Automatic MRI Bone Segmentation

Toki Migimatsu
takatoki@stanford.edu

***Abstract*—Automatically segmenting bone tissue in MRI scans requires robustness against poor signal-to-noise ratios, highly inconsistent lighting conditions, and variability within bone tissues. Because of these difficulties, very little literature exists on automatic MRI bone segmentation. Current state-of-the-art methods are either semi-automatic or rely on databases of prior manual segmentations. Here, we propose a method that segments the radius and ulna bones using multiple passes of the Maximally Stable Extremal Regions (MSER) [1] algorithm on 2D slices of the volumetric MRI scans. On two MRI scans, our method achieved a Dice similarity coefficient (DSC) of 0.98, beating present state-of-the-art techniques.**

I. Introduction and Motivation

Magnetic Resonance Imaging (MRI) provides a safe and non-invasive way to study internal tissues and create detailed musculoskeletal models of the body. Aside from clinical applications, these models can be used in areas ranging from character animation to assistive robotics, where accurate models of human motion are important. However, obtaining these musculoskeletal models requires manually segmenting bones and muscles, a prohibitively time-consuming process; segmenting one knee alone takes an expert hours [2]. Automatic segmentation could allow the mass generation of musculoskeletal models.

Fully automatic segmentation is difficult for three primary reasons. First, the low signal-to-noise ratio of MRI scans makes identifying different tissues difficult, even for human labelers [2]. Second, highly inconsistent lighting in the MRI scanning workspace causes white trabecular bone tissue in one region to be darker than the background in other regions, making global image processing difficult. Finally, different bone tissues tend to vary more in appearance with each other than with the surrounding muscle tissue, meaning segmentation methods must be robust to inconsistencies.

In this paper, we outline a pipeline for automatically segmenting the radius and ulna. However, this method is immediately generalizable to any bone with a relatively thick cortical layer, such as the humerus.

II. Related Work

Current methods for automatic segmentation rely on statistical techniques to generate new segmentations by warping prior datasets. However, these methods require the existence of prior manual segmentations, which are only available from clinics for frequently segmented regions such as hip and knee joints. Seim et al., for example, uses 60 MRI knee segmentations to create a statistical shape model of the knee, which is then transformed based on voxel intensities [2]. Xia et al. performs Principal Component Analysis (PCA) on the shape variations of 28 hip bone segmentations and transforms the features using gradient information of the image [3]. Schmid et al. similarly uses 29 hip bone segmentations to perform PCA, but warps the features using Markov Random Field forces [4]. By their statistical nature, these methods cannot be easily used to segment infrequently segmented bones. Thus, a method for automatic segmentation that does not require prior segmentations is desired.

Other methods start from very rough manual segmentations and use boundary-seeking algorithms to refine the segmentations [5], [6]. Although these methods lift a significant amount of labor from precise manual segmentation, they still do not make the mass generation of musculoskeletal models possible.

III. Methodology

A. Background Removal

Because the backgrounds of MRI scans are so noisy and contain significant lighting artifacts, it is important to filter out the background before proceeding with other image processing techniques. Otherwise, the background compromises the effectiveness of these methods.

The skin boundary is relatively distinct, so we can outline the foreground with a a simple edge detector. However, due to inconsistent lighting, the magnitudes of the image gradients at this boundary vary significantly. Thus, we choose to use a Laplacian of Gaussian edge detector because it responds equally to strong and weak edges. Furthermore, it is robust against blurry images, which is a necessary feature as most MRI scans have low resolutions, and miniscule movement of the subject during the scanning process adds additional blur. The Laplacian of Gaussian edge detector also conveniently produces complete loops that we can fill in with simple hole filling methods to create a foreground mask.

The foreground mask can afford to be approximate, since bones do not typically lie close to the skin's surface. Thus, for speed, we simply slice the 3D volume along one axis and process each slice as 2D images. On each slice, we first perform a median filter to reduce the grainy noise. Then, we apply the Laplacian of Gaussian edge detector, which typically produces the result in Figure 1.

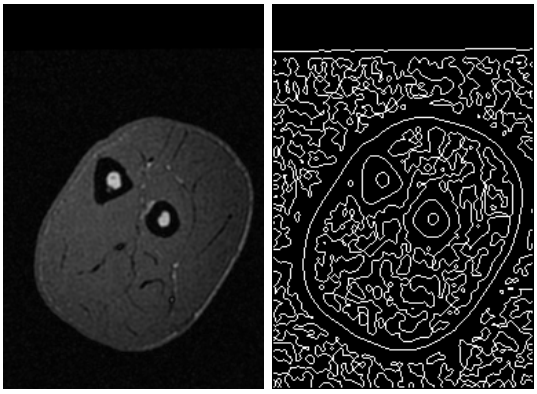

Fig. 1. Using Laplacian of Gaussian to detect the skin boundary. This image is a cross section of the forearm taken perpendicular to the arm's axis.

Clearing all edge loops connected to the border typically removes the edges in the background. Then, performing a morphological close operation with a disk followed by hole filling consolidates the edges within the foreground into a solid mask. To remove extraneous foreground slices that stick out from neighboring slices, we can perform morphological opening with a square along one of the other two axes. Finally, to fill in gaps between slices and to simultaneously smooth out the final mask, we perform a morphological open with a circle. This produces the mask in Figure 2.

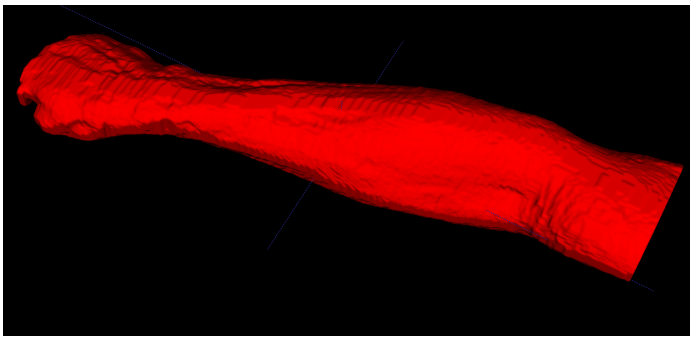

Fig. 2. Foreground mask extracted from Laplacian of Gaussian edge detection.

B. Adaptive Histogram Equalization

By masking out the noisy background of the MRI image, we can then use adaptive histogram equalization to compensate for unequal lighting in the image. For a 3D image, we simply extend standard adaptive histogram equalization with 3D window boxes instead of squares. Figure 3 shows the result of equalization. Running histogram equalization with small step sizes is computationally expensive, so we aim to use as large window/step sizes as possible.

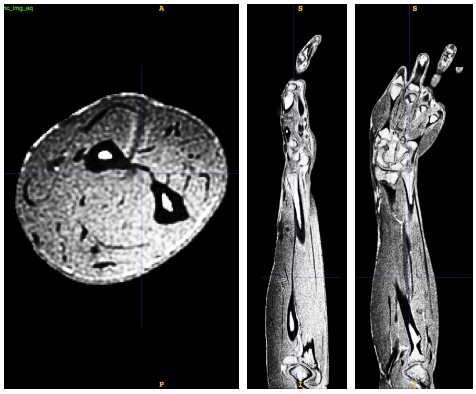

Fig. 3. 3D adaptive histogram equalization to mitigate unequal lighting.

C. Trabecular Tissue Detection with MSER

The trabecular tissue of bones shows up as white in the MRI scan and is surrounded by a black cortical tissue layer, which also belongs to the bone. Because the trabecular tissue is isolated from the surrounding muscle tissue, detecting this tissue first provides a reliable method to identify separate bones.

We use bright-on-dark Maximally Stable Extremal Regions (MSER) [1] to mark the trabecular regions on slices of the MRI volume along a particular axis. In situations where the trabecular tissue's surface is tangential to the 2D image slice, MSER has trouble identifying trabecular tissue. To resolve this issue, we perform a second pass of MSER on slices along a different axis. The result is shown in Figure 4.

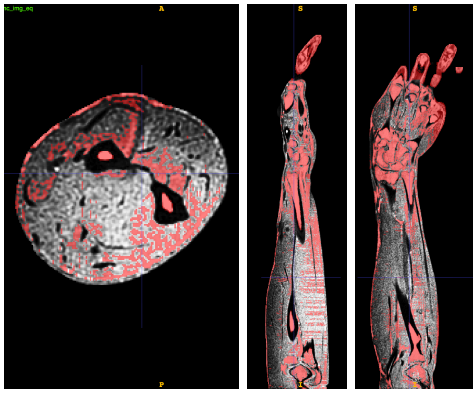

Fig. 4. Bright-on-dark MSER reveals trabecular tissue.

