Automatic and Reliable Segmentation of Spinal Canals in Low-Resolution, Low-Contrast CT Images

Qian Wang, Le Lu, Dijia Wu, Noha El-Zehiry, Dinggang Shen, and Kevin S. Zhou

1 Introduction

To segment spinal canals is desirable in many studies because it facilitates analysis, diagnosis, and therapy planning related to spines. Segmentation of spinal canal provides helpful references to parcellate other anatomical structures and contributes to the understandings of full-body scans essentially [1]. Given spinal canal, it is much easier to delineate spinal cord, which is vulnerable to dosage tolerance and crucial for radiotherapy [2]. More previous works on spinal canal/cord focus on MR images, partly due to the better capability of MRI in rendering soft tissues. However, in this paper, we present an automatic method to segment spinal canals in low-resolution, low-contrast CT images. In particular, our highly diverse datasets are acquired from the CT channel in PET-CT and on pathological subjects. They are collected from eight different sites and vary significantly in Field-of-View (FOV), resolution, SNR, pathology, etc. Sagittal views of two typical datasets with different FOVs are shown

Qian Wang
Department of Computer Science, Department of Radiology and BRIC, University of North Carolina at Chapel Hill, e-mail: qianwang@cs.unc.edu

Le Lu
Radiology and Imaging Science, National Institutes of Health (NIH) Clinical Center, e-mail: le.lu@nih.gov

Dijia Wu
Siemens Corporate Research, e-mail: dijia.wu@siemens.con

Noha El-Zehiry
Siemens Corporate Research, e-mail: noha.el-zehiry@siemens.con

Dinggang Shen
Department of Radiology and BRIC, University of North Carolina at Chapel Hill, e-mail: dgshen@med.unc.edu

Kevin S. Zhou
Siemens Corporate Research, e-mail: shaohua.zhou@siemens.con

in Fig. 1(a) and (c), respectively. The coronal plane of another patient, whose spine twists due to diseases, is also provided in Fig. 1(b). High variation and limited quality of the datasets have incurred additional difficulty in segmenting spinal canals.

Most spinal canal segmentation methods in the literature are semi-automatic [3, 4, 5], which require manual initializations or interactions. Archip et al. [2] present a fully automatic pipeline by parsing objects in a recursive manner. Specifically, body con-tour and bones are extracted first. Then region growing is employed to segment spinal canal on each slice independently. When the boundary of spinal canal is relatively weak as shown in top-right of Fig. 1(c), this approach does not suffice and thus Snakes [6] is used to incorporate segmentation results from neighboring slices. Fol-lowing similar top-down parcellation strategy, [7] uses watershed and graph search to segment spinal canals. However, this top-down parcellation depends on locating the spine column first to provide rough but important spatial reference, which can be nevertheless non-trivial.

Interactive segmentation has also developed rapidly and drawn many successes in past decades. By allowing users to define initial seeds, the interactive mechanism is able to understand image content better and generate improved segmentation results in the end. We refer readers to [8] for a comprehensive survey of interactive segmentation methods. Among them, random walks (RW) [9] has been widely applied in various studies. RW asks users to specify seeding voxels of different labels, and then assigns labels to non-seeding voxels by embedding the image into a graph and utilizing intensity similarity between voxels. Users can edit the placement of seeds in order to acquire more satisfactory results.

In this paper, we adapt the idea of interactive segmentation to form a fully automatic approach that segments spinal canals from CT images. Different from manually editing seeds in the interactive mode, our method refines the topology of the spinal canal and improves segmentation in the automatic and iterative manner. To start the automatic pipeline, we identify voxels that are inside the spinal canal according to their appearance features [10]. For convenience, we will denote the voxels inside the spinal canal as foreground, and background otherwise. Then the detected seeds are input to RW and produce the segmentation of foreground/background. Based on the tentative segmentation, we extract and further refine the topology of the spinal canal by considering both geometry and appearance constraints. Seeds are adjusted accordingly and fed back to RW for better segmentation. By iteratively applying this scheme, we are able to cascade several RW solvers and build a highly reliable method to segment spinal canals from CT images, even under challenging conditions.

