FULLY AUTOMATIC AND FAST SEGMENTATION OF THE FEMUR BONE FROM 3D-CT IMAGES WITH NO SHAPE PRIOR

Marcel Krčah[12] *Gábor Székely*[1] *Rémi Blanc*[1]

[1] Computer Vision Laboratory, ETH Zurich, Zurich, Switzerland
[2] Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic

ABSTRACT

Statistical shape and intensity modeling have been subject to an increasing interest within the past decade. However, construction of such models requires large number of initial samples. Accurate and automatic segmentation techniques that do not require any explicit prior model are therefore of high interest. We propose a fully-automatic method for segmenting femur in 3D CT volumes, based on graph-cuts and a novel bone boundary enhancement filter. The presented framework has been analyzed and compared to other three automatic bone segmentation techniques in large-scale experiments, conducted on 197 femur samples. Our approach was capable of segmenting the femur accurately in 81% of cases without any shape prior or user interaction.

Index Terms— CT, bone segmentation, graph-cuts, sheetness measure, femur

1. INTRODUCTION

Statistical shape and intensity modeling have been subject to an increasing interest within the past decade in a wide variety of applications [1]. In particular in the domain of orthopaedics, they are of high interest for intra-operative guidance, reconstructive surgery or implant design.

However, learning such models require that large number of samples are available and pre-processed. Segmentation, in particular, remains a serious bottleneck and manual processing remains common practice. Accurate and automatic segmentation methods that do not require any explicit prior model remain therefore of high interest in this context.

While segmentation of long bones in CT images is relatively simple in the diaphysis, due to a high contrast between the thick high-intensity cortical layer and the encasing low-intensity soft tissue, the segmentation still remains a challenge in the joint areas. In the epiphysis, the cortical layer becomes much thinner and cancellous bone does not present as much contrast with soft tissues. Additionally, the inter-bone space can become very narrow and partial volume effects result in a very weak contrast in these regions (Fig. 1).

This work has been supported by the Swiss National Science Foundation.

As pointed out in [2], despite several years of active research, the challenges posed by bone segmentation are not yet fully solved. Intensity-based methods, such as (local) binary thresholding or region growing, tend to produce discontinuous contours and "leakages" to soft-tissues or adjacent bones. Active contour models, such as snakes or level-set methods [3], are sensitive to initialization, and can have problems in areas of low gradient. More recently, the graph-cut [4, 5] was shown to provide an elegant and efficient framework for segmentation. It was applied to bone segmentation in [6] with promising results. We extend this approach to a fully 3D formulation. In particular, we propose new energy terms for the graph-cut, including a novel bone boundary detector relying on the sheetness measure [7]. A post-processing step is also proposed for the automatic separation of adjacent bones.

The details of the methods are given in Section 2. In section 3, we evaluate our method on a database of 197 cases, for which both the original CT and the ground-truth manual segmentation were available, and compare it with 3 other automatic methods. Section 4 concludes the paper.

2. OVERVIEW OF THE SEGMENTATION PROCEDURE

Given a 3D CT image $\mathcal{I} : \Omega \to R$, the goal is to find a binary labelling $L : \Omega \to \{0, 1\}$ assigning 1 to all voxels belonging to the femur and 0 to all other (background) voxels.

The proposed method is a two-step procedure. (Fig. 1) First, all bone voxels in the input image are located. Then, the connected component containing the femur is identified and other bones adjacent to the femur (patella, tibia and pelvis) are removed.

Both steps are based on Boykov and Jolly's graph-cut method [5], an optimization algorithm from the combinatorial graph theory. The graph-cut segmentation requires definition of two functions representing soft-constraints imposed on the segmentation. The *regional* cost function $R_p(A_p) : \Omega \times \{0, 1\} \to R_0^+$ specifies an individual penalty for assigning the label A_p to the voxel p. The *boundary* cost function $B(p, q) : \mathcal{N} \to R_0^+$ reflects a penalty for classifying the voxel p as an object and the voxel q as background, within a symmetric neighborhood system $\mathcal{N} \subset \Omega^2$. The boundary cost

Fig. 1: Individual steps of the proposed procedure illustrated on an axial slice of the acetabulofemoral joint: original image (left), identification of all bone voxels (middle), final result after removing bones adjacent to femur (right).

function is usually defined as symmetric; however, we exploit directed edges of the underlying combinatorial graph (see [5]) to obtain more precise bone boundaries. Result of the graph-cut segmentation is a binary labelling A globally minimizing the function $E(A)$:

$$E(A) = \sum_{p \in \Omega} R_p(A_p) + \lambda \sum_{(p,q) \in \mathcal{N}} \delta(A_p, A_q) B(p,q), \quad (1)$$

where δ is the Kronecker delta and λ defines the relative importance of the regional term and the boundary term.

