AUTOMATED BONE AND JOINT-REGION SEGMENTATION IN VOLUMETRIC ULTRASOUND

Kedar A. Patwardhan[⋆], *Kunlin Cao*[⋆†], *David Mills*[⋆], *Ralf Thiele*[‡]

[⋆] GE Global Research Center, Niskayuna, NY 12309 USA
[⋆†] University of Iowa, Iowa City, IA 52242, USA
[‡] University of Rochester School of Medicine and Dentistry, Rochester, NY 14642 USA
{patwardh,millsda}@ge.com

ABSTRACT

Ultrasound (US) is being increasingly used in the assessment of joints for disorders such as rheumatoid arthritis (RA). Pathologies related to RA in joints, manifest themselves commonly as changes in the bone (e.g. erosions) and the region enclosed by the joint-capsule (e.g. synovitis). Automated tools for detecting and segmenting such structures would help significantly towards objective and quantitative assessment of RA in joints. In this paper we present a novel method for automatic extraction of the 3D bone surface and segmentation of the joint-capsule region from volumetric (3D) scans of commonly affected joints such as the metacarpophalangeal (MCP) and metatarsophalangeal (MTP) joints. We improve and extend the cost function proposed in [1] to 3D such that it is more robust to loss of acoustic intensity due to the surface normals facing away from the incident acoustic beam. The extracted bone coupled with a simple anatomical model of the MCP joint provides a coarse localization of the joint-capsule region. A probabilistic speckle model is then used to iteratively refine the capsule segmentation. We illustrate the performance of proposed methods through quantitative comparisons with expert annotations as well as qualitative results on over 30 scans obtained from 11 subjects.

Index Terms— musculoskeletal, ultrasound, 3D, segmentation, bone surface, joint capsule

1. MOTIVATION AND PRIOR ART

Rheumatoid arthritis (RA) is a common chronic disease all over the world. Conventional 2D US is becoming a popular *point-of-care* imaging modality for the assessment of RA [2]. Nevertheless, 2D imaging does not capture the underlying 3D anatomy accurately, does not allow accurate quantification and requires training in order to make an objective assessment. Recent research ([3, 4]) has shown that 3D US is more sensitive to pathologies like bone-erosion than 2D US.

KP, DM & KC would like to thank GE Healthcare Ultrasound Point-of-Care for funding this research. KC was a summer intern at GE Global Research when she contributed to this work.

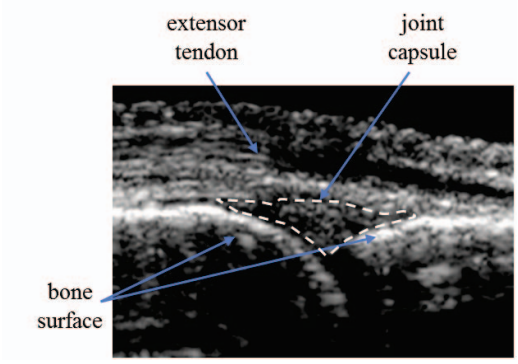


Fig. 1. Sonographically observed anatomy of a Metacarpophalangeal (MCP) joint (derived from [9]).

We believe that automated analysis and quantification of such 3D images could improve the treatment work-flow as well as reduce subjectivity in the assessment of the disease condition (e.g. erosions, hypertrophic synovium, etc). Numerous works [1, 5, 6, 7, 8] and references therein have addressed automated bone extraction in 2D as well as 3D images for different clinical applications. In [5], authors have proposed the use of local phase asymmetry in detecting image edges caused by bone. Local phase in general is well suited for US edge detection because of its robustness to variation in contrast which is commonly observed as one moves away from the transducer surface. In our application though, we have noticed that local phase based edge detection detects spurious edges other than bone equally well and since this is a multi-resolution, multi-orientation filtering algorithm, it generally takes a longer time for its edge computations in 3D. Authors in [8] propose a method for fast detection of bone contours in freehand 2D sweeps, which can be used to reconstruct a 3D bone surface. In practice, the quality of the extracted bone surface can vary from scan to scan because of the inherent 'freehand' nature of scanning and this may also lead to non-uniform bone-surface resolution depending on the speed of the freehand scanning. In this paper, we propose to improve

