Preliminary version – final version appeared in
G. Fichtinger, A. Martel, T. Peters, editors, Medical Image Computing and Computer Assisted Intervention – MICCAI 2011, Lecture Notes in Computer Science, vol. 6893, 554–561, Springer, Berlin, 2011.

3-D Graph Cut Segmentation with Riemannian Metrics to Avoid the Shrinking Problem

Shouhei Hanaoka[12], Karl Fritscher[1], Martin Welk[1], Mitsutaka Nemoto[2], Yoshitaka Masutani[2], Naoto Hayashi[3], Kuni Ohtomo[2], Rainer Schubert[1],
[1] The Institute of Biomedical Image Analysis, the Health and Life Sciences University (UMIT), Eduard-Wallnöfer-Zentrum 1, Hall in Tirol, Austria
[2] Dept. of Radiology and [3] Dept. of Computational Diagnostic Radiology and Preventive Medicine, the Univ. of Tokyo Hospital, 7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan
hanaoka-tky@umin.ac.jp

Abstract. Though graph cut based segmentation is a widely-used technique, it is known that segmentation of a thin, elongated structure is challenging due to the "shrinking problem". On the other hand, many segmentation targets in medical image analysis have such thin structures. Therefore, the conventional graph cut method is not suitable to be applied to them. In this study, we developed a graph cut segmentation method with novel Riemannian metrics. The Riemannian metrics are determined from the given "initial contour," so that any level-set surface of the distance transformation of the contour has the same surface area in the Riemannian space. This will ensure that any shape similar to the initial contour will not be affected by the shrinking problem. The method was evaluated with clinical CT datasets and showed a fair result in segmenting vertebral bones.

Keywords: Graph cut, