AUTOMATIC THREE-LABEL BONE SEGMENTATION FROM KNEE MR IMAGES

Liang Shan[†]*, Christopher Zach*[+]*, Marc Niethammer*[†]

[†]Department of Computer Science, University of North Carolina, Chapel Hill, NC, USA
[+]Department of Computer Science, ETH Zürich, Switzerland

ABSTRACT

We propose a novel fully automatic three-label bone segmentation approach applied to knee segmentation (femur and tibia) from T1 and T2* magnetic resonance (MR) images. The three-label segmentation approach guarantees separate segmentations of femur and tibia which cannot be assured by general binary segmentation methods. The proposed approach is based on a convex optimization problem by embedding label assignment into higher dimensions . Appearance information is used in the segmentation to favor the segmentation of the cortical bone. We validate the proposed three-label segmentation method on nine knee MR images against manual segmentations for femur and tibia.

***Index Terms*—** three-label segmentation, shape, appearance, convex optimization, globally optimal

1. INTRODUCTION

Osteoarthritis (OA) is one of the major causes of long-term disability affecting millions of people. With no cure available further research into potential treatments is necessary. To this end large longitudinal image databases have been acquired for example by the Osteoarthritis Initiative. However, a comprehensive analysis of this imaging data is hindered by a lack of fully-automatic analysis methods. Cartilage loss [1] is believed to be the dominating factor in OA and hip and knee involvement are most common. With magnetic resonance imaging (MRI) becoming increasingly accepted for the assessment of OA, we propose a new method for the segmentation of femur and tibia from MR images. Our segmentation method is designed for three-label segmentation problems and is therefore particularly well suited for the simultaneous segmentation of bones within a joint, because it guarantees (unlike binary segmentation approaches) that bones are segmented as separate entities even if they touch. The segmentation framework is general and should therefore also be applicable to segmentation problems where touching structures are frequent, such as the segmentation of femoral and tibial cartilage.

The authors thank Dr. Cecil Charles, Duke University, for providing data sets through the Pfizer Longitudinal study and gratefully acknowledge support by NSF EECS-0925875.

Many methods have been applied to bone and cartilage segmentation, e.g., region growing approaches [2], medial models, active shape models [3], general deformable models (such as live-wire, active contour or active surface models) [4], clustering methods [5], and graph-based approaches [6]. Most of these methods require user interaction to place seeds or identify landmarks or are not designed for the simultaneous segmentation of potentially touching objects. Further, they are frequently based on non-convex optimization problems which are difficult to solve. Graph cuts and their recently proposed continuous formulations [7, 8] allow for the computation of globally optimal solutions: the three-label segmentation method described in this paper falls into the same class of approaches. We make use of a formulation for depth-map computations proposed in [9] and use it to extend our previously proposed (binary) bone segmentation method [10]. This results in a fully-automatic bone segmentation method, which is easy to compute and allows for reliable segmentations. Sec. 2 introduces the three-label approach, results are given in Sec. 3. The paper concludes with a summary and future work.

2. THREE-LABEL BONE SEGMENTATION

The energy minimized for a general binary segmentation problem proposed in [10] is

$$E(w) = \int_\Omega g\|\nabla w\| + rw + F \cdot \nabla w \, d\Omega, \; w \in [0,1]. \quad (1)$$

This is a convex optimization problem (for which a globally optimal solution can be obtained), where Ω is the image domain, w the indicator function (0 for background and 1 foreground), g is regularization weight, r is a regional bias, and F a vector field which can be used to incorporate orientation-dependent appearance information. In [10] we proposed an automatic bone segmentation method based on this framework. For OA, clear separation of bones may not always be observed in MR images for severely pathological cases. Consequentially, a binary segmentation cannot guarantee a separation of femur and tibia. We therefore propose a three-label segmentation method, which guarantees this separation. Fig. 1 illustrates the limitations of a binary versus a three-label segmentation method for a synthetic case.

(a) (b)

Fig. 1. Synthetic example. (a) Binary segmentation result. Femur and tibia are segmented as one object and the boundary in the joint region is not captured well due to regularization effects. (b) Proposed three-label segmentation. The boundaries between bones and background are preserved.