Our method and its bottom-up design, significantly different from the top-down parcellation in other solutions, utilize both population-based appearance information and subject-specific geometry model. With limited training subjects, we are able to locate enough seeding voxels to initialize segmentation and iteratively improve the results by learning spinal canal topologies that vary significantly across patients. We will detail our method in Section 2, and show experimental results in Section 3.

2 Method

We treat segmenting spinal canal as a binary segmentation problem. Let p_x denote the probability of the voxel x being foreground (inside spinal canal) after voxel classification and $\bar{p}_x$ for background, respectively. In general, we have $p_x + \bar{p}_x = 1$ after normalization. The binary segmentation can be acquired by applying a threshold to p_x. Although shapes of spinal canals can vary significantly across the population, they are tubular structures in general. We start from a small set of foreground voxels with very high classification confidences. These voxels act as positive seeds in RW to generate conservative segmentation with relatively low sensitivity but also low false positives (FP). All foreground voxels are assumed to form a continuous and smooth anatomic topology, which refines the seed points in order to better approximate the structure of the spinal canal. Hence the sensitivity of the RW segmentation increases with the new seeds. By iteratively feeding the improved seeds to RW, we have successfully formed an automatic pipeline that yields satisfactory segmentation of spinal canals.

2.1 Voxelwise Classification

In order to identify highly reliable foreground voxels as positive seeds, we turn to voxelwise classification via supervised learning. We have manually annotated the medial lines of spinal canals on 20 CT datasets. Voxels exactly along the medial lines are sampled as foreground, while background candidates are obtained from a constant distance away to the medial lines. We further use 3D Haar features as voxel descriptors. With varying sizes of detection windows, an abundant collection of Haar features is efficiently computed for each voxel. The probabilistic boosting tree (PBT) classifiers are then trained with AdaBoost nodes [11]. We have cascaded multiple PBT classifiers that work in coarse-to-fine resolutions. In this way, we not

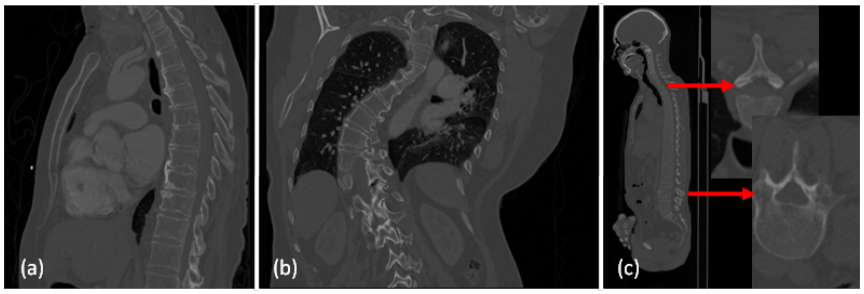

Fig. 1 Examples of datasets in our studies: (a) Sagittal view of restricted FOV near the chest area only; (b) Coronal view of disease-affected spine; (c) Sagittal view of full-body scan. Two additional transverse planes show that the spinal canal is not always contoured by bones.

only speed up the detection in early stage by reducing the number of samples, but also exploit features benefiting from higher scales of Haar wavelets in coarse resolution. Note that similar strategy is also successfully applied in other studies [10]. The well-performing foreground voxel confidence map (as well as the measuring color map) with respect to a training subject is displayed in Fig. 2(a). However, when applied to a new testing dataset (e.g., Fig. 2(c)-(d)), the classifiers may suffer from both false negative (FN) and FP errors. For instance, an FP artifact is highlighted in Fig. 2(b). Fig. 2(c) shows discontinuity of foreground confidence due to FN errors. Since the purpose here is to preserve highly reliable foreground voxels only (i.e., Fig. 2(d)), we have adopted a high confidence threshold (> 0.9) empirically to suppress most FP errors. The detection sensitivity will be subsequently improved as follows.