Sections 2.1 and 2.2 detail the boundary and regional cost function employed in the first segmentation step, respectively. Section 2.3 explains the procedure of adjacent bones removal.

2.1. Boundary cost function

In the case of bone segmentation, the boundary cost function should encourage spatial coherence within the bone and within the background. Thus, the penalty should be low on bone boundaries and high elsewhere. As explained in Section 1, the cost function based on intensities leads to unsatisfactory results. Contrast in the epiphysis region is therefore enhanced by successively applying two filters.

First, the input image $\mathcal{I}$ is filtered with the unsharp masking, defined as $\mathcal{I}^U = \mathcal{I} + k(\mathcal{I} - \mathcal{I} \star G_s)$, where $\star$ denotes convolution, G_s is a Gaussian kernel with variance s^2 and k denotes a scaling constant. The bone boundaries are further enhanced by employing a novel sheetness filter, inspired by [7]. For each voxel, a sheetness score is computed from the eigenvalues $|\lambda_1| \leq |\lambda_2| \leq |\lambda_3|$ of the local Hessian matrix as:

$$S_\sigma(x) = -\operatorname{sgn}(\lambda_3) \exp\left\{-\frac{R_{sheet}^2}{\alpha^2}\right\} \exp\left\{-\frac{R_{tube}^2}{\beta^2}\right\} \left(1 - \exp\left\{-\frac{R_{noise}^2}{\gamma^2}\right\}\right),$$

with $R_{sheet} = |\lambda_2|/|\lambda_3|$, $R_{tube} = |\lambda_1|/(|\lambda_2||\lambda_3|)$ and $R_{noise} = (|\lambda_1| + |\lambda_2| + |\lambda_3|)/T$, where T denotes mean trace of all computed Hessians. In order to further enhance the contrast on the bone boundaries, we have incorporated the sign of the largest eigenvalue into the sheetness score. Finally, we improve the robustness of the sheetness through a multi-scale implementation, by computing the Hessian matrix at different scales Σ and retaining, for each voxel, the corresponding sheetness response that is the highest in absolute value.

The result of these filters, simply denoted S in the following, greatly enhances bone boundaries, especially in narrow inter-bone channels (Fig. 2a). We employ it for defining the boundary cost function of the graph-cut algorithm:

$$B(p,q) \propto \begin{cases} \exp\left\{-\dfrac{|S(p) - S(q)|}{\sigma_s}\right\} & \text{for } S(p) \geq S(q), \\ 1 & \text{otherwise,} \end{cases}$$

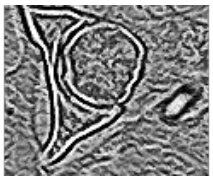
where σ_s is a parameter. This function encourages boundaries in regions with abrupt variations of the sheetness value. Additionally, more precise bone boundaries are achieved through exploiting the asymmetry since the cost function forces transitions from bone to background in regions with decreasing sheetness score.

2.2. Regional cost function

The regional cost function $R_p(A_p)$ reflects penalty for assigning the voxel p the label A_p. For the bone segmentation purpose, the cost should be low if (a) p belongs to a bone and $A_p = 1$ or (b) if p belongs to background and $A_p = 0$; otherwise the cost should be high. Weak bone boundaries, narrow inter-bone space and low intensities in the trabecular bone make image intensity alone a relatively poor feature for discriminating bone from background. Nevertheless, exclusion regions E_{notbkg} and $E_{notbone}$ can, with good confidence, be defined as:

$$E_{notbkg} = \{x \in \Omega | \mathcal{I}(x) \geq 400 \wedge S(x) > 0\},$$
$$E_{notbone} = \operatorname{lcc}(\{x \in \Omega | \mathcal{I}(x) < -50\}),$$

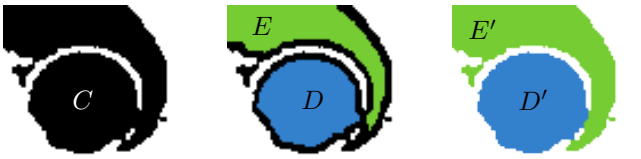
where lcc denotes largest connected component of the input binary argument. The term $\mathcal{I}(x) \geq 400$ selects high-intensity voxels present mostly in cortical bone while the term $S(x) > 0$ removes soft-tissue voxels in narrow inter-bone channels.



(a) The input image filtered successively with the unsharp masking and the multi-scale sheetness measure.

(b) Estimation of exclusion regions E_{notbkg} (blue) and $E_{notbone}$ (green).

Fig. 2: Illustration of intermediate results in the first step of the bone segmentation.



(a) Binary input with the femur not properly separated from adjacent bones.

(b) Morphological erosion divides the component C into two subcomponents D_1, D_2.

