Published in final edited form as:

Med Image Anal. 2011 August ; 15(4): 438–448. doi:10.1016/j.media.2011.01.007.

Automatic graph-cut based segmentation of bones from knee magnetic resonance images for osteoarthritis research

Sufyan Y. Ababneh, Jeff W. Prescott, and Metin N. Gurcan

Department of Biomedical Informatics, The Ohio State University, 333 W. Tenth Avenue, Columbus, OH, USA

Abstract

In this paper, a new, fully automated, content-based system is proposed for knee bone segmentation from magnetic resonance images (MRI). The purpose of the bone segmentation is to support the discovery and characterization of imaging biomarkers for the incidence and progression of osteoarthritis, a debilitating joint disease, which affects a large portion of the aging population. The segmentation algorithm includes a novel content-based, two-pass disjoint block discovery mechanism, which is designed to support automation, segmentation initialization, and post processing. The block discovery is achieved by classifying the image content to bone and background blocks according to their similarity to the categories in the training data collected from typical bone structures. The classified blocks are then used to design an efficient graph-cut based segmentation algorithm. This algorithm requires constructing a graph using image pixel data followed by applying a maximum-flow algorithm which generates a minimum graph-cut that corresponds to an initial image segmentation. Content-based refinements and morphological operations are then applied to obtain the final segmentation. The proposed segmentation technique does not require any user interaction and can distinguish between bone and highly similar adjacent structures, such as fat tissues with high accuracy. The performance of the proposed system is evaluated by testing it on 376 MR images from the Osteoarthritis Initiative (OAI) database. This database included a selection of single images containing the femur and tibia from 200 subjects with varying levels of osteoarthritis severity. Additionally, a full three-dimensional segmentation of the bones from ten subjects with 14 slices each, and synthetic images with background having intensity and spatial characteristics similar to those of bone are used to assess the robustness and consistency of the developed algorithm. The results show an automatic bone detection rate of 0.99 and an average segmentation accuracy of 0.95 using the Dice similarity index.

Keywords

Segmentation; graph-cut algorithm; osteoarthritis; tibia; femur

1. Introduction

The rapid growth of diagnostic medical imaging studies has led to enormous strides in the effective diagnosis and treatment of myriad diseases, from appendicitis to brain cancer (Brant and Holms, 2006). The rise of imaging as a major factor in medical decision making has directly led to a drive towards quantification of image findings, in order to augment the mostly visual evaluation of trained medical professionals. To this end, the National Institutes of Health (NIH) and the Food and Drug Administration (FDA) have declared the pursuit of quantitative imaging biomarkers to be a major task of current and future research in medical

imaging (Smith et al, 2003). *Imaging biomarkers* are defined as “any anatomic, physiologic, biochemical, or molecular parameter detectable with one or more imaging methods used to help establish the presence and/or severity of disease.” Therefore, many recent research efforts are focused on discovering imaging biomarkers by analyzing the statistical relationship between image features and clinical phenotypes for a variety of diseases including chronic diseases such as osteoarthritis (Baker et al, 2004) (Prescott et al, 2009b).

OA is a degenerative disease of the joints (Buckwalter et al, 2000). It is characterized by degeneration of the articular cartilage with concomitant cartilage repair and remodeling, subchondral bone sclerosis (hardening), and osteophyte (“bone spur”) and bone cyst formation. Clinical signs and symptoms that may accompany these joint changes include joint pain, effusion, movement restriction, and gait abnormalities. These characteristics make OA an extremely debilitating disease. OA is also the most common form of arthritis, affecting approximately 21 million American adults (Felson et al, 2000). In addition, more than half of the 35 million Americans age 65 and older have radiographic evidence of OA in at least one joint. These facts, coupled with the large cohort of aging Americans known as the “baby boomers,” make OA a serious public health concern. To address this issue, a collection of public and private interest groups, led by the NIH, funded the Osteoarthritis Initiative (OAI). The purpose of the OAI is to discover biochemical, genetic, and imaging biomarkers for the development and progression of OA, such that therapies can be directed at the underlying causes of the disease, instead of the current modes of treatment focused mainly on reducing symptoms (Nevitt et al, 2006). From 2004–2006, the OAI enrolled 4796 volunteers to participate in longitudinal clinical studies over a four-year period. The massive amount of imaging data to be produced highlights the need for the development of computer-assisted image analysis methods to investigate imaging biomarkers efficiently and accurately.

The work presented in this paper is part of a larger effort to investigate the role of joint structures such as bone, cartilage, meniscus, and muscle on the course of knee OA progression through characterization of relevant anatomical structures (Swanson et al 2009) (Prescott et al, 2010). The overarching goals are (1) to develop an automated segmentation methodology to be used as an anchor for segmentation of other joint structures, such as cartilage and meniscus, and (2) to discover and characterize imaging biomarkers extracted from the bones. After the bone segmentation is performed, a set of features can be calculated by quantifying the bone’s content or its morphological characteristics. The relationships are then explored between the extracted features and the clinical phenotypes available for the test subjects. This paper presents an automated framework for the solution of aim (1), and offers a means to develop future work towards aim (2). Specifically, the current work is focused on segmentation of the femur and tibia bones in the knee joint.

A typical knee MR image obtained from the OAI database is shown in Figure 1. In these kinds of images, there is a presence of fat tissues whose intensity characteristics are very similar to those of bone structures. This ambiguity can lead to misclassifying fat as bone structures. Therefore, segmentation of bones is not a straightforward process. Previous methods for automated segmentation of bone in MR imaging have included the use of texture features (Bourgeat et al, 2007; Lorigo et al, 1999), active contours (Fripp et al 2007; Lorigo et al, 1999), shape models (Schmid and Magnenat-Thalmann, 2008), normalized cuts (Carballido-Gamio et al, 2004), and graph cut (GC) (Li et al 2005) (Liu et al, 2008). GC based segmentation algorithms have shown high accuracy, simultaneous ROI detection, and scalability to three-dimensions (3-D) in segmenting structures (Boykov, 2008). GC algorithms are combinatorial optimization techniques, which have been used extensively for applications in image processing, including segmentation (Li et al, 2005)(Boykov and Kolmogorov, 2004). For example, (Li et al, 2005) use multi-object GC as part of multi-stage

procedure to perform semi-automatic simultaneous segmentation of bone and cartilage surfaces in 3D images with application on human ankle bone with osteoarthritis. Using the same method, (Yin et al. 2008) segment knee bone and cartilage, starting from rough initial segmentations, which they generate by adapting single-object statistical shape models of the knee bones. They evaluated 17 MR datasets and report a successful segmentation accuracy of 0.94 in which 10% of the surface requires further manual editing of the border or the surface.

It is essential to be able to provide an automatic and robust image segmentation to support image analysis and processing of massive volumes of medical images datasets. Most of the segmentation algorithms are “semi-automated,” they require initialization and user input. The amount of information needed for initialization varies from one segmentation algorithm to another. For example, the region-growing based segmentation algorithms (Adams et al, 1994)(Gonzalez and Woods, 2002) require only a few seed points to perform segmentation. Other algorithms demand as many seed points as possible from both the ROI and the background (i.e., everything else). Examples of such algorithms are the GC based algorithms (Boykov, 2001) and the grow-cut algorithm (Vezhnevets and Konouchine, 2005). Segmentation can be automated by providing the segmentation algorithm with the necessary initialization information without the need for user initialization (Wang et al, 2007)(Fripp et al 2007). Since the goal is to have a fully automated segmentation system that can lead to high-throughput image analysis, it becomes necessary to eliminate the need for user-provided constraints and seed points to perform GC segmentation. However, many GC based algorithms require an extensive number of user-provided seed points and constraints to initialize segmentation (Boykov, 2006)(Boykov et al 2007).

The segmentation framework proposed herein addresses the initialization and seed selection by using training data collected for both ROI, (i.e. bones) and background regions (e.g. fat tissue, muscles) to guide a two-pass content-based block classification (discovery) process. The ROI in this case consists of the femur and tibia bones, and the background is everything else in the knee area (Figure 1). A multi-feature weight formula is defined and used to classify the scanned image blocks. The block discovery process is performed in two passes. In the first pass, training data is collected and used to classify the image blocks. This is followed by a second pass where the detected ROI blocks in the first pass are used as training data to augment the block classification achieved in the first pass. The proposed weight formula uses texture features and pixel intensity distribution. Training data blocks, which represent both ROI and background, are collected from a selected set of images to use as reference blocks. The results of applying the automatic block classification are a collection of ROI, background and undecided blocks. The detected object and background blocks are then used to initialize and support a subsequent segmentation process achieved by an efficient GC algorithm. However, the proposed framework is flexible to allow for the use of other algorithms, such as region growing. For this project, a GC algorithm is used due to its accuracy and scalability to 3-D in segmenting structures.

GC algorithms use global cost functions to find an optimal segmentation. A lack of discriminative power in the cost function may lead to ambiguous segmentation output where undesired structures are included as part of the final segmentation output. In that case, a refinement operation is needed to filter out undesired structures and refine the raw segmentation mask. Some algorithms allow for an editing stage in which an additional user input helps with conducting a partial second execution of the algorithm (Boykov, 2006). The approach proposed in this paper involves a content-based refinement procedure to improve the segmentation mask resulting from the GC algorithm, in order to distinguish between the bone regions and the fat tissue regions. Content-based criteria are used as part of the refinement procedure to filter out the unwanted objects. The criteria uses the block

classification results obtained during the automation stage in addition to image edge information.

A preliminary version of this work, previously described in (Ababneh and Gurcan, 2010a, 2010b), proposed the framework for automatic bone detection and GC based segmentation. This paper expands on the previous effort by integrating the automatic bone detection framework with the GC based segmentation to form a unified bone segmentation system, by presenting both a leak detection and correction mechanism and a segmentation refinement operation. In addition, it provides a unified framework for the evaluation of these kinds of systems for accuracy, robustness and consistency. Using this framework, we validate the robustness of the system by conducting extensive evaluation on 376 individual images, synthetic images as well as whole volumes (3-D) of MR image data. Furthermore, variation that may occur due to OA disease progression is taken into account by testing the algorithm on cases with all Kellgren-Lawrence (KL) grades (0 to 4), a radiographic measure of OA severity (Kellgren and Lawrence, 1957). This serves the end goal of deriving features from images that belong to patients with different levels of OA severity.

The rest of this paper is organized as follows: Section 2 describes the proposed automatic segmentation framework with five subsections to describe the main components of automatic bone detection, block discovery, GC segmentation and segmentation refinement. This is followed by Section 3, which describes an evaluation framework and gives the results of the performance analysis. Section 4 presents a discussion and conclusion.

2. Automatic segmentation framework

An overview of the automated segmentation framework is shown in Figure 2. The proposed segmentation framework can be described in two phases: Phase I: The pre-segmentation phase, and Phase II: the segmentation and post processing phase. The first phase starts with image pre-processing followed by training data collection to use as a reference for block-wise feature extraction. A block-wise ROI discovery operation is then applied to classify image blocks to ROI and background blocks. The classification outcome of the ROI discovery operation is then used by the second phase in two ways: (1) to initialize the subsequent segmentation algorithm and (2) to support the post processing steps. The second phase starts with initializing and applying a segmentation algorithm (e.g. GC, region grow). In the case of GC, a graph is first reconstructed based on seed and feature information obtained from the previous steps. A GC algorithm using maximum-flow minimum-cut is then applied followed by a content-based refinement procedure. Finally, a post-processing operation including morphological operations and leak detection and correction are applied to improve final segmentation.

2.1 Significant-block discovery

The discovery of ROI blocks (*significant blocks*) is performed by disjoint (non-overlapping) block-wise scanning of the image. The image is divided to $n \times n$ square blocks; the content of each block in the image is compared to those of *training blocks* representative of either ROI (bone) or background (not bone) blocks from a set of training images using a similarity metric. This metric is obtained by combining multiple similarity weight functions. The features used to generate the similarity metric are empirically selected after showing an effective discrimination power. The combined weight of a given image block i , $W(i)$ is used to classify that block as ROI, background, or undecided by using multiple thresholds :

$$W(i) = \alpha_1 \cdot W_1(i) + \alpha_2 \cdot W_2(i) + \dots + \alpha_n \cdot W_n(i), \quad (1)$$

where $\{W_1(i) \cdots W_n(i)\}$ are the n weight components of block i based on given quantifiable similarity metrics such as the distance between two intensity distributions. $\{\alpha_1 \dots \alpha_n\}$ are scaling factors used to control the importance of each individual weight component. Following the weight formulation presented in Equation (1), the weight function is formed by a combination of three feature-based sub-weights, each with some level of discrimination power. The resulting weight function $W(i)$ is defined as

$$W(i) = \alpha_1 \cdot H(i) + \alpha_2 \cdot C(i) - \alpha_3 \cdot e^{-\mu B(i)}, \quad (2)$$

where $\alpha_1, \alpha_2, \alpha_3$ and μ are scaling factors. $H(i)$, $C(i)$ and $B(i)$ are non-negative individual similarity metrics. The first term in Equation (2) is defined as

$$H(i) = \frac{h(i)}{N}$$

where $h(i)$ denotes the number of the ROI training blocks that are considered similar to block i . $h(i)$ is obtained by using the distance between two normalized intensity distributions.

$$h(i) = \sum_{k=1}^N \delta_k(p_i, q_k)$$

$$\delta_k(p_i, q_k) = \begin{cases} 1, & d_k(p_i, q_k) < T_d \\ 0, & d_k(p_i, q_k) \geq T_d \end{cases}$$

where T_d denotes a threshold value and d_k denotes the distance calculated using the Kullback–Leibler divergence between two normalized intensity distributions p and q over a range of gray levels L (Cover and Thomas, 1991),

$$d_k(p, q) = \sum_{j=1}^L p(j) \log \frac{p(j)}{q(j)}$$

The *second term* $C(i)$ in Equation (2) is derived based on Gray Level Co-occurrence matrix (GLCM) (Haralick et al, 1973). It is defined as

$$C(i) = \frac{c(i)}{N}$$

where N is the number of ROI training blocks and $c(i)$ is the number of ROI training blocks that are similar to block i using the distance between two feature values. The feature values are obtained using the *Information Measure of Correlation 1*, which is an entropy-based metric derived from the GLCM.

The discrimination power of several GLCM-derived features has been evaluated by monitoring the range of their values for both the bone and fat tissues. We observed that the *Information Measure of Correlation 1* feature provides a good discrimination between the knee fat tissues and bones,

$$\delta_k = \begin{cases} 1, & f_k < T_f \\ 0, & f_k \geq T_f \end{cases}$$

where T_f denotes a threshold value and f_k denotes *Information Measure of Correlation 1* as formally described in (Haralick et al, 1973). The *third term* in Equation (2) is evaluated using $B(i)$ which is computed in a way similar to $H(i)$, with the difference of using M background training blocks and using different threshold values. Therefore, a high value of $B(i)$ corresponds to a high dissimilarity with the ROI regions. The term is formed with an exponential to properly balance its effect with the other two terms in the weight function.

After the combined weight value is calculated for each block, threshold values T_1 and T_2 are used to classify each block as follows:

$$class(i) = \begin{cases} 0, & W(i) > T_1 \\ 1, & W(i) < T_2 \\ 2, & T_2 \leq W(i) \leq T_1 \end{cases}$$

where 0, 1 and 2 denote a ROI block, a background block or an undecided block, respectively. To conduct training, N ROI blocks and M distinct background blocks are obtained from a set of training images. The training images are chosen to represent typical images. The ROI training blocks are obtained by manually selecting a number of typical femur and tibia bone blocks (e.g. size of 24×24 pixels) from the training images. Once a block is selected, the content of the block is saved and the relevant features are extracted. The distinct background blocks are obtained by choosing blocks that are visually different from the ROI (bone) blocks. Since the background does not have a one distinctive feature, the set of background blocks are not necessarily representative of all possible background blocks. However, they are used to classify blocks that are highly likely to belong to the background.

2.2 Two-pass block discovery

Our segmentation method uses a two-pass block discovery. The first pass provides information about the image blocks and determines which are most likely to be ROI blocks. In the second pass, these blocks with high-likelihood of being ROIs can replace the training data used in the first pass. The second pass helps to discover new ROI blocks that did not show enough similarity with the original ROI training blocks, but they show higher similarity with the newly discovered ROI blocks due to their locality. The training data at the second pass is purely derived from the ROI blocks discovered at the first pass. An average value of the ROI blocks' normalized intensity distribution is calculated and used in the calculation of $H(i)$ in the second pass. The same procedure is followed for the calculation of $C(i)$.

2.3 GC-based segmentation

The ROI and background blocks discovered from the two-pass algorithm can now be employed as the initial input to a refined segmentation procedure. In this paper, for a refined segmentation algorithm, a GC algorithm (Boykov, 2006) is used as shown in Figure 4. Because the block discovery operation classifies some of the image blocks as ROI and background blocks; ROI and background seed points can be used to initialize graph construction. In addition, this classification allows for calculating ROI and background features, which are used to assign regional edge, cost values during graph construction. A

graph construction operation is performed using a global cost function, which uses both regional and boundary weights (explained later in this section). The constructed graph is then used by a graph minimum-cut (maximum-flow) algorithm (Boykov and Kolmogorov, 2004) to generate the minimum-cut mask image, which is a binary mask that represents all the segmented regions. The mask image is then subjected to a *content-based refinement process* that filters out background objects. Finally, morphological operations and leak detection are used as post processing steps to further enhance the segmentation outcome (Section 2.4).

The GC algorithm requires the construction of an s-k graph, where each pixel in the image is represented by a vertex in the graph (Boykov, 2006). In addition, two terminal vertices, source s and sink k are added to the graph to represent the ROI and background, respectively. One way to construct the graph is by creating an edge between each image-pixel vertex and its four immediate neighbors (neighbor edge). In addition, an edge is created between each pixel vertex and both the s and k terminal vertices (terminal edges). An illustration of graph construction is shown in Figure 5, with nine vertices representing a 3×3 image space, and both source (ROI) and sink (background) vertices. A non-negative weight value is assigned to each edge. Each weight value assigned to an n-edge reflects the degree of similarity between the two connected pixels (the similarity is defined later in this section); higher edge weight values represent pixels with higher similarity. On the other hand, the weight value assigned to each t-edge reflects each pixel's similarity with both the ROI and the background (the source and the sink, respectively) (Boykov, 2006). The value assigned to each pixel-source (s) edge represents the regional penalty or cost that is added to the global cost if that specific pixel is classified as a background during the GC segmentation process. The opposite is true for the k-edge. In the current application, the source, s , is set to represent the bone region and the sink, k , to represent the background region.

After the graph is constructed, the maximum flow (minimum cut) algorithms are applied on the graph with the goal of computing the GC that yields an optimal (minimum cost) segmentation. The global cost function is comprised of both regional and boundary constraints and can be defined as

$$C(A) = \gamma T(A) + N(A),$$

where $C(A)$ is the cost of a possible segmentation A , where A is a binary vector whose components specify assignments to the image pixels. $T(A)$ and $N(A)$ are regional and boundary cost functions respectively, and γ is a scaling factor to control the contribution of the regional cost function. Before defining the terms of the cost function, let $\mathcal{P}$ denote the set which includes all pixels p , and $\mathcal{N}$ denote the set which includes all neighboring pixels $\{p, q\}$.

$$\begin{aligned} T(A) &= \sum_{p \in \mathcal{P}} T_p(A_p) \\ N(A) &= \sum_{\{p, q\} \in \mathcal{N}} N_{p, q} \cdot \delta(A_p, A_q) \\ \delta(A_p, A_q) &= \begin{cases} 1, & \text{if } A_p \neq A_q \\ 0, & \text{otherwise} \end{cases} \end{aligned}$$

The specific regional cost function used in this paper is obtained by measuring the distance between each pixel value and the mean pixel intensity value of both the object (ROI) and

background regions. $T_p(A_p)$ for a given pixel p consists of both the object and background edge cost values $T_p(\text{"object"})$ and $T_p(\text{"background"})$, which are assigned to the s and k edges, respectively.

$$\begin{aligned} T_p(\text{"object"}) &= \text{abs}(I_p - \mu_{obj}), \\ T_p(\text{"background"}) &= \text{abs}(I_p - \mu_{bg}), \end{aligned}$$

where I_p denotes the gray-level intensity value of pixel p , μ_{obj} and μ_{bg} denote the mean pixel intensity values of the object and background pixels, respectively. The boundary cost function is obtained by measuring the similarity between each two neighboring pixels. The cost value is inversely proportional to the dissimilarity between the two pixels, which means that a higher penalty is paid in the GC algorithm when two very similar pixels are assigned to two different regions.

$$N_{p,q} \propto \exp\left(-\frac{(I_p - I_q)^2}{2\sigma^2}\right) \cdot \frac{1}{\text{distance}(p, q)}$$

where $\text{distance}(p, q)$ denotes the distance between the neighboring pixels (e.g. Euclidean distance) and σ denotes a sensitivity factor, which is a real-valued positive number that can be selected experimentally.

2.4 Content-based refinement and morphological operations

Content-based refinement is a post-processing step after the initial segmentation. In this work, the refinement is focused on the output of the GC algorithm. However, the content-based characteristic of the refinement stage is applicable to the output of other segmentation methods, such as region growing. The GC algorithm does not always result in a clear segmentation whose only segments are the ROI objects, partly due to a lack of an exclusive set of seed points that represent all the areas in the background and partly due to the limited discriminative power of the cost function that defines both the regional and boundary edges. Therefore, a content-based refinement operation is used as a solution to improve the segmentation output of the GC algorithm. An example of this is provided in Figure 6.

The refinement operation is performed on a segmentation mask, as shown in Figure 6 (d), which contains a mix of ROI and background segments. The block classification results, obtained during the automated block discovery, are used at this stage for refinement. These results are obtained using *Information Measure of Correlation 1*, derived from the GLCM as described earlier in Section 2.1. The refinement criteria also include using edge content to rule out segments with a high level of edge presence. The edge information is obtained using the Canny edge detector (Gonzalez and Woods, 2002). It is observed that the edge content is higher in fat tissues than the bone regions. Therefore, the Canny edge detector is applied and the number of edge pixels within each region is obtained. The Canny method detects edges by finding local maxima of the image gradient, which is calculated using the derivative of a Gaussian filter. The method detects strong and weak edges, and considers the weak edges only if they are connected to strong edges. Therefore, it is more likely than other methods to detect actual weak edges and not to confuse them with noise. Based on the reference feature measurements obtained from the training data as shown in Figure 6 (c), only the ROI segments are selected and the background segments are filtered out.

2.5 Leak detection and correction

The bone segmentation may occasionally include regions with similar intensity characteristics such as fat. We will call these segmentation problems as “leaks.” Such occasional leak problems can take place when ROI and background regions become joined together (usually in a narrow area) due to lack of strong edges between them. An example is shown in Figure 6(g) with a leak in the tibia bone. This situation requires an operation to separate the two regions.

A leak check is applied by performing an image opening using “disk” structuring elements of size 30. In case no leak is present, this operation results in shrinking or reducing the bone structure and leaving out small fragments that form the actual boundaries of the bone. A threshold of 400 pixels is used to identify the small fragments. If a leak is present, then the opening operation results in two reduced objects and filtering out fragments with varying sizes. Only one of the reduced objects represents bone and the other represents fat, which necessitates the leak correction operation. The correction operation starts by removing the reduced object that does not contain blocks classified as ROI (as discussed in Section 2.1). The fragments with large sizes are then examined using criteria based on edge content, intensity mean value distance and histogram distance. As a result, the fragments that are attached to the bone object are selected and the other fragments are removed. An example of leak detection is shown in Figure 6 with a leak in the tibia bone. The steps in Figure (g) to (k) illustrate a successful leak correction procedure.

3. Performance evaluation

The performance of the algorithm was evaluated by measuring the accuracies of segmentation as well as block detection rate. Several tests have been conducted to target a variety of images, including 3-D stacks which belong to a given patient, and arbitrary slice number for patients with varying degrees of OA progression.

3.1 Methods and dataset description

The images used in this work are obtained from the publicly available OAI dataset (Nevitt et al, 2006). A total of 376 MR images have been used to test the performance of the proposed segmentation scheme. They are sagittal T2 map 120 mm FOV MR images obtained from 3 Tesla Siemens machines. The image resolutions are 384×384 pixels (pixel spacing = . 3125/.3125 mm) with a slice thickness of 3 mm resulting in a total of 25–35 slices per series.

The accuracy of the proposed segmentation scheme has been objectively measured by comparing it with ground-truth regions obtained by manual segmentation performed by two trained readers. The *Dice Similarity Index* (DSI) has been used to quantify the region overlap between the automatic and the manual segmentations (Dice, 1945; Zijdenbos, 1994):

$$DSI=2 \cdot \frac{|A1 \cap A2|}{|A1|+|A2|}$$

where A_1 and A_2 refer to the manual and computer generated segmentations, respectively. In addition to segmentation accuracy defined by DSI, we define an additional measure of accuracy called “automation detection rate (ADR)” to represent the ratio of the successfully detected slices using the ROI block discovery process relative to the total tested slices. A successfully detected slice is the one with ROI blocks detected for both femur and tibia bones. Furthermore, we define yet another measure called “segmentation success rate (SSR)” to represent the ratio of the successfully segmented slices relative to the total tested slices. This is determined by considering any segmented slice with DSI higher than a

threshold of 0.6 as a successful segmentation. Choosing this threshold means that the statistics obtained are based on acceptable cases > 0.6 , the cases less than 0.6 are considered unworthy of including them in the statistics. SSR is needed to determine the number of cases in which the resulted segmentation is unacceptable.

