faster results.

Disclosures

No conflicts of interest, financial or otherwise, are declared by the authors.

References

1. C. Cooper et al., "Epidemiology of osteoarthritis," *Medicographia* **35**(2), 145–151 (2013).
2. T. Aigner and N. Schmitz, "Pathogenesis and pathology of osteoarthritis," *Rheumatology* **414**, 1741–1759 (2011).
3. F. Eckstein et al., "Magnetic resonance imaging (MRI) of articular cartilage in knee osteoarthritis (OA): morphological assessment," *Osteoarthritis Cartilage* **14**, 46–75 (2006).
4. F. Eckstein, D. Burstein, and T. M. Link, "Quantitative MRI of cartilage and bone: degenerative changes in osteoarthritis," *NMR Biomed.* **19**(7), 822–854 (2006).
5. S. Asada et al., "Comparison of MRI and CT-based patient-specific guides for total knee arthroplasty," *Knee* **21**(6), 1238–1243 (2014).
6. T. Pfizner et al., "Small improvements in mechanical axis alignment achieved with MRI versus CT-based patient-specific instruments in TKA: a randomized clinical trial," *Clin. Orthop. Relat. Res.* **472**(10), 2913–2922 (2014).
7. B. M. Frye et al., "MRI is more accurate than CT for patient-specific total knee arthroplasty," *Knee* **22**(6), 609–612 (2015).
8. P. Stirling et al., "The choice of CT versus MRI for modeling of patient-specific instrumentation in total knee arthroplasty: a review of current evidence," *Bone Joint J.* **98**(Suppl. 4), 28–28 (2016).
9. A. Silva, E. Pinto, and R. Sampaio, "Rotational alignment in patient-specific instrumentation in TKA: MRI or CT?" *Knee Surg. Sports Traumatol. Arthroscopy* **24**(11), 3648–3652 (2014).
10. J. Pang et al., "Automatic articular cartilage segmentation based on pattern recognition from knee MRI images," *J. Digital Imaging* **28**(6), 695–703 (2015).
11. A. Aproxitola and L. Gallo, "Knee bone segmentation from MRI: a classification and literature review," *Biocybern. Biomed. Eng.* **36**(2), 437–449 (2016).
12. L. M. Lorigo et al., "Segmentation of bone in clinical knee MRI using texture-based geodesic active contours," in *Medical Image Computing and Computer-Assisted Intervention*, pp. 1195–1204, Springer, New York (1998).
13. J. Frapp et al., "Automatic segmentation of the bone and extraction of the bone-cartilage interface from magnetic resonance images of the knee," *Phys. Med. Biol.* **52**(6), 1617–1631 (2007).
14. J. Frapp et al., "Segmentation of the bones in MRIs of the knee using phase, magnitude, and shape information," *Acad. Radiol.* **14**(10), 1201–1208 (2007).
15. P. Dodin et al., "A fully automated human knee 3D MRI bone segmentation using the ray casting technique," *Med. Biol. Eng. Comput.* **49**(12), 1413–1424 (2011).
16. P. Dodin et al., "Automatic human knee cartilage segmentation from 3-D magnetic resonance images," *IEEE Trans. Biomed. Eng.* **57**(11), 2699–2711 (2010).
17. G. Vincent et al., "Fully automatic segmentation of the knee joint using active appearance models," in *Proc. MICCAI Workshop on Medical Image Analysis for the Clinic: A Grand Challenge*, pp. 224–230 (2010).
18. J. Schmid and N. Magnenat-Thalmann, "MRI bone segmentation using deformable models and shape priors," in *Medical Image Computing and Computer-Assisted Intervention*, pp. 119–126, Springer, New York (2008).
19. H. Seim et al., "Model-based auto-segmentation of knee bones and cartilage in MRI data," in *Proc. Medical Image Analysis for the Clinic: A Grand Challenge, in conjunction with MICCAI 2010*, pp. 215–223 (2010).
20. T. G. Williams et al., "Automatic segmentation of bones and inter-image anatomical correspondence by volumetric statistical modelling of knee MRI," in *IEEE Int. Symp. on Biomedical Imaging: From Nano to Macro*, pp. 432–435, IEEE (2010).
21. 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).
22. L. Shan et al., "Automatic atlas-based three-label cartilage segmentation from MR knee images," *Med. Image Anal.* **18**(7), 1233–1246 (2014).
23. S. Y. Ababneh, J. W. Prescott, and M. N. Gurcan, "Automatic graph-cut based segmentation of bones from knee magnetic resonance images for osteoarthritis research," *Med. Image Anal.* **15**(4), 438–448 (2011).
24. L. Shan, C. Zach, and M. Niethammer, "Automatic three-label bone segmentation from knee MR images," in *IEEE Int. Symp. on Biomedical Imaging: From Nano to Macro*, pp. 1325–1328, IEEE (2010).
25. P. Bourgeat et al., "Texture-based segmentation of the knee bones in MRI using phase information," in *3rd IEEE Int. Symp. on Biomedical Imaging: Nano to Macro*, pp. 742–745, IEEE (2006).
26. T. Heimann et al., "Segmentation of knee images: a grand challenge," in *Proc. MICCAI Workshop on Medical Image Analysis for the Clinic*, pp. 207–214 (2010).
27. L. Xu et al., "Structure extraction from texture via relative total variation," *ACM Trans. Graphics* **31**(6), 139 (2012).
28. S. M. Pizer et al., "Adaptive histogram equalization and its variations," *Comput. Vision, Graphics Image Process.* **39**(3), 355–368 (1987).
29. L. Xu et al., "Image smoothing via L_0 gradient minimization," *ACM Trans. Graphics*, **30**(6), 174 (2011).
30. Y. Zhao et al., "Retinal vessel segmentation: an efficient graph cut approach with retinex and local phase," *PLoS One* **10**, e0122332 (2015).
31. I. Hacıhaliloglu et al., "Bone surface localization in ultrasound using image phase-based features," *Ultrasound Med. Biol.* **35**(9), 1475–1487 (2009).
32. M. Felsberg and G. Sommer, "The monogenic signal," *IEEE Trans. Signal Process.* **49**(12), 3136–3144 (2001).
33. P. Kovesi et al., "Symmetry and asymmetry from local phase," in *10th Australian Joint Conf. on Artificial Intelligence*, Vol. 190, Citeseer (1997).

34. I. Hacıhalilolu et al., "Automatic bone localization and fracture detection from volumetric ultrasound images using 3-D local phase features," *Ultrasound Med. Biol.* **38**(1), 128–144 (2012).
35. S. Andrews, G. Hamarneh, and A. Saad, "Fast random walker with priors using precomputation for interactive medical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention*, pp. 9–16, Springer, New York (2010).
36. A. Fedorov et al., "3D Slicer as an image computing platform for the quantitative imaging network," *Magn. Reson. Imaging* **30**(9), 1323–1341 (2012).
37. 3D Slicer, 2016, www.slicer.org (29 November 2016).

Mikhiel Lim earned his bachelor's degree in bioengineering from Binghamton University in 2014. He recently earned his master's

degree in biomedical engineering from Rutgers University in 2016, working under Dr. Ilker Hacıhalilolu. His interest is in medical imaging.

Ilker Hacıhalilolu is an assistant professor at the Department of Biomedical Engineering and an adjunct assistant professor at the Department of Radiology at Rutgers University. From 2010 till 2013, he was a postdoctoral fellow at the Centre for Hip Health and Mobility, Vancouver General Hospital, Vancouver, British Columbia, Canada. He received his PhD from the Department of Electrical and Computer Engineering at the University of British Columbia, Vancouver, British Columbia, Canada.