2.2 Random Walks

Similar to PBT-based classification, RW also produces voxelwise likelihoods of being foreground/background [9]. After users have specified foreground/background seeds, RW departs from a certain non-seeding voxel and calculates its probabilities to reach foreground and background seeds, as p_x and $\bar{p}_x$, respectively. Usually the non-seeding voxel x is assigned to foreground if $p_x > \bar{p}_x$. In the context of RW, the image is embedded into a graph where vertices correspond to individual voxels and edges link neighboring voxels. The weight w_{xy} of the edge e_{xy}, which measures the

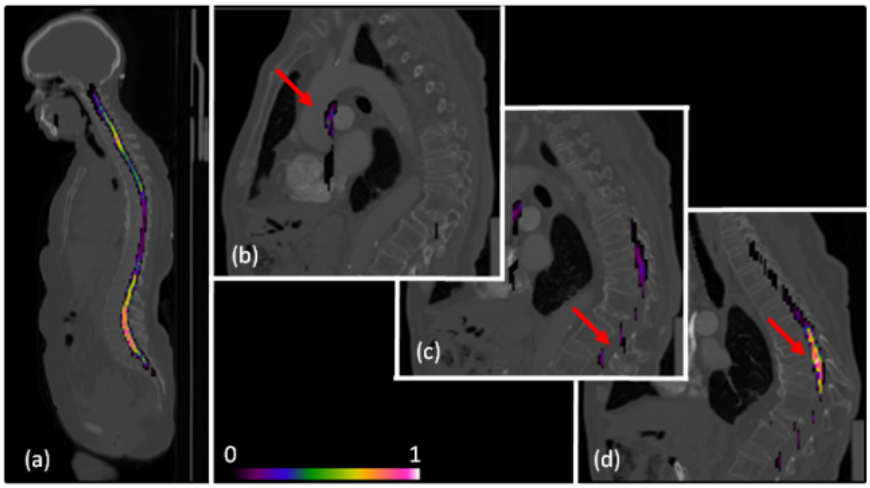

Fig. 2 Panel (a) shows the confidence map output by voxelwise classification on a *training* subject; panels (b)-(d) are for the voxelwise confidences of another *testing* dataset. Among them, FP errors and FN errors are highlighted in (b) and (c), respectively. We use a high confidence threshold to preserve reliable foreground voxels only as in (d).

similarity between two neighboring voxels x and y, is defined as

$$w_{xy} = \exp\left(-\beta(I_x - I_y)^2\right), \tag{1}$$

where I_x and I_y represent intensities at two locations; β a positive constant. Assuming segmentation boundaries to be coincident with intensity changes, RW aims to estimate p_x that satisfies to minimize the following energy term

$$E = \sum_{\forall e_{xy}} w_{xy}(p_x - p_y)^2. \tag{2}$$

To optimize the above is equivalent to solving a Dirichlet problem with boundary conditions defined by the seeds. Specifically, p_x is set to 1 if x is a foreground seed, and 0 for background. The calculated p_x incorporates spatial information of neighboring voxels, which differs from the independent voxelwise classification (Section 2.1).

The probability of each voxel in RW is associated with the paths from the voxel to seeds. Hence p_x is dependent not only on the weights of the edges forming the path, but also the length of each path. This potentially undermines RW that is sensitive to the seed locations. In the toy example of Fig. 3(a), there are three vertical stripes. The intensity of the middle stripe is slightly different and approximates the spinal canal surrounded by other tissues in CT data. We highlight certain sections of stripe boundaries in very high intensity to simulate the existence of vertebra, whose presence can be discontinuous as in Fig. 1(c). Foreground seeds and background seeds are colored in red and green, respectively. The calculated probability p_x in RW and the binary segmentation ($p_x > 0.5$) are shown. We observe from Fig. 3(c) that the segmented foreground falls into two segments undesirably. Though increasing the threshold on p_x and refining β to modify edge weights might improve the segmentation results, this becomes very ad-hoc. On the other hand, RW provides an interactive remedy by simply allowing users to place more seeds in proximity to the desired segmentation boundaries. The few additional seeds in Fig. 3(d) yield better discrepancy of foreground/background and lead to more satisfactory segmentation results (Fig. 3(e)-(f)).