Finally, we identify the 3D connected components of the volume with a 6-neighborhood. In the case of the radius and ulna, the longest connected components with the greatest eccentricity are the relevant segmentations (Figure 5).

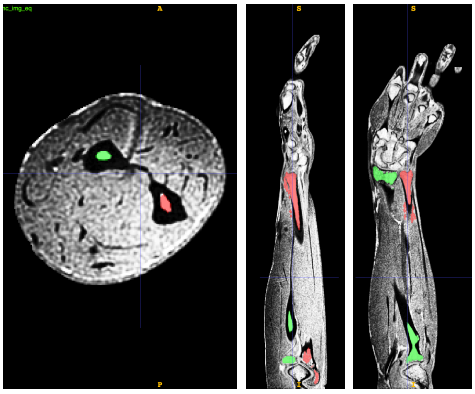

Fig. 5. The longest and most eccentric connected components correspond to the trabecular tissues of the radius and ulna.

D. Cortical Tissue Detection with MSER

To detect the dark cortical regions of the bone, we apply dark-on-bright MSER on slices of the MRI volume along only one axis. For each slice, we identify the cortical regions by the dark segmentations that are connected to the trabecular segmentations with a 4-neighborhood.

After filling in the holes and applying a 3D morphological close with a sphere for some post-processing smoothing, we obtain the final segmentation of the radius and ulna (Figure 6).

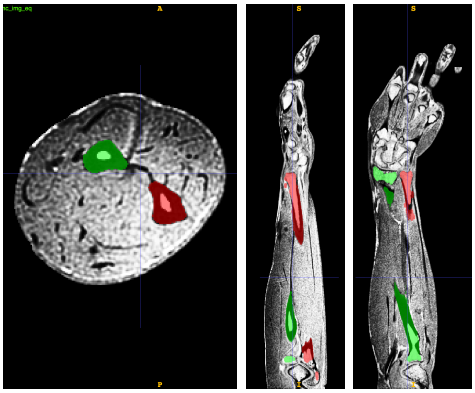

Fig. 6. Dark-on-bright MSER adds cortical tissues to the segmentation.

IV. Evaluation

To score the accuracy of the automatic segmentations, we calculated the ratio of foreground voxels that matched the corresponding voxels in the manually segmented reference image. This is equivalent to the Dice similarity coefficient (DSC) [7]. In Table I, the false positive score represents the percentage of voxels that were incorrectly labeled as bone tissue, and false negative score represents the percentage of voxels that were incorrectly labeled as non-bone tissue. Because of the difficulty of manually segmenting bones, we only evaluated two automatic segmentations.

Both scans scored DSCs of around 0.98, achieving a state-of-the-art record. Most of the error comes from false negatives; in other words, the automatic segmentation is more conservative with bone segmentation. For comparison, Seim et al. and Xia et al. achieved DSCs of 0.80 and 0.93, respectively, for automatic knee and hip bone segmentation [2] [3].

TABLE I
Performance of proposed segmentation method.

Scan ID	False Positives	False Negatives	DSC
1	0.02%	2.20%	0.9878
2	0.60%	1.10%	0.9830

The two scans were taken with the same MRI scanner and resolution settings but with different subjects. Without changing any parameters, the automatic segmentation performed very well for both scans. This suggests that the parameters can be tuned once for a particular MRI scanner, and then left alone to automatically segment many scans across different subjects.

As Figure 7 shows, the errors mostly occur at the ends of the bones, where the trabecular tissue is darker and the cortical tissue is thinner. Sometimes, the cortical tissue is thinner than a pixel. Thus, the edges are less pronounced and do not show up as easily with MSER.