Zach et al.[9] formulated the associated multi-label segmentation problem by expressing the labels through a separate spatial dimension. A multi-label segmentation is a mapping from an image domain Ω to a label space represented by a set of non-negative integers, i.e. $\mathcal{L} = \{0, ..., L-1\}$. The labeling function $\Lambda : \Omega \rightarrow \mathcal{L}, \mathbf{x} \mapsto \Lambda(\mathbf{x})$ maps a pixel $\mathbf{x}$ in the image domain Ω to label $\Lambda(\mathbf{x})$ in label space. The goal is to find a labeling function that minimizes an energy functional of the form:

$$E(\Lambda) = \int c(\mathbf{x}, \Lambda(\mathbf{x})) + V(\nabla\Lambda, \nabla^2\Lambda, ...)\, d\Omega, \quad (2)$$

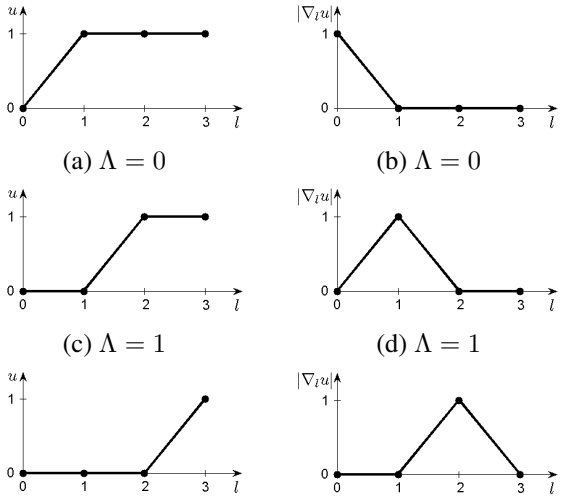
where $c(\mathbf{x}, \Lambda(x))$ is the cost of assigning label $\Lambda(x)$ to pixel $\mathbf{x}$ and $V(\cdot)$ is a regularizing term. The different labellings can be encoded through a level function u

$$u(\mathbf{x}, l) = 1 \text{ if } \Lambda(\mathbf{x}) < l, \quad 0 \text{ otherwise}, \quad (3)$$

which maps the Cartesian product of the image domain Ω and the labeling space $\mathcal{L}$ to $\{0, 1\}$. By definition, we have $u(\mathbf{x}, 0) = 0$ and $u(\mathbf{x}, L) = 1$. Of note, unlike the indicator function w in Eq.(1), u does not directly encode labels, but instead defines them through its discontinuity set. Fig.(2) illustrates the relation between u and Λ for the three-label case. This setup is in general asymmetric with respect to the labels, since the design of the level function implies a specific label ordering. However, for the three-label case, the background label (1) can be symmetrically positioned between the two object labels (i.e., femur (0) and tibia (2)). The three-label segmentation energy becomes

$$E(u) = \int_{\mathcal{D}} g\|\nabla_{\mathbf{x}} u\| + c\,|\nabla_l u|\; d\mathbf{x}dl,\; \mathcal{D} = \Omega \times \mathcal{L}, \quad (4)$$

where, $\nabla_{\mathbf{x}} u$ denotes the spatial gradient of u, $\nabla_{\mathbf{x}} u = (\partial u/\partial x, \partial u/\partial y, \partial u/\partial y)^T$ and similarly $\nabla_l u$ the gradient in the label direction, $\nabla_l u = \partial u/\partial l$. Functions g, u and c are four-dimensional scalar fields. In our implementation,



(a) $\Lambda = 0$ (b) $\Lambda = 0$

(c) $\Lambda = 1$ (d) $\Lambda = 1$

(e) $\Lambda = 2$ (f) $\Lambda = 2$

Fig. 2. Values of u and $\nabla_l u$ for different label assignments in a three-label segmentation. Assuming a discretization with forward differences. $\nabla_l u$ determines the label assignment.

we set g to a positive constant. The data-costs are defined by the log-likelihoods of femur ($l = 0$), tibia ($l = 2$), and background ($l = 1$)