(c) Final result after applying a simplified graph-cut.

Fig. 3: Identification of adjacent bones.

The term $\{x|\mathcal{I}(x) < -50\}$ chooses fat and air voxels of low intensity values, present rarely in the bone interior.

The regional cost function for the bone segmentation is then defined as

$$R_p(A_p) \propto \begin{cases} 1 & \text{if } A_p = \text{“bone“ and } p \in E_{notbone}, \\ 1 & \text{if } A_p = \text{”bkg” and } p \in E_{notbkg}, \\ 0 & \text{otherwise.} \end{cases}$$

Since the regional cost is zero for all non-estimated voxels $\Omega \setminus (E_{notbkg} \cap E_{notbone})$, it is the boundary cost function that arbitrates on the classification.

2.3. Postprocessing - Femur Separation

Applying the graph-cut segmentation with the cost functions defined above generates a binary volume where adjacent bones (pelvis, tibia and patella) are likely to be connected. Experiments have shown that when such a defect occurs, the adjacent bone is connected to the femur generally by only a few voxels (Fig. 1 and 3a)

To separate the neighboring bones, a morphological erosion with a spherical element of radius R is applied on the binary image (Fig. 3b). If the erosion divides a component $C \subset \Omega$ into two subcomponents $D, E \subset C$, a bottleneck has to be present in C . Our goal is to find disjoint sets $D', E' \subset \Omega$ such that $D' \cup E' = C, D \subset D', E \subset E'$ and that the number of voxels on the boundary between both sets is minimal. To accurately find D', E', a simplified graph-cut is employed on the image part C. The hard-constraints $D \subset D', E \subset E'$ are imposed on the classification via the regional cost function:

$$\forall p \in C: \quad R_p(A_p) = \begin{cases} \infty & \text{if } A_p = \text{”D” and } p \in E, \\ \infty & \text{if } A_p = \text{”E” and } p \in D, \\ 0 & \text{otherwise.} \end{cases}$$

The boundary cost function $B(p,q) = B(q,p) = 1$ encourages minimal boundary between D' and E'. This procedure can be employed to either segment all bones in the image by assigning them different labels or to extract one of them (Fig. 3c).

3. EXPERIMENTS AND RESULTS

Experiments were conducted on 197 CT scans, roughly cropped around the femur, with image intensities expressed as Hounsfield units. The volume spacing ranged from 0.6 to 1.17mm in-plane and from 0.8 to 1.25mm out-plane.

For all datasets, a manual segmentation by a medical expert was available and served as a reference for quantitative evaluations. For each sample, we measured the True Positive Rate TPR = TP/(TP+FN), False Positive Rate FPR = FP/(FP+TN) and the Hausdorff distance (HD) between the estimated and the ground-truth surfaces. Additionally, the time to process and consumed RAM memory were recorded.

3.1. Experiment design

The proposed method has been implemented in C++ using the ITK library (http://itk.org) and the Kolmogorov max-flow library [4] for efficient computation of the graph-cut. The following parameters have been used: $\lambda = 5, k = 10, s = 1, \alpha = \beta = 0.5, \gamma = 0.25, \Sigma = \{0.75, 1.0\}, \sigma_s = 0.2, R = 3$, and were kept constant for all datasets.

We compared our method to three existing fully-automatic 3D-based bone segmentation schemes: gradient-based geometric active contour (GeomAC) [3], Zhang iterative adaptive thresholding (ZIAT) [8] and an intensity-based graph-cut method (IBGC). For GeomAC, the ITK implementation was employed with the initial level-set set being the signed distance of the binary image $E_{notbone}$, propagation parameter 5.0, curvature parameter 1.0 and max. number of iterations 400. The ZIAT method was implemented with a window size 15x15x3 and initial threshold value 70. For IBGC, the graph-cut algorithm was employed with the following cost functions: $R_p(\text{bone}) = 1$ for $p \in E_{notbone}$, 0 otherwise; $R_p(\text{bkg}) = 1$ for $\mathcal{I}(p) > 500$, 0 otherwise; $B(p,q) = B(q,p) = \exp\{|I(p) - I(q)|/100\}$.

3.2. Results

Results (Fig. 4 and 5) show that all tested approaches, except ZIAT, are able to correctly discriminate the true femur voxels with satisfactory precision (TPR > 0.85). However, our approach is clearly superior in detecting bone boundaries in narrow inter-bone regions and in suppressing leakages into bones adjacent to femur. Our approach correctly segmented and separated femur from adjacent bones (TPR > 0.85, FPR < 0.001, HD $< 8mm$) in 81%, with average HD within this set of 5.4mm. The bone removal procedure was effective in 57% of cases.