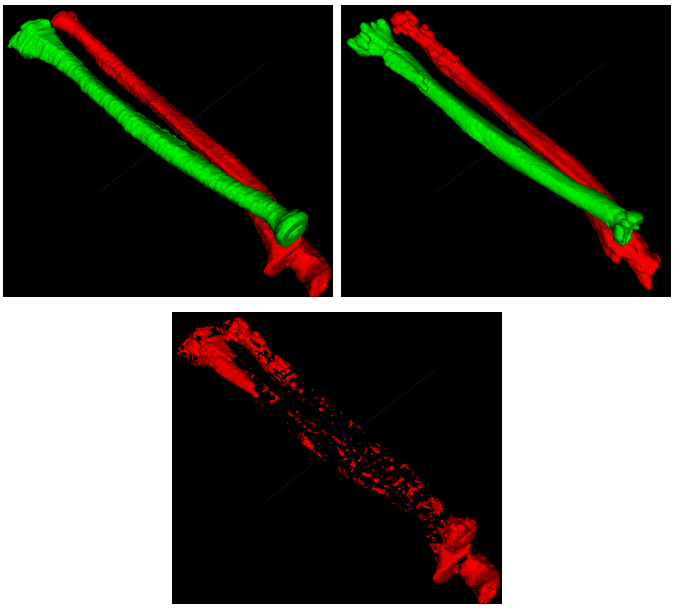

Fig. 7. Top left: manual reference segmentation. Top right: automatic segmentation. Bottom: mislabeled voxels.

V. Comparison to Alternative Approaches

A. 3D Otsu Thresholding

A simpler implementation for automatic segmentation might use adaptive Otsu thresholding. In theory, this would work well to separate the bright trabecular tissue or the dark cortical tissue of the bone from the other tissues. We tested 3D Otsu thresholding with various window box sizes and step sizes, but failed to produce segmentations that could be further processed to isolate the radius and ulna.

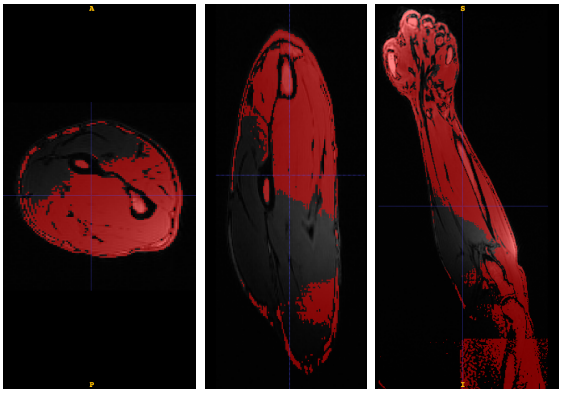

Fig. 8. 3D adaptive Otsu thresholding performs poorly on all window sizes.

The issue is that in areas where the dark cortical tissue is lighter, it gets labeled as a bright area, thus connecting the bright trabecular and muscle tissues. Like MSER, this is particularly an issue at the ends of the bones, where the trabecular and cortical tissues have similar brightness levels. However, unlike MSER, because the areas of connection between bone and muscle tissue are so large at the ends of the bones, they cannot be easily split with operations such as morphological opening. With such poor performance on the radius and ulna, adaptive Otsu thresholding fails to compete with MSER.

B. 3D Sobel Surface Detection

Another implementation might use an edge detection algorithm such as Sobel, but extended to detect oriented surfaces in 3 dimensions.

The standard 2D Sobel operator can be defined as follows:

$$h_a(x) = \begin{bmatrix} -1 & 0 & 1 \end{bmatrix} \qquad h_b(x) = \begin{bmatrix} 1 & 2 & 1 \end{bmatrix}$$

$$h_a(y) = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} \qquad h_b(y) = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

$$h_{sobel_x}(x,y) = h_a(x) * h_b(y) = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

$$h_{sobel_y}(x,y) = h_a(y) * h_b(x) = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

With this definition, extending to three dimensions is straightforward:

$$h_{sobel_x}(x,y,z) = h_a(x) * h_b(y) * h_b(z)$$
$$h_{sobel_y}(x,y,z) = h_a(y) * h_b(z) * h_b(x)$$
$$h_{sobel_z}(x,y,z) = h_a(z) * h_b(x) * h_b(y)$$

As an example, $h_{sobel_z}(x,y,z)$ is:

$$h_{sobel_z}(x,y,-1) = \begin{bmatrix} -1 & -2 & -1 \\ -2 & -4 & -2 \\ -1 & -2 & -1 \end{bmatrix}$$

$$h_{sobel_z}(x,y,0) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$h_{sobel_z}(x,y,1) = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix}$$

Applying the three filters to the image, taking their sum of squares, and then thresholding the voxels with the largest gradients results in a segmentation as shown in Figure 9.