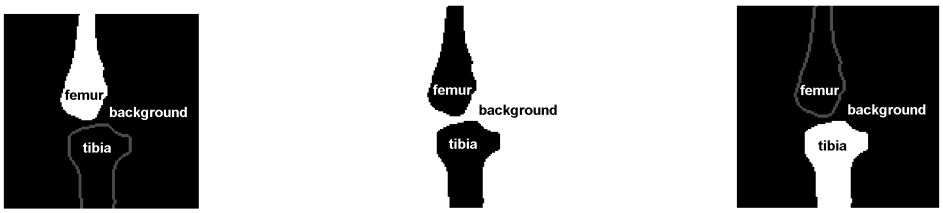
$$c(\mathbf{x}, l) = -log(P(l|I(\mathbf{x}))) = -log\left(\frac{P(I(\mathbf{x})|l) \cdot P(l)}{P(I(\mathbf{x}))}\right),$$

where $P(I(\mathbf{x})|l)$ are obtained given a reference segmentation through the joint histogram for $I = \{I_{T1}, I_{T2}\}$ by Parzen windowing. Initially, $P(l) = \frac{1}{3}$. The prior probabilities are updated by the alignment of a shape model [10]. Shape model alignment (to the likelihood image in the first iteration and to the following segmentations in the following iterations) and segmentation are iterated to convergence. In practice, satisfying segmentations are typically obtained after two iterations. For the binary segmentation [10], appearance information $F = \beta(\nabla I_{T1} + \nabla I_{T2})$, $\beta \in \mathbb{R}^+$ can be introduced to capture the outer boundary of the cortical bone which appears dark in both T1 and T2* images. This information can be incorporated into the three-label segmentation through the modified energy

$$E(u) = \int_{\mathcal{D}} g\,\|\nabla_{\mathbf{x}} u\| + c\,|\nabla_l u| + F \cdot \nabla_{\mathbf{x}}(|\nabla_l u|)\; d\mathbf{x}dl, \quad (5)$$

where F is a four-dimensional vector field, representing the appearance information. Since u is a level function the design of F depends on $|\nabla_l u|$, which is illustrated in Fig.3. In our bone segmentation task, we choose F as (Femur is labeled 0, background 1 and tibia 2)

$$F(\mathbf{x}, l) = \begin{cases} \beta(\nabla I_{T1}(\mathbf{x}) + \nabla I_{T2}(\mathbf{x})),\ \beta \in \mathbb{R}^+, & \text{if } l = 0, \\ \mathbf{0}, & \text{if } l = 1, \\ \beta(\nabla I_{T1}(\mathbf{x}) + \nabla I_{T2}(\mathbf{x})),\ \beta \in \mathbb{R}^+, & \text{if } l = 2. \end{cases} \quad (6)$$



(a) $l = 0$ (b) $l = 1$ (c) $l = 2$

Fig. 3. $|\nabla_l u|$ for different labels (femur is labeled 0, background 1 and tibia 2) to justify the choice for F in Eq. (6). Black is 0 and white is 1. Boundaries of femur and tibia in (a) and (c) are shown in gray for visualization only and correspond to 0.

We rewrite Eq. (5) as

$$E(u) = \int_{\mathcal{D}} g \left\|\nabla_{\mathbf{x}} u\right\| + (c - \operatorname{div}(F)) \left|\nabla_l u\right| \, d\mathbf{x}dl, \quad (7)$$

and shift $c - \operatorname{div}(F)$ to non-negative range by adding a common offset, which does not change the global minimizer. Then we introduce the two dual variables $\mathbf{p}$ and q and obtain

$$E(u, \mathbf{p}, q) = \int_{\mathcal{D}} -\mathbf{p} \cdot \nabla_{\mathbf{x}} u - q \cdot \nabla_l u \, d\mathbf{x}dl, \quad (8)$$
$$\text{subject to } \|\mathbf{p}\| \leq g, \ |q| \leq c - \operatorname{div}(F),$$