3.2 Segmentation accuracy in 3-D

For training, three volumes of slices which represent three patients were selected. To obtain training data, for each slice four ROI blocks and two background blocks were selected and their features were extracted to be used during ROI block discovery. To test the performance of segmenting volumes of slices (3D), whole datasets from ten patients were used. For each patient, 14 consecutive slices have been selected (slice 8 to slice 21), which represent the slices with identifiable femur and tibia bone structures. Ten test patients were chosen to represent different KL grades (0–4). To test inter-reader variability, two readers provided ground truth segmentation masks for all ten patients. Tables I and II show the performance results of the segmentation by comparing against the ground truth. The average segmentation accuracy (DSI) achieved is 0.95 with a segmentation success rate (SSR) of 0.99. The automation detection rate (ADR) is 0.99. The misses encountered during the automated seed block extraction usually occur on images that have an extreme acquisition artifact or atypical bone texture. The DSI between the two readers (inter-reader reliability) is shown in Table III. The table shows a very low difference between the two DSI values measured using the ground truth segmentation of the two readers.

3.3 Segmentation robustness in relation to OA severity (KL grade)

Since the goal of this study is to develop an algorithm to be used for images of patients with varying levels of disease progression, a test on 206 patients has been conducted. The 206 patients have a range of radiographic OA severity levels in the range of 0 to 4 (i.e. KL 0–4). Therefore, a representative 41 patients have been selected to represent each grade, and a different slice number is selected for each patient starting from slice 8 and ending with slice 21. The goal of this experiment is to test the system's robustness against different KL grades because the bone structure may change as the KL grade increases. Two readers have provided the ground truth for segmentation. The training data has been selected to represent 50 patients among the 206 patients. The results are shown in Table IV with average segmentation accuracy achieved is with $DSI=0.94$ with a successful segmentation rate (SSR) of 0.93. The automated detection rate (ADR) is 0.99.

3.4 Robustness and consistency test

This test is performed on a set of images regardless of their KL grades to assess the system's robustness and consistency. It is performed on two representative boundary slices (slice 10 and 20) from 15 patients. The results are shown in Table V with $DSI = 0.94$, $SSR = 1.0$, $ADR = 0.99$.

An example of real knee MR images is shown in Figure 7, which provides a normal segmentation case. In this figure, the weight block shows the resulting weight value (y-axis) for each block in the image (x-axis). The highest values in the weight plot represent the blocks classified as bone and the lowest values are classified as background. In addition, synthetic images have been created for testing because the ground truth information can be accurately monitored. The simulated background has been generated by applying noise, Gaussian blurring, statistical filtering and pixel level shifting. The foreground is composed of actual bone and fat texture taken from real MR bone images. Figure 8 shows the set of images for regular shape structures for bone texture, background-like texture and fat tissue texture. The test shows the system's ability to precisely distinguish between the different

structures, and achieve a very accurate segmentation. Figure 9 shows another synthetic image test with structures representing real bone structures with accurate segmentation.

4. Discussion and conclusion

The goal of this work is to provide automated and accurate bone segmentation in knee MRI images of the OAI dataset as a means for imaging biomarkers discovery. In this paper, an efficient automated segmentation system is proposed that uses a content-based image block classification mechanism in conjunction with graph-cut methodology. The classified image blocks are utilized as inputs to a cost function, which is needed for graph construction and as an integral part of a post processing refinement operation. A two-phase approach has been proposed, the first phase of which identifies areas of bone with a high detection rate (0.99 accuracy). The output of this phase provides the seed points for the second phase, which accurately segments the bones and separates it from other high intensity structures such as fat and muscle with the help of a set of content-based features. The final segmentation is accurate with a mean DSI value of 0.95. The low standard deviation demonstrates the consistency of the results. The robustness of the system was tested with 376 images that come from cases with KL grades from 0 to 4. The cases with low DSI are due to a combination of factors such as the inconsistency in the bone texture of some images and a high fat presence that is strongly fused with the bone structures without clear-cut separation.

Novel aspects of this work include a two-pass approach, multi-feature weight formula, and the block discovery process. In all test cases, the second pass of the two-pass approach proved efficient in either augmenting the existing ROI block cluster (see Figure 3) or by discovering a new cluster that indicates the presence of a ROI region. In some cases, for example, the tibia bone is not detected in the first pass, but is detected in the second pass. The multi-feature weight formula enables combining the discrimination power of multiple features and can be easily expanded to include more features. The current selection of three features provided adequate level of discrimination between ROI and background with many blocks classified as undecided. The number of undecided blocks are likely be minimized by adding more relevant features. Some undecided blocks are resolved in the 2nd pass and the rest are discarded from further analysis. The automation, provided in this paper by the means of image block classification, may be suitable for many segmentation algorithms that rely on obtaining seed information from the image such as region growing. Even algorithms that require an approximate ROI contour instead of seed points for initialization (e.g. level set based algorithms) may benefit from our automation approach.

Four test scenarios have been used in this work with different emphasis in each scenario. In the first test scenario (segmentation in 3-D), ten stacks of slices (14 slices each) that belong to ten patients have been used. We found out from testing with this scenario that obtaining segmentation results in 3-D (see Table I and II) is achievable with consistent accuracy that allows for volume rendering and 3-D bone viewing. In addition, inter-reader variability is tested by obtaining the ground truth segmentation by two trained readers, the results in Table III show that the system maintains almost the same level of accuracy with both readers with a negligible change in DICE. The second test scenario is conducted using over 200 slices that represent 200 patients spanning all the KL grades. We learned from testing in this scenario that the proposed segmentation system works with all images irrespective of the OA disease progression measured by the KL grade, as shown by the consistent results in Table IV. In the third test scenario, two boundary slices are used (slice 10 and 20) in each of the 15 test volumes. This represents a sample test in which the boundary conditions are examined. Table V shows robust results that are further supported by the first test scenario. The fourth test scenario uses synthetic images by constructing images that represent ROI and non-ROI objects with known boundaries. With boundaries and textures known in

advance, the accuracy of the system can be measured and this helps provide a reference point to the expected accuracy when testing with actual images. The synthetic images we constructed showed the accuracy of ROI detection and segmentation as shown in Figures 8 and 9.

Previous work in quantifying bone changes on MRI has focused on pathologic changes in the subchondral bone, such as cysts (Pouder et al, 2008), bone marrow edema (Brem et al, 2008), and attrition (Neogi et al, 2009). There has also been some analysis of the effect of anatomical variations among patient populations and OA. These variations include morphology of the femoral condyles (Matsuda et al, 2004), surface area of the tibial plateau (Ding et al, 2007), and the morphology of the tibial tuberosity (Unlu et al, 2006). These previous studies have found promising associations between these pathological and anatomical measurements and OA onset, severity, and progression. While some of these studies have used MR imaging protocols, which differ from that of the dataset used in the automated segmentation described in this work, these relationships provide a reasonable starting point for future biomarker analysis.