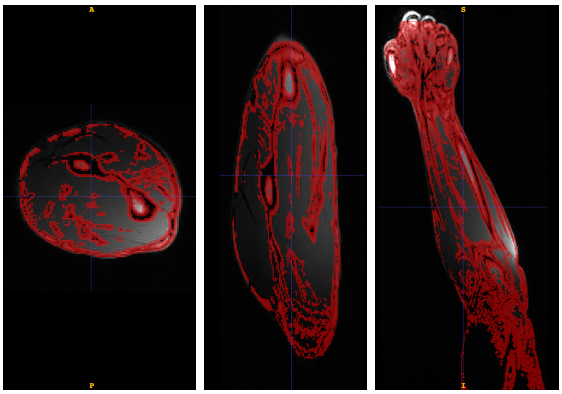

Fig. 9. Sobel filtering with a magnitude threshold set to highlight the top 10% brightest pixels.

The Sobel segmentation is cleaner than adaptive Otsu segmentation, but like the Otsu method, Sobel connects the bone tissue with muscle tissue at the ends of the bones where the cortical layer is thin. Thus, the bone segmentations cannot be extracted easily from the muscle with Sobel, and MSER remains the most robust method.

VI. Discussion & Future Work

With an average DSC of 0.98, the method presented in this paper beats state-of-the-art records with a fully automated approach. The high accuracy of this method means that the amount of manual post-processing required to bring the automatic segmentation up to the golden standard is minimal. Furthermore, the robustness of the parameters to different subjects means that many MRI scans can be automatically segmented without constant supervision.

Like all the segmentation methods explored in this paper, 2D MSER's weakest point is the segmentation of the ends of the ulna and radius. As a result, it likely cannot achieve 0.98 DSC on smaller bones such as hand bones with thin cortical tissues. Further refinement is required to resolve this issue.

One possible future direction may be to apply a fully 3-dimensional MSER method that uses 3D level sets to find closed 3D volumes. 3D MSER would likely perform better

than 2D MSER in cases where the thin cortical tissue tissue is tangential to the 2D slices, since it could still detect the surface from a different angle. However, it would also require more consistent lighting throughout the MRI volume; performing 2D MSER has the benefit of ignoring brightness changes between slices. Thus, 3D MSER would place greater importance on adaptive histogram equalization.

Another possibility is using the method proposed in this paper as input for semi-automatic boundary-seeking segmentation methods such as those presented by Rusu [5] and Jolivet et al. [6]. The combination of these methods could result in a fully automatic pipeline for robust bone segmentation.

REFERENCES

[1] J. Matas, O. Chum, M. Urban, and T. Pajdla, "Robust wide-baseline stereo from maximally stable extremal regions," *Image and vision computing*, vol. 22, no. 10, pp. 761–767, 2004.

[2] H. Seim, D. Kainmueller, H. Lamecker, M. Bindernagel, J. Malinowski, and S. Zachow, "Model-based auto-segmentation of knee bones and cartilage in mri data," in *Proc. MICCAI Workshop Medical Image Analysis for the Clinic*, B. v. Ginneken, Ed., 2010, pp. 215 –223.

[3] Y. Xia, S. Chandra, O. Salvado, J. Fripp, R. Schwarz, L. Lauer, C. Engstrom, and S. Crozier, "Automated mr hip bone segmentation," in *Digital Image Computing Techniques and Applications (DICTA), 2011 International Conference on*, Dec. 2011, pp. 25–30.

[4] J. Schmid and N. Magnenat-Thalmann, "Mri bone segmentation using deformable models and shape priors," English, in *Medical Image Computing and Computer-Assisted Intervention MICCAI 2008*, ser. Lecture Notes in Computer Science, D. Metaxas, L. Axel, G. Fichtinger, and G. Szkely, Eds., vol. 5241, Springer Berlin Heidelberg, 2008, pp. 119–126, ISBN: 978-3-540-85987-1.

[5] A. Rusu, *Segmentation of bone structures in magnetic resonance images (mri) for human hand skeletal kinematics modelling*, 2011.

[6] E. Jolivet, E. Dion, P. Rouch, G. Dubois, R. Charrier, C. Payan, and W. Skalli, "Skeletal muscle segmentation from mri dataset using a model-based approach," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 2, no. 3, pp. 138–145, 2014.

[7] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.