which is minimized with respect to u and maximized with respect to $\mathbf{p}$ and q to obtain a minimizer for Eq. (5). Here, $\mathbf{p}$ is a 4-dimensional vector field and q is a 4-dimensional scalar field. The solution can be computed by the following gradient descent/ascent

$$\mathbf{p}_t = -\nabla_{\mathbf{x}} u, \ \|\mathbf{p}\| \leq g, \quad (9)$$

$$q_t = -\nabla_l u, \ |q| \leq c - \operatorname{div}(F), \quad (10)$$

$$u_t = -(\operatorname{div}(\mathbf{p}) + \nabla_l q), \ u \in [0, 1]. \quad (11)$$

Because of the convexity of the optimization problem the iterative scheme yields a global optimum. The dual energy to (5) is given [9] by

$$E^*(u) = \int_S \operatorname{div}(\mathbf{p}) + \nabla_l q \, d\mathbf{x}dl + \int_{\mathring{\mathcal{D}}} \min(0, \operatorname{div}(\mathbf{p}) + \nabla_l q) \, d\mathbf{x}dl, \quad (12)$$

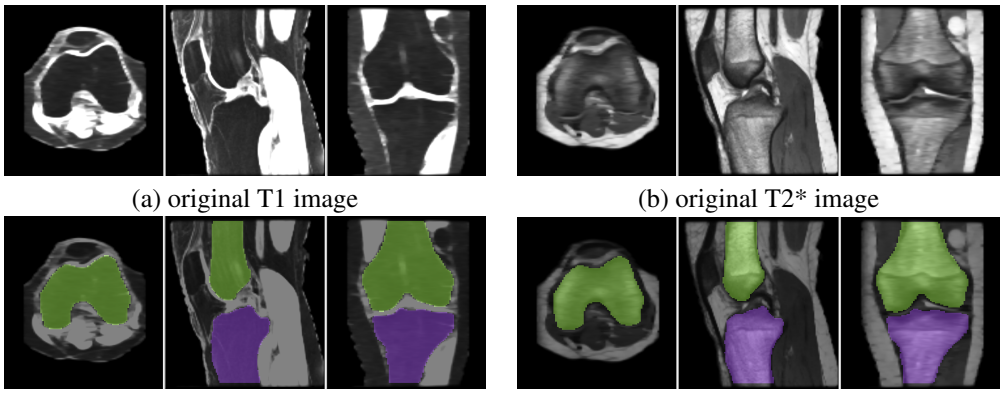
where $S = \Omega \times \{3\}$ is the source set (where $u = 1$ is enforced), $T = \Omega \times \{0\}$ is the sink set (where $u = 0$ is enforced) and $\mathring{\mathcal{D}} = \mathcal{D} \setminus (S \cup T) = \Omega \times \{1, 2\}$. The dual energy can be used to terminate iterations once a sufficiently small duality gap (difference between primal and dual energy) is reached. After convergence, the solution u is essentially binary and monotonically increasing. The three-label segmentation can be recovered from the discontinuity set of u.

3. RESULTS AND VALIDATION

We demonstrate the behavior of the three-label segmentation on a synthetic case and of the complete proposed segmentation algorithm on a set of nine real MR images (T1 weighted 3D SPGR and T2* 3D GRE images acquired sagittally at a resolution of $0.29 \times 0.29 \times 1.5$ mm^3 at 3T). Preliminary results were obtained on images downsampled to an approximately isotropic resolution of $1.17 \times 1.17 \times 1.5$ mm^3. Fig. (1) shows the beneficial behavior of the three-label segmentation method compared to binary segmentation for the synthetic test case. Fig. 4 shows an exemplary three-label bone segmentation overlaid onto the original T1 and T2* images. Femur and tibia are segmented simultaneously and the outer boundary of the cortical is captured as desired (see (c) and (d)). We validated the obtained segmentation results against manual segmentations of femur and tibia. Each of the nine cases was segmented eight times using shape models generated from the manual segmentations for femur and tibia of the remaining eight cases. Three axial slices of the manual and three-label segmentations were cropped at the superior and inferior image boundaries. We obtained a median Dice coefficient of 0.953 for femur and 0.938 tibia (see Table 1).