Due to the low contrast in narrow inter-bone channels and weak bone boundaries, leakages (GeomAC, IBGC, ZIAT) and contour discontinuities (ZIAT) always occurred in segmentations of the three methods. None of them could separate the femur accurately (HD $> 5cm$), nor segment the challenging

(a) Original image (b) Ground truth (c) GAC (d) Zhang (e) IBGC (f) Our aprroach

Fig. 4: Acetabulofemoral joint segmentation. (todo - resize, colors, text, choose maybe better examples)

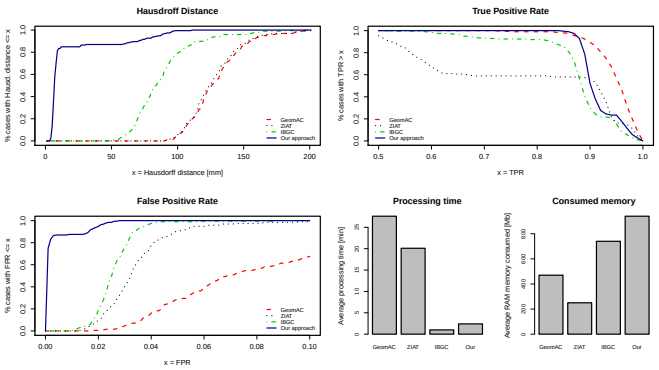


Fig. 5: Results of the experiments conducted on 197 femoral samples (todo-larger labels, TPR and FPR on the same line.

weak bone contours (visual evaluation). Graph-cut-based approaches have also proven to be much faster (only a few minutes per dataset), albeit consuming more memory. However, the high memory requirements can be decreased by employing alternative graph-cut algorithms (e.g. [9]).

4. CONCLUSION

We have presented a fully automatic method for segmenting the femur from 3D-CT images, based on the graph-cut algorithm and a novel bone boundary detector. The strengths of the new sheetness score tied with the automatic bone separator have been proven in the large-scale experiments conducted on 197 femur CT samples. The femur has been segmented accurately and separated from other adjacent bones in 81% of cases. The method has been also compared to other existing fully-automatic solutions; however, all of them require manual postprocessing in order to achieve satisfactory results.

As no explicit priors are required, the method is also suitable for segmenting other bones in CT. Preliminary experiments on the forearm showed promising results (Fig. 6), without any parameter tuning.

Future work aims at embedding this automatic segmentation framework into a statistical model generation approach, by iteratively improving the model and refining the segmentations.

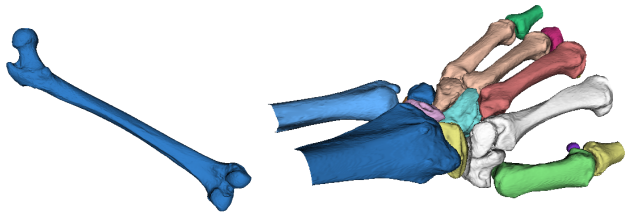


Fig. 6: todo caption, add radius/ulna.

5. REFERENCES

[1] Tobias Heimann and Hans-Peter Meinzer, "Statistical shape models for 3d medical image segmentation: A review," *Medical Image Analysis*, vol. 13, no. 4, pp. 543 – 563, 2009.

[2] L.I. Wang, M. Greenspan, and R. Ellis, "Validation of bone segmentation and improved 3-D registration using contour coherency in CT data," *IEEE transactions on medical imaging*, vol. 25, no. 3, pp. 324, 2006.

[3] R. Malladi, J.A. Sethian, B.C. Vemuri, et al., "Shape modeling with front propagation: A level set approach," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, no. 2, pp. 158–175, 1995.

[4] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1124–1137, 2004.

[5] Y. Boykov and G. Funka-Lea, "Graph cuts and efficient nd image segmentation," *International Journal of Computer Vision*, vol. 70, no. 2, pp. 109–131, 2006.

[6] P. Fürnstahl, T. Fuchs, A. Schweizer, L. Nagy, G. Székely, and M. Harders, "Automatic and robust forearm segmentation using graph cuts," *Biomedical Imaging: From Nano to Macro*, 2008.

[7] M. Descoteaux, M. Audette, K. Chinzei, and K. Siddiqi, "Bone enhancement filtering: Application to sinus bone segmentation and simulation of pituitary surgery," *Lecture Notes in Computer Science*, vol. 3749, pp. 9, 2005.

[8] J. Zhang, C.H. Yan, C.K. Chui, and S.H. Ong, "Fast segmentation of bone in CT images using 3D adaptive thresholding," *Computers in Biology and Medicine*, vol. 40, no. 2, pp. 231–236, 2010.

[9] A. Delong and Y. Boykov, "A scalable graph-cut algorithm for nd grids," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. IEEE, 2008, pp. 1–8.