The proposed system is designed to work on MR images obtained from the OAI (Section 3.1). The OAI has been acquiring images using different MR protocols (e.g. Coronal IW TSE) with the purpose of extracting properties of each knee joint structure specifically (Peterfy et al, 2008). Although the weighting equation used in this paper may not be directly applicable to the other studies, there are several potential ways of adapting this framework so that it will work with those images: (1) adjusting the scaling variables in the terms in the weighting equation, or (2) using the segmentation obtained from the current study as an atlas for segmentation of the bone structures in the other studies.

To further improve the existing work, a number of other approaches can be integrated with the current framework to yield better results. For example, some constraints, such as spatial information specific to a given type of images (e.g. ROI location), can be used to further guide or restrict the number of blocks included in the classification process. In addition, cost functions can be enhanced by investigating new features with enhanced discriminative power. A possible improvement can be achieved by exploiting the 3-D aspect of the data. This can be done during ROI block classification, where the result of one slice can be used as an input to enhance the result of the other slices in the dataset. In the present work, the segmentation is performed in 2D (not 3D) due to the anisotropic nature of the image data in the third dimension which is due to the greater slice spacing versus the in-plane resolution (3 mm vs. 0.3 mm, respectively). Applying 3-D segmentation could also help to avoid missing a bone in a single image slice. Another example of possible improvement is adding a shape priori or using shape models to guide the location and spatial arrangement of the blocks and to help improve the segmentation results at later steps (Fritscher, et, al, 2007; Leventon, 2000).

We believe that the proposed algorithm produces results suitable to be applied in the extraction and characterization of imaging biomarkers of OA. However, our group is still researching possible biomarkers, which correspond to radiographic severity and progression of OA (with the severity of OA measured by the KL grade). Biomarker measurements obtained using this automated method will still need to be validated against measurements obtained by manual methods to make sure that the accuracy of this algorithm can be taken as clinically meaningful. Based on our experience and the experience of clinicians working in our group, however, we believe that this automated method will produce segmentations that can properly characterize clinically meaningful imaging biomarkers. Additionally, current work is underway to use the segmentation results of this research to assist in the segmentation of other structures in the knee, such as identifying meniscus locations.

In conclusion, an efficient automated segmentation system is proposed in this paper by using a content-based image block classification mechanism in conjunction with GC segmentation. The classified image blocks are utilized as inputs for a cost function, which is needed for graph construction and as an integral part of a post processing refinement operation. The proposed system is designed to work on MR images obtained from the OAI. Validation results are presented to show the system's robustness and accuracy.

Acknowledgments

This project was supported in part by Award Number R01LM010119 from the National Library of Medicine. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Library of Medicine, or the National Institutes of Health. The authors would like to thank Drs. Thomas Best, David Flanigan, and Rebecca Jackson of the Ohio State University for useful discussions and guidance on the clinical aspects of this work.

References

- Ababneh, SY.; Gurcan, MN. An Automated Content-Based Segmentation Framework: Application to MR Images of Knee for Osteoarthritis Research; IEEE Inter. Conf. on Electro/Information Tech; 2010a.
- Ababneh, SY.; Gurcan, MN. An Efficient Graph-Cut Segmentation for Knee Bone Osteoarthritis Medical Images; IEEE Inter. Conf. on Electro/Information Tech; 2010b.
- Adams R, Bischof L. Seeded region growing. IEEE Trans. Pattern Anal. Machine Intell. 1994; vol. 16:641–647.
- Baker K, Xu L, Zhang Y, Nevitt M, Niu J, Aliabadi P, Yu W, Felson D. Quadriceps weakness and its relationship to tibiofemoral and patellafemoral knee osteoarthritis in Chinese. Arthritis Rheum. 2004; vol. 50:1815–1821. [PubMed: 15188358]
- Bourgeat P, Fripp J, Stanwell P, Ramadan S, Ourselin S. MR image segmentation of the knee bone using phase information. Med Image Anal. 2007; 11(4):325–335. [PubMed: 17482502]
- Boykov Y. Graph Cuts and Efficient N-D Image Segmentation. Int. Journal of Computer Vision. 2006; vol. 70(issue 2):109–131.
- Boykov Y, Kolmogorov V. An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in vision. IEEE Trans. On PAMI. 2004; vol. 26(no. 9):1124–1137.
- Boykov, Y.; Jolly, M. Interactive Graph Cuts for Optimal Boundary & Region Segmentation of Objects in N-D Images, Computer Vision, 2001. ICCV 2001; Proceedings. Eighth IEEE Int. Conf. in Comp. Vision; 2001.
- Brant, WE.; Holms, CA. Fundamentals of Diagnostic Radiology. third ed.. Philadelphia, PA: Lippincott Williams & Wilkins; 2006.
- Brem MH, Schlechtweg PM, Bhagwat J, Genovese M, Dillingham MF, Yoshioka H, Lang P. Longitudinal evaluation of the occurrence of MRI-detectable bone marrow edema in osteoarthritis of the knee. Acta Radiol. 2008; vol. 49(issue 9):1031–1037. [PubMed: 18720084]
- Buckwalter, JA.; Einhorn, TA.; Simon, SR. American Academy of Orthopaedic Surgeons. Orthopaedic Basic Science: Biology and Biomechanics of the Musculoskeletal System. Rosemont, IL: American Academy of Orthopaedic Surgeons; 2000.
- Carballido-Gamio J, Belongie S, Majumdar S. Normalized cuts in 3-D for spinal MRI segmentation. IEEE Trans on Med Img. 2004; vol. 23(issue 1):36–44.
- Cover, TM.; Thomas, JA. Elements of Information Theory. New York, NY: John Wiley and Sons, Inc; 1991.
- Dice L. Measure of the amount of Ecologic Association Between Species. Ecology. 1945; 26:297–302.
- Ding C, Cicuttini F, Jones G. Tibial subchondral bone size and knee cartilage defects: relevance to knee osteoarthritis. Osteoarthritis Cartilage. 2007; vol. 15(issue 5):479–486. [PubMed: 17291789]
- Felson D, Lawrence R, et al. Osteoarthritis: new insights. Part 1: the disease and its risk factors. Ann Intern Med. 2000; vol. 133:635–646. [PubMed: 11033593]