and extend to 3D the method presented in [1] for bone extraction in 2D images using a dynamic programming approach. As suggested in [10], we utilize the basic physical principles of US imaging coupled with information of the scanning geometry to automatically extract the bone surface from the 3D scan of a joint. After extracting the bone surface, we then utilize a simple anatomical model of the metacarpophalangeal (MCP) joints [9] to automatically segment the joint capsule region which lies between the bones forming a joint. The rest of the paper is organized as follows: Section 2 describes the 3D bone extraction in detail. The joint-capsule segmentation approach is covered in Section 3. We illustrate the proposed techniques using results on real patient data in Section 4 and conclude with a discussion about the limitations and future directions of our work in Section 5.

2. BONE SURFACE EXTRACTION

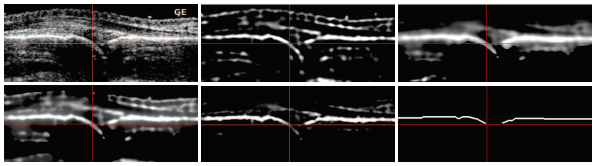


Fig. 2. Bone extraction steps *as-is* from [1]. Top-row: b-mode image (left), negated Laplacian-of-Gaussian (LoG) filtered image (center), *Shadow* image (right). Bottom-row: *reflection* image (left), bone likelihood (center) and detected bone contour (right). The linked cursor is shown in red to indicate 3D correspondence.

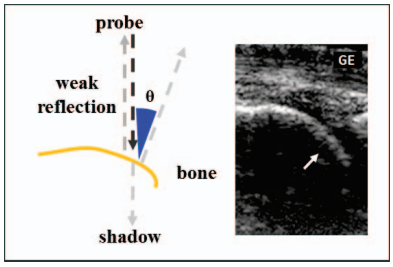


Fig. 3. Illustration of intensity drop-off as the bone curves away from the incident acoustic beam direction. White arrow (right) indicates areas of intensity drop-off.

The basic steps for bone extraction used in [1] are shown in Figure 2. For extending this method to 3D, merely applying the 2D approach to individual slices would be simple but inefficient in terms of utilizing the 3D information inherent in the data. Moreover we would be unable to address segmentation errors caused by intensity drop-off as shown in Figure 3, because the surface normal may not lie in the 2D slice, which cannot be accounted for without implicit 3D information. We propose to correct the loss in intensity by assuming a very simple model, given by:

$$I_p = \bar{I}_0 \cdot \hat{n}_p \tag{1}$$

where, $\bar{I}_0$ is the incident acoustic intensity which we want to estimate (vector notation used to incorporate direction) on the bone-surface point p, I_p is the observed image intensity and $\hat{n}_p$ is the implicit normal computed at p (where, $\hat{n}_p = \frac{\nabla I(p)}{||\nabla I(p)||}$, I denoting the 3D volume). Figure 4 illustrates the improvement in bone surface extraction after applying this correction. The subsequent filtering steps (LoG filtering, *shadow* and *reflection* image) are carried out on this corrected image. The cost function to be optimized is then given by the following set of equations:[1]

$$Cost(p) = (w_x + w_y)E_{int} + (1 - w_x)E_{ext_x} + (1 - w_y)E_{ext_y} \tag{2}$$

$$E_{int} = \begin{cases} 1 - \mathbf{L}_{bone}, & \text{bone region.} \\ 0.6, & \text{otherwise.} \end{cases} \tag{3}$$

$$E_{ext_x} = \begin{cases} \alpha_x || \frac{dC_z}{dx} - \tan\theta_{xz} ||^2 + \gamma_x, & \text{if bony.} \\ J_x, & \text{if jumping.} \\ \alpha_x {D_x}^2, & \text{otherwise.} \end{cases} \tag{4}$$

$$E_{ext_y} = \begin{cases} \alpha_y || \frac{dC_z}{dy} - \tan\theta_{yz} ||^2 + \gamma_y, & \text{if bony.} \\ J_y, & \text{if jumping.} \\ \alpha_y {D_y}^2, & \text{otherwise.} \end{cases} \tag{5}$$