		AVDP	ASD	MSD	DC
Femur	Median	2.14%	0.61	5.19	95.3%
	Mean	2.46%	0.65	5.45	95.1%
	SD	2.15%	0.16	2.22	1.23%
Tibia	Median	6.60%	0.78	6.04	93.8%
	Mean	6.02%	0.84	7.79	93.4%
	SD	3.77%	0.23	4.37	1.49%

Table 1. Statistics of validation results. SD is standard deviation. AVDP is absolute volume difference percentage, ASD is average surface distance, MSD is maximum surface distance and DC is Dice coefficient. All distances in mm.



(a) original T1 image (b) original T2* image

(c) three-label segmentation overlaid onto the original T1 image (d) three-label segmentation overlaid onto the original T2* image

Fig. 4. Three-label segmentation overlaid onto original images.

4. CONCLUSION AND FUTURE WORK

In this paper, we proposed a fully automatic three-label segmentation method for bone segmentation and applied it to the segmentation of femur and tibia. Unlike binary segmentation, the three-label segmentation approach guarantees the separation of bones even if they touch. We formulated the segmentation approach as a convex optimization problem in four dimensions assuring (given an aligned shape model) a globally optimal solution. We further proposed a method to integrate simple appearance information to allow for segmentations of the outside of the cortical bone in MR images. Validation results confirm the performance of the proposed algorithm. The obtained bone segmentations may subsequently be used [10] to automatically transform MR images of the knee to a reference configuration. Further, the three-label segmentation algorithm should be particularly well suited for knee cartilage segmentations (to disambiguate the touching femoral and tibial cartilage) which we will explore in future work.

5. REFERENCES

[1] D. T. Felson et al., "Osteoarthritis: New insights, part 1: The disease and risk factors," *Annals of Internal Medicine*, vol. 133, pp. 635–646, 2000.

[2] J. G. Tamez-Pena, S. Totterman, and K. J. Parker, "Unsupervised statistical segmentation of multispectral volumetric MR images," *Proceedings of the SPIE Conference on Image Processing*, vol. 3661, pp. 300–311, 1999.

[3] J. Fripp, S. Crozier, S. K. Warfield, and S. Ourselin, "Automatic segmentation of bone and extraction of the bone-cartilage interface from magnetic resonance images of the knee," *Physics in Medicine and Biology*, vol. 52, pp. 1617–1631, 2007.

[4] M. H. Brem et al., "Magnetic resonance image segmentation using semi-automated software for quantification of knee articular cartilage - initial evaluation of a technique for paired scans," *Skeletal Radiology*, vol. 28, no. 5, pp. 505–511, 2009.

[5] J. Folkesson, E. B. Dam, O. F. Olsen, P. C. Pettersen, and C. Christiansen, "Segmenting articular cartilage automatically using a voxel classification approach," *IEEE Transactions on Medical Imaging*, vol. 26, no. 1, pp. 106–115, 2007.

[6] H. Shim, S. Chang, C. Tao, J.-H. Wang, C. K. Kwoh, and K. T. Bae, "Knee cartilage: Efficient and reproducible segmentation on high-spatial-resolution MR images with the semiautomated graph-cut method," *Radiology*, vol. 251, no. 2, pp. 548–556, 2009.

[7] B. Appleton and H. Talbot, "Globally mimimal surfaces by continuous maximal flows," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 28, no. 1, pp. 106–118, 2006.

[8] X. Bresson, S. Esedoglu, P. Vandergheynst, J.-P. Thiran, and S. Osher, "Fast global minimization of the active contour/snake model," *Journal of Mathematical Imaging and Vision*, vol. 28, pp. 151–167, 2007.

[9] C. Zach, M. Niethammer, and J. M. Frahm, "Continuous maximal flows and Wulff shapes: Application to MRFs," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1911–1918, 2009.

[10] L. Shan, C. Zach, M. Styner, C. Charles, and M. Niethammer, "Automatic bone segmentation and alignment from MR knee images," *SPIE Medical Imaging*, 2009, accepted for publication.