- Fritscher K, Grunerbl A, Schubert R. 3D image segmentation using combined shape-intensity prior models. *Inter. Journal of Computer Assisted Radiology and Surgery*. 2007; vol. 1(Number 6):341–350.
- Fripp J, Crozier S, Warfield SK, Ourselin S. Automatic segmentation of the bone and extraction of the bone-cartilage interface from magnetic resonance images of the knee. *Phys Med Biol*. 2007; vol. 52:1617–1631. [PubMed: 17327652]
- Gonzalez, RC.; Woods, RE. *Digital Image Processing using Matlab*. Upper Saddle River, NJ: Prentice Hall; 2002.
- Haralick RM, Shanmugam K, Dinstein I. Textural Features of Image Classification. *IEEE Transactions on Systems, Man and Cybernetics*. SMC. 1973; vol. 3(issue 6):610–621.
- Kellgren J, Lawrence J. Radiological assessment of osteoarthritis. *Ann Rheum Dis*. 1957; vol. 16:494–501. [PubMed: 13498604]
- Lindsey C, Narasimhan A, Adolfo J, Jin H, Steinbach S, Link T, Ries M, Majumdar S. Magnetic resonance evaluation of the interrelationship between articular cartilage and trabecular bone of the osteoarthritis knee. *Osteoarthritis and Cartilage*. 2004; vol. 12:86–96.
- Leventon, ME. PhD Dissertation. MIT; 2000. *Statistical Models in Medical Image Analysis*.
- Lorigo LM, Faugeras O, Grimson WEL, Keriven R, Kikinis R. Segmentation of bone in clinical knee MRI using texture-based geodesic active contours. *MICCAI*. 1998; vol. 1496:1195–1204.
- Li K, Millington, Wu S, Chen X, Sonka D. Simultaneous Segmentation of Multiple Closed Surfaces Using Optimal Graph Searching. *IPMI*. 2005; vol. 3565:406–417.
- Liu, L.; Raber, D.; Nopachai, D.; Commean, P.; Sinacore, P.; Prior, F.; Pless, R.; Ju, T. Interactive separation of segmented bones in ct volumes using graph cut. In: Metaxas, DN.; Axel, L.; Fichtinger, G.; Székely, G., editors. *MICCAI*, ser. LNCS. Vol. vol. 5241. 2008. p. 296-304.
- Matsuda S, Miura H, Nagamine R, Mawatari T, Tokunaga M, Nabeyama R, Iwamoto Y. Anatomical analysis of the femoral condyle in normal and osteoarthritic knees. *J Orthop Res*. 2004; vol. 22(issue 1):104–109. [PubMed: 14656667]
- Neogi T, Felson D, Niu J, Lynch J, Nevitt M, Guermazi A, Roemer F, Lewis CE, Wallace B, Zhang Y. Cartilage loss occurs in the same subregions as subchondral bone attrition: a within-knee subregion-matched approach from the Multicenter Osteoarthritis Study. *Arthritis Rheum*. 2009; vol. 61(issue 11):1539–1544. [PubMed: 19877101]
- Nevitt, MC.; Felson, DT.; Lester, G. The Osteoarthritis Initiative: Protocol for the Cohort Study, v. 1.1. 2006. <http://www.oai.ucsf.edu/datarelease/>
- Peterfy CG, Schneider E, Nevitt M. The osteoarthritis initiative: report on the design rationale for the magnetic resonance imaging protocol for the knee. *Osteoarthritis Cartilage*. 2008; vol. 16(issue 12):1433–1441. [PubMed: 18786841]
- Pouders C, De Maeseneer M, Van Roy P, Gielen J, Goossens A, Shahabpour M. Prevalence and MRI-anatomic correlation of bone cysts in osteoarthritic knees. *AJR Am J Roentgenol*. 2008; vol. 190(issue 1):17–21. [PubMed: 18094288]
- Prescott, J.; Swanson, MS.; Powell, K.; Gurcan, MN.; Haq, F.; Best, TM.; Jackson, RD. Template-Based Level Set Segmentation using Anatomical Information: Application to Quadriceps Muscles in MR Images from the Osteoarthritis Initiative; 24th International Symposium on Computer and Information Sciences; 2009a.
- Prescott J, Pennell M, Best TM, Swanson MS, Haq F, Jackson R, Gurcan MN. An automated method to segment the femur for osteoarthritis research. *IEEE Eng. Med. Biol. Soc (IEMBS)*. 2009b
- Prescott J, Best TM, Swanson MS, Haq F, Jackson RD, Gurcan MN. Anatomically-anchored template-based level set segmentation: Application to quadriceps muscles in MR images from the Osteoarthritis Initiative. *Journal of Digital Imaging*. 2010 e-published.
- Schmid J, Magnenat-Thalmann N. MRI bone segmentation using deformable models and shape priors. *Medical Image Computing and Computer-Assisted Intervention*. 2008:119–126. [PubMed: 18979739]
- Smith JJ, Sorensen AG, Thrall JH. Biomarkers in imaging: Realizing radiology's future. *Radiology*. 2003; vol. 227(issue 3):633–638. [PubMed: 12663828]

- Swanson MS, Prescott J, Best TM, Powell K, Jackson RD, Haq F, Gurcan MN. Semi-automated segmentation to assess the lateral meniscus in normal and osteoarthritis knees. *Osteoarthritis and Cartilage*. 2010; vol. 18(issue 3):344–353. [PubMed: 19857510]
- Tao, Y.; Xuan, J.; Freedman, MT.; Chepko, G.; Shields, PG.; Wang, Y. Imaging biomarker analysis of rat mammary fat pads and glandular tissues in MRI images; *Inter. Conf. on pattern recognition*; 2008.
- Unlu Z, Tarhan S, Goktan C, Tuzun C. The correlation between magnetic resonance detected cartilage defects and spiking of tibial tubercles in osteoarthritis of the knee joint. *Acta Med Okayama*. 2006; vol. 60(issue 4):207–214. [PubMed: 16943857]
- Vezhnevets V, Konouchine V. *GrowCut–Interactive Multi-Label N-D Image Segmentation By Cellular Automata*. Graphicon. 2005
- Wang, C.; Guo, X.; Yao, L.; Li, K.; Jin, Z. A Practical Method for Muscles Extraction and Automatic Segmentation of Leg Magnetic Resonance Images; *Inter. Conf. on Complex Medical Eng*; 2007.
- Yin, Y.; Zhang, X.; Sonka, M. Optimal multi-object multi-surface graph search segmentation: Full-joint cartilage delineation in 3d; *Proc. of the 12th Annual Conf. on Medical Image Understanding and Analysis*; 2008.
- Zijdenbos A, Dawant B, Margolin R, Palmer A. Morphometric analysis of white matter lesions in MR images: Method and validation. *IEEE Trans. Med. Imag*. 1994; vol. 13:716–724.