where E_{int} is the internal cost defined as one minus the bone-likelihood $\mathbf{L}_{bone}$, w_x and w_y are weights on the internal cost when processing the bone curves in the azimuthal and elevational planes respectively and E_{ext_x} and E_{ext_y} are the *external* costs that enforce smoothness constraints. The derivative of the 2D bone contour in the $x-z$ and $y-z$ planes is given by $\frac{dC_z}{dx}$ and $\frac{dC_z}{dy}$ respectively. Small negative scalars (γ_x and γ_y) are used to encourage larger connected bone curves. In order to add 3D constraints we added a $\tan\theta_{xz}$ and $\tan\theta_{yz}$ term in E_{ext_x} and E_{ext_y} respectively, where θ_{xz} and θ_{yz} are the angles between the incident acoustic ray and the projection of 3D normal $\hat{n}_p$ on the azimuthal ($x-z$) and elevational ($y-z$) planes respectively. D and J correspond to the smoothness cost in the boneless region and the jump-cost respectively (refer to [1]). The dynamic programming search for the best fitting bone-surface is then carried out similar to [1]. We begin with the central (azimuthal) slice of the volume and assume that E_{ext_y} is zero. Once we have a bone contour in the central slice we can proceed with the dynamic programming search moving away from the central slice in both directions using the cost function specified in Equation 2.[2]

[1] Here x, y and z are co-ordinates of a voxel (p) along the azimuthal, elevational and axial axes respectively, where the cost function is being evaluated.

[2] Note that though the dynamic programming search proceeds slice by slice to maintain computational efficiency, the cost function to be minimized is designed to incorporate 3D information.

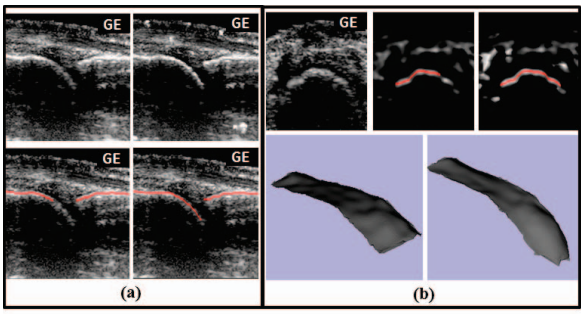


Fig. 4. Correcting drop-off effects. (a) B-mode image and detected bone-surface (red overlay) shown before (left-column) and after (right-column) correction. (b) Orthogonal cut-plane images from the same volume as in (a). Original slice (left), bone detection before (center) and after (right) correction. Bottom-row: 3D surface rendering (from one of the two bones forming the joint) before (left) and after (right) correction.

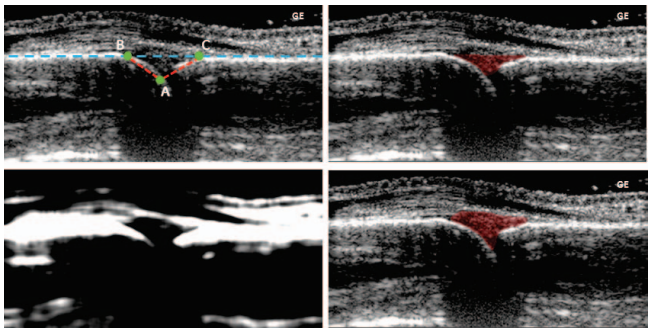


Fig. 5. Capsule segmentation. Top-left: line (blue) fitted to bone surface using weighted least-squares intersects the bone surface at two points (green). A third point is given by the *lowest* bone-surface detection. Top-right: the region inside the triangle formed by these points and *above* the bone surface is used as an initial guess for the capsule segmentation. Bottom-left: Kullback-Leibler distance (low values imply high statistical similarity with the the initial seed region). The final segmentation of the capsule region using geodesic active contour level-sets with Gradient-Vector-Flow [12].

3. JOINT-CAPSULE SEGMENTATION

Once we have the bone surface from the scan of a joint, a joint-specific anatomical model may be used to design priors or seeds for segmenting the capsule region. We have used the sonographically observed anatomical structure of an MCP joint described in [9] and illustrated in Figure 1, to design a simple algorithm for automatically segmenting the joint capsule using the following steps (refer to Figure 5).