2.3 Pipeline of Cascaded Random Walks

As mentioned above, we are able to identify seeds in voxel classification and feed them to RW for estimating segmentation. The initial segmented spinal canal usually breaks into several disconnected segments that imply high FN errors. This is because the initial seeds with high confidences from voxelwise classification are usually not sufficient to cover everywhere of the spinal canal.

Topology Refinement: To refine and acquire complete segmentation, we introduce the topology constraints of the spinal canal to segmentation. Specifically, we use the medial line of the spinal canal to represent its topology. After calculat-

ing all segments of the medial line given the tentative segmentation, we interleave them into a single connected curve. Fig. 4 illustrates the four sub-steps to refine the topology of the spinal canal with regard to its medial line. Based on tentative segmentation (including outputs from voxelwise classification), we calculate the medial point of foreground voxels on each transverse slice in Fig. 4(1). The medial point is defined to have the least sum of distances to all other foreground voxels on each slice. Assuming that the medial line connects all medial points, we then connect the medial points into several segments in Fig. 4(2). The medial line may break into several segments since medial points can be missing. Also certain medial point would be rejected as outlier if it incurs too high curvature to the medial line. With all computed segments, we interleave them by filling gaps with smooth virtual segments (as dotted curves) in Fig. 4(3). Each virtual segment $c(s)$ minimizes $\int \|\nabla^2 c(s)\|^2 ds$ to keep smooth as $s \in [0,1]$ for normalized arc-length. The stationary solution to the above holds when $\nabla^4 c(s) = 0$, and the Cauchy boundary conditions are defined by both two ends of the virtual segment as well as tangent directions at the ends. Though the numerical solution is non-unique, we apply the cubic Bzier curve for fast estimation of the virtual segment. For a certain virtual segment, we denote its ends as $\mathbf{P}_0$ and $\mathbf{P}_3$. An additional control point $\mathbf{P}_1$ is placed so that the direction from $\mathbf{P}_0$ to $\mathbf{P}_1$ is identical to the tangent direction at $\mathbf{P}_0$. Similarly, we can define $\mathbf{P}_2$ according to $\mathbf{P}_3$ and the associated tangent direction. We further require that the four control points are equally spaced. The virtual segment is $c(s) = (1-s)^3\mathbf{P}_0 + 3(1-s)^2 s\mathbf{P}_1 + 3(1-s)s^2\mathbf{P}_2 + s^3\mathbf{P}_3$.

After predicting the virtual segment in Fig. 4(3), we finally place more virtual medial points along the virtual segment. Besides the subject-specific geometry constraints to keep the virtual segments smooth, we further incorporate appearance criterion in Fig. 4(4). Upon all existing medial points (red dots), we calculate their intensity mean and the standard deviation (STD). The univariate Gaussian intensity model allows us to examine whether a new voxel is highly possible to be foreground given its simple appearance value. In particular, we start from both two ends of each virtual segment, and admit virtual medial point (purple dot) if its intensity is within

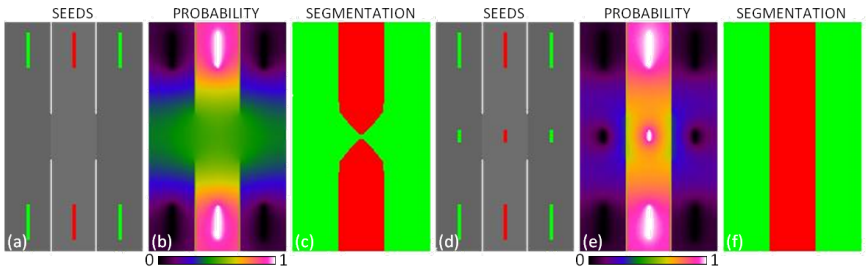

Fig. 3 With foreground seeds (in red) and background seeds (in green) (a), the calculated probabilities (b) and the corresponding binary segmentation (c) are not satisfactory. However, by manually placing more seeds (d), the segmentation results (e-f) are improved significantly.

the single STD range of the intensity model. The process to admit virtual medial points aborts when a disqualified candidate has been encountered.