Figure 1.
A sample MRI image for the knee with annotated bone and fat tissues. the image shown is a
slice of a typical (3D) MR dataset

Figure 2. Illustration of system components (the dotted line represents the interface between the two phases.)

Figure 3.
Illustration of the block discovery outcome (from left to right: MRI image, first pass, second pass)

Figure 4.
Block diagram for the GC based segmentation algorithm

Figure 5.

An illustration of GC edge assignments, where the vertices 1, 3 and 4 represent pixels with high similarity to ROI, vertices 8 and 9 represent pixels with high similarity to the background, and vertices 2, 5, 6 and 7 represent pixels that have some level of similarity with both the ROI and the background. The thickness of the edge lines reflect the cost values.

Figure 6.

An example to illustrate both ROI detection failure/recovery and leak detection/correction. (a) MRI test image (b) ROI block discovery with only the femur bone detected; the tibia is not detected. (c) both bones are detected after lowering the ROI detection threshold (d) GC output mask (e) after morphological operations (f) the resulting two potential bone structures with a leak obvious in the tibia bone (g) tibia bone with its leak connecting fat and other structures to the tibia (h) first step to check for a leak by applying a morphological opening (i) the remaining content which resulted from subtracting (h) from (g). (j) After examining the remains in (i), the leak detection algorithm detects a leak and determines that only the pixels in (j) are relevant to the tibia (k) after adding the relevant pixels in j to h which results

in a leak-free tibia (i) the mask for the femur (m) after applying a morphological opening to check for leak presence (n) the remaining pixels after subtracting (m) from (i), based on this it is determined that there is no leak and the pixels are put back in (o). (o) the resulting femur and tibia masks (p) delineation of the GC segmentation in white and the manual segmentation in yellow with accuracy achieved of DICE= 0.95 and 0.96 for the femur and tibia bones, respectively.

Figure 7.

GC image segmentation, top row: MR image, block weights plot, x-axis: block number, y-axis: weight, block discovery first pass, block discovery second pass. Bottom row: raw mask generated by the GC algorithm, after morphological operations, after refinement and delineated segmentation with DSI of 0.98 and 0.96 for the femur and tibia bones, respectively

Figure 8.

Synthetic image segmentation with both bone and fat tissues (top row: bone and fat regions with artificial background, block weight image, block discovery, plot of weight blocks, x-axis: block number, y-axis: weight. bottom row: histogram of the ROI and background, x-axis: pixel intensity, y-axis: number of pixels, raw mask generated by the GC algorithm, after refinement and morphological operations, bone delineated image).

Figure 9.

Synthetic image segmentation with femur and tibia bones (top row: femur and tibia bone image with artificial background, block weight image, block discovery, plot of weight blocks (x-axis: block number, y-axis: weight). Bottom row: histogram of the ROI and background (x-axis: pixel intensity, y-axis: number of pixels),, raw mask generated by the GC algorithm, after refinement and morphological operations, and the segmented image where the bone is delineated.

TABLE I
Results on individual 3D datasets (140 slices) 10 patients with 14 slices each. Compared against segmentation by first reader.

Patient Num	KL	Success rate (SSR)	Femur Avg DSI	Tibia Avg DSI	both ^a Avg DSI	Femur DSI Std.	Tibia DSI Std.	both ^a DSI Std.
1	0	0.93	0.896	0.958	0.927	0.101	0.022	0.059
2	0	1.0	0.965	0.969	0.967	0.028	0.010	0.013
3	0	1.0	0.966	0.968	0.967	0.019	0.007	0.011
4	1	1.0	0.962	0.964	0.963	0.020	0.027	0.014
5	2	1.0	0.955	0.905	0.930	0.034	0.080	0.041
6	2	1.0	0.944	0.969	0.956	0.042	0.009	0.020
7	3	1.0	0.907	0.962	0.934	0.199	0.013	0.099
8	3	1.0	0.955	0.953	0.954	0.033	0.027	0.022
9	4	1.0	0.965	0.951	0.958	0.025	0.059	0.031
10	4	1.0	0.961	0.969	0.965	0.030	0.005	0.014

both^a means that the DSI values, for both femur and tibia, are used in a given image (by obtaining their average).

TABLE II

Results on individual 3D datasets (140 slices) 10 patients with 14 slices each. Compared against segmentation by second reader

Patient Num	KL	Success rate (SSR)	Femur Avg DSI	Tibia Avg DSI	both ^a Avg DSI	Femur DSI Std.	Tibia DSI Std.	both ^a DSI Std.
1	0	0.93	0.891	0.961	0.926	0.097	0.024	0.056
2	0	1.0	0.967	0.977	0.972	0.026	0.005	0.013
3	0	1.0	0.970	0.969	0.970	0.013	0.007	0.008
4	1	1.0	0.964	0.957	0.961	0.024	0.027	0.018
5	2	1.0	0.962	0.914	0.938	0.035	0.078	0.041
6	2	1.0	0.945	0.970	0.958	0.041	0.013	0.017
7	3	1.0	0.968	0.971	0.970	0.022	0.008	0.013
8	3	1.0	0.954	0.963	0.959	0.040	0.024	0.023
9	4	1.0	0.966	0.959	0.962	0.025	0.052	0.025
10	4	1.0	0.970	0.979	0.974	0.023	0.006	0.011

TABLE III

Inter-reader variability across 10 3D datasets (10 patients) with 14 slices each

Patient	Femur Delta DSI	Tibia Delta DSI	both ^a Delta DSI
1	0.005	0.003	0.001
2	0.002	0.008	0.005
3	0.004	0.001	0.003
4	0.002	0.007	0.002
5	0.007	0.009	0.008
6	0.001	0.001	0.002
7	0.061	0.009	0.036
8	0.001	0.01	0.005
9	0.001	0.008	0.004
10	0.009	0.01	0.009

TABLE IV

Results across disease severity grade (KL 0–4) with 41 patients per grade, coverage of 190 patients with one slice for each patient (the slices are sequenced from 8 to 21).

Bone	DSI Avg.	DSI Std.	DSI Median
femur	0.929	0.070	0.962
tibia	0.941	0.056	0.962
both	0.936	0.046	0.954

TABLE V

30 slices total, 15 patients with slices 10 and 20 for each patient

Bone	DSI Avg.	DSI Std.	DSI Median
femur	0.936	0.049	0.962
tibia	0.946	0.026	0.957
both	0.941	0.030	0.947