1. In each azimuthal slice, find the lowest point (say A) of the detected bone surface.
2. Fit a line to the bone-contour (we use weighted least-squares fitting) such that it intersects the bone surface at two points (say B and C).
3. All voxels that lie (a) inside the triangle ABC and (b) *above* the detected bone surface, form the initial seed for the segmentation.
4. We use a probabilistic speckle model using Fisher-Tippett distribution [11]. The capsule region is assumed to be homogeneous and hence we can compute a single Fisher-Tippet reflectivity parameter σ_c for all voxels in the initial seed region:

$$\sigma_c{}^2 = \frac{\int_{\Omega_c}(e^{2I(x,y)} - 1)dxdy}{2\int_{\Omega_c} dxdy}, \qquad (6)$$

   where Ω_c denotes the initial (or current) capsule region.
5. The distribution parameter σ_i also is computed using equation 6 for all other voxels (denoted by suffix i) using a small 3D patch (Ω_i) around them.
6. A Kullback-Leibler distance (J-divergence) is computed for each voxel [11]:

$$J_{ci} = -1 + \frac{\sigma_c^2}{2\sigma_i^2} + \frac{\sigma_i^2}{2\sigma_c^2} \qquad (7)$$

7. This distance image is passed through a sigmoid filter and then used as a *speed* image (see Figure 5 bottom-left) for a level-sets [12] based segmentation.

Steps 4 to 7 are repeated as long as change in σ_c is nominal and there is a significant change in the capsule volume.

4. RESULTS

Figure 6 shows the extracted bone surface contours from representative slices of some 3D volumes. We applied the proposed bone-extraction method to over 30 volumes of different joints from 11 subjects. For a quantitative validation, we used expert[3] annotation of randomly selected 2D slices from 7 of the 3D volumes to compare our results (approximately 5 slices per volume). The expert annotations (red overlay) and corresponding detections (green-overlay) can be qualitatively evaluated from Figure 6 (yellow overlay indicates exact match between reference and the automated algorithm). The maximum average distance of the extracted bone with respect to the expert labelling was 0.3mm (approximately 1-2 voxels). For a detection tolerance of 1mm, we observed maximum false-positive (FP) and false-negative (FN) rates of 1.2% and 1.4% respectively, with a maximum and minimum detection rate (DR) of 98.3% and 82.8% respectively.[4] Figure 7 shows the initial seed (left-column) using our simple anatomical prior and the final segmentation based on probabilistic speckle modelling (right-column). Again, for quantitative validation we used expert annotation from representative slices containing the joint capsule. The proposed algorithm achieved average precision and recall rates

[3] Over 25 years of experience with musculoskeletal US.

[4] These statistics (average, maximum, minimum) are computed over the evaluation results from all 7 labelled datasets.

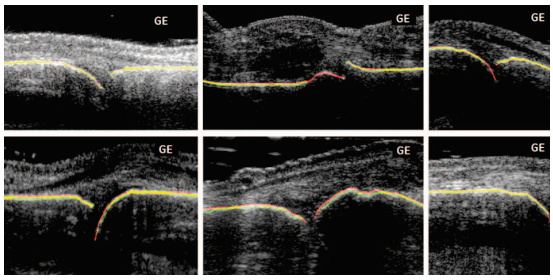


Fig. 6. A collage of 2D slices from 3D bone-extraction results on scans from multiple subjects. Expert annotations (red) and detections (green) are overlaid. Yellow overlay indicates a perfect match between the two.

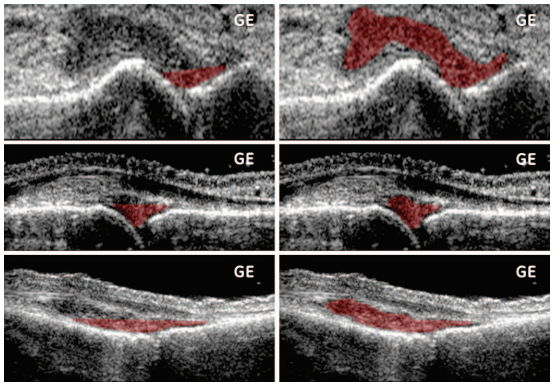


Fig. 7. Capsule segmentation results from multiple subjects. Initial seeds (left-column) and final segmentation (right-column).

of about 87% and 80% respectively. Our algorithms were implemented using experimental non-optimized code leveraging the Insight Toolkit (ITK) library in C++. The bone extraction as well as capsule segmentation took approximately 5 minutes for a 200x100x100 size volume on a laptop with an Intel© i5 processor and 4GB of RAM.