Seed Sampling: After the topology of the spinal canal has been refined, we are able to provide better seeds for RW to use. All points along the refined medial line, including the newly admitted virtual ones, will act as foreground seeds. Moreover, we qualify more voxels as foreground seeds if (1) they have been classified as foreground in previous segmentation; (2) their intensities are within the single STD range of the appearance model introduced above; and (3) they are connected to the medial line via other foreground seeds. In this way, we have inherited previous segmentation in areas of high confidence, and saved computation since RW can simply treat them as boundary conditions. Surrounding voxels with high intensities will be regarded as bones and then counted as the background seeds.

Cascaded Random Walks: By repeating the procedures above, we have cascaded several RW solvers in order to generate the final segmentation. The pipeline will terminate automatically when the topology of the spinal canal, or the length of the medial line, has become stable. Remaining medial segments that are isolated from others will then be excluded from the foreground, in that they usually reflect artifacts especially from legs. During the iterative refinement, we also allow the medial line to grow at its both ends and thus admit more virtual medial points. The growth can stop automatically at the tail and terminate in the head by limiting the maximal radial size of the spinal canal. In Fig. 5(a), we show the foreground probability on four slices of a subject after the first iteration of our method. Improvement can be observed in Fig. 5(b) that shows the output after the second iteration. The final probability after the fourth iteration is shown in Fig. 5(c), with the corresponding segmentation in Fig. 5(d). The results above demonstrate that our method can efficiently utilize the topology of the spinal canal and generate satisfactory segmentation in the final.

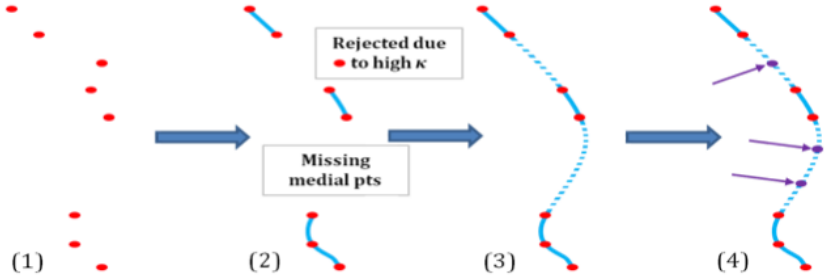

Fig. 4 Four sub-steps in refining the topology of the spinal canal include (1) estimating medial points; (2) determining medial segments; (3) calculating virtual medial segments (in dotted curves); and (4) placing more virtual medial points (in purple).

3 Experimental Result

We have conducted a study on 110 individual images from eight medical sites, as the largest scale reported in the literature, to verify the capability of our method. The training data are not used for the sake of testing. Even though the image quality of our data is low and the appearance variation is extremely high, we successfully generate good segmentation results on all datasets by our method with a fixed configuration. Sagittal views of segmentation results for 4 randomly selected subjects are shown in Fig. 6(a)-(d), respectively. In Fig. 6(e)-(g), we also show the segmentation result in 3 coronal slices for the extreme case in Fig. 1(b). This patient is under influences from severe diseases, which cause an unusual twist to the spine. However, though the topology of the spinal canal under consideration is abnormal, our method is still capable to well segment the whole structure. All results above confirm that our method is robust in dealing with the challenging data.