5. DISCUSSION

We have shown a novel method for automatic extraction of bone-surface and joint-capsule from the 3D US of a joint. Though these are preliminary findings, they encourage us to believe that automated tools can help objective assessment of joint pathology. Automatically extracted bone surfaces can serve as landmarks for longitudinal registration allowing for improved longitudinal assessment and quantitative tracking. Research in these directions will be reported elsewhere.

6. REFERENCES

[1] P. Foroughi, E. Boctor, M. J. Swartz, R. H. Taylor, and G. Fichtinger, "Ultrasound bone segmentation using dynamic programming," in *Proc. IEEE Ultrasonics Symp*, 2007, pp. 2523–2526.

[2] R. J. Wakefield, W. W. Gibbon, P. G. Conaghan, P. O'Connor, D. McGonagle, C. Pease, M. J. Green, D. J. Veale, J. D. Isaacs, and P. Emery, "The value of sonography in the detection of bone erosions in patients with rheumatoid arthritis: a comparison with conventional radiography.," *Arthritis Rheum*, vol. 43, no. 12, pp. 2762–2770, Dec 2000.

[3] M. Backhaus, G. R. Burmester, D. Sandrock, D. Loreck, D. Hess, A. Scholz, S. Blind, B. Hamm, and M. Bollow, "Prospective two year follow up study comparing novel and conventional imaging procedures in patients with arthritic finger joints.," *Ann Rheum Dis*, vol. 61, no. 10, pp. 895–904, Oct 2002.

[4] U. M. Döhn et al, "Are bone erosions detected by magnetic resonance imaging and ultrasonography true erosions? a comparison with computed tomography in rheumatoid arthritis metacarpophalangeal joints.," *Arthritis Res Ther*, vol. 8, no. 4, pp. R110, 2006.

[5] I. Hacihaliloglu, R. Abugharbieh, A. J. Hodgson, and R. N. Rohling, "Enhancement of bone surface visualization from 3D ultrasound based on local phase information," in *Proc. IEEE Ultrasonics Symp*, 2006, pp. 21–24.

[6] L. Lopez-Perez, J. Lemaitre, A. Alfiansyah, and M. Bellemare, "Bone surface reconstruction using localized freehand ultrasound imaging," in *Proc. 30th Annual Int. Conf. of the IEEE Engineering in Medicine and Biology Society EMBS 2008*, 2008, pp. 2964–2967.

[7] H. Overhoff and D. Lazovic, "Visualization of the newborn's hip joint using 3-d ultrasound and automatic image processing," 1999.

[8] Y. Zhang, R. Rohling, and D. K. Pai, "Direct surface extraction from 3d freehand ultrasound images," in *Proc. IEEE Visualization VIS 2002*, 2002, pp. 45–52.

[9] N. Boutry, A. Lard, X. Demondion, B. Cortet, H. Cotten, and A. Cotten, "Metacarpophalangeal joints at us in asymptomatic volunteers and cadaveric specimens.," *Radiology*, vol. 232, no. 3, pp. 716–724, Sep 2004.

[10] J. A. Noble and D. Boukerroui, "Ultrasound image segmentation: a survey," vol. 25, no. 8, pp. 987–1010, 2006.

[11] Z. Wang, G. Slabaugh, G. Unal, and T. Fang, "Registration of ultrasound images using an information-theoretic feature detector," in *Proc. 4th IEEE Int. Symp. Biomedical Imaging: From Nano to Macro ISBI 2007*, 2007, pp. 736–739.

[12] X. Han, Xu. C., and J. L. Prince, "A topology preserving level set method for geometric deformable models," vol. 25, no. 6, pp. 755–768, 2003.