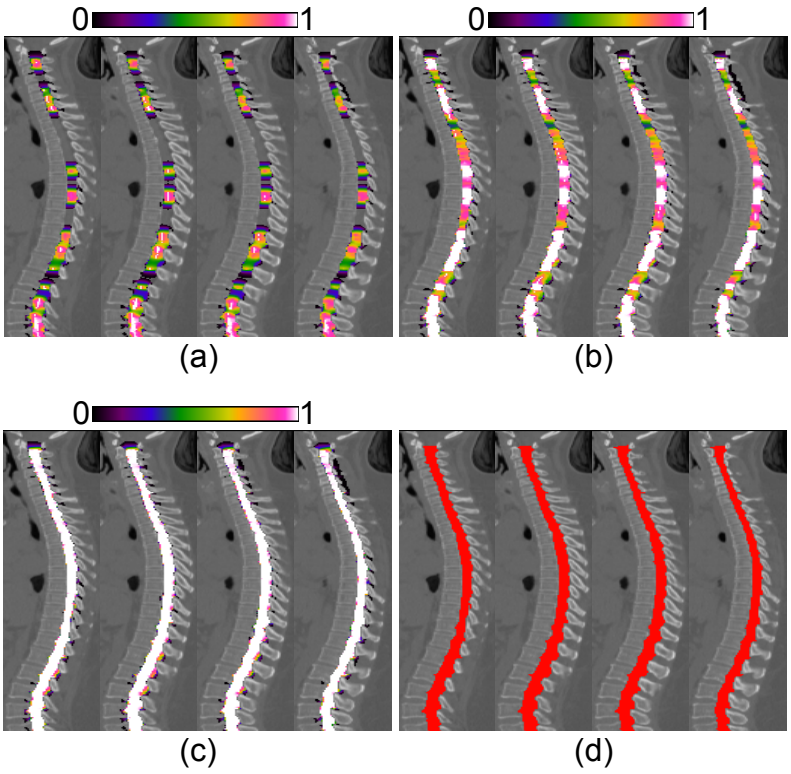

Fig. 5 Panels (a)-(c) show foreground probability on 4 consecutive slices of a certain subject output after the first, second, and the final (fourth) iteration, respectively. The binary segmentation in (d) is corresponding to the final probability in (c).

We have further manually annotated 20 datasets for quantitative evaluation. For the manually delineated parts, the Dice overlapping ratio between the segmentation of our method and the ground truth reaches 92.791.55%. By visual inspection, robust and good segmentation results are achieved on all 110 datasets, especially including many highly pathological cases. Note that to deal with image data with severe pathologies is not addressed and validated in the previous literature [2, 3, 4, 5, 7].

Our method achieves final segmentation in 2-5 iterations for all datasets, and typically costs 20-60 seconds per volume depending on the image size. With more sophisticated RW method that is better designed for editing seeds [12], the speed performance of our method can be further improved.

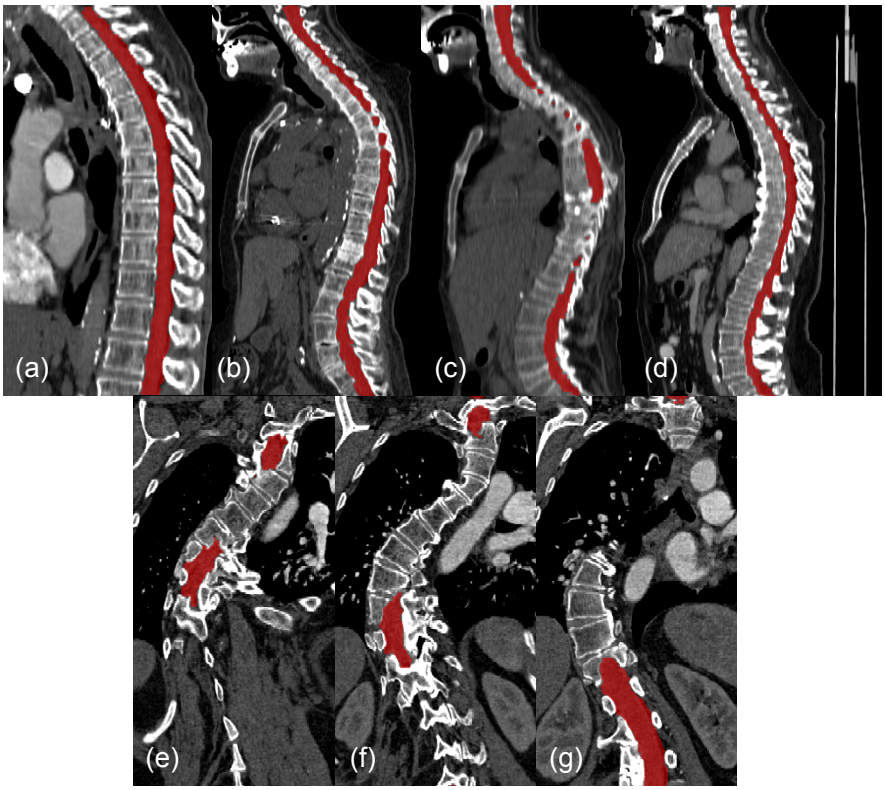

Fig. 6 Panels (a)-(d) show segmentation results (in sagittal views) on 4 randomly selected images. Coronal views of the segmentation results on an extreme case, whose spine twists due to diseases, are shown in (e)-(g).

4 Discussion

In this work, an automatic method to segment spinal canals from low-quality CT images is proposed. With initial seeds provided by PBT-based classification, we introduce topology constraints into segmentation via RW. Our iterative optimization has successfully enhanced the capability of RW in dealing with tubular spinal canals, in that the boundary conditions can be improved to guarantee better segmentation results. Our large-scale evaluation shows that the proposed method is highly accurate and robust even if the datasets are very diverse and challenging.

References

1. T. Klinder, J. Ostermann, M. Ehm, A. Franz, R. Kneser, C. Lorenz, Medical Image Analysis **13**(3), 471 (2009)
2. N. Archip, P.J. Erard, M. Egmont-Petersen, J.M. Haefliger, J.F. Germond, Medical Imaging, IEEE Transactions on **21**(12), 1504 (2002)
3. S.S.C. Burnett, G. Starkschall, C.W. Stevens, Z. Liao, Medical Physics **31**(2), 251 (2004)
4. G. Karangelis, S. Zimeras, in *Bildverarbeitung für die Medizin 2002*, ed. by M. Meiler, D. Saupe, F. Kruggel, H. Handels, T. Lehmann (Springer Berlin Heidelberg, 2002), pp. 370–373
5. L. Nyúl, J. Kanyó, E. Máté, G. Makay, E. Balogh, M. Fidrich, A. Kuba, in *Computer Analysis of Images and Patterns*, vol. 3691, ed. by A. Gagalowicz, W. Philips (Springer Berlin Heidelberg, 2005), vol. 3691, pp. 456–463
6. M. Kass, A. Witkin, D. Terzopoulos, International Journal of Computer Vision **1**(4), 321 (1988)
7. J. Yao, S. O'Connor, R. Summers, in *Biomedical Imaging: Nano to Macro, 2006. 3rd IEEE International Symposium on* (2006), pp. 390–393
8. D. Cremers, M. Rousson, R. Deriche, International Journal of Computer Vision **72**(2), 195 (2007)
9. L. Grady, Pattern Analysis and Machine Intelligence, IEEE Transactions on **28**(11), 1768 (2006)
10. D. Wu, D. Liu, Z. Puskas, C. Lu, A. Wimmer, C. Tietjen, G. Soza, S. Zhou, in *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on* (2012), pp. 980–987
11. Z. Tu, in *Computer Vision, 2005. ICCV 2005. Tenth IEEE International Conference on*, vol. 2 (2005), vol. 2, pp. 1589–1596 Vol. 2
12. S. Andrews, G. Hamarneh, A. Saad, in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2010*, vol. 6363, ed. by T. Jiang, N. Navab, J. Pluim, M. Viergever (Springer Berlin Heidelberg, 2010), vol. 6363, pp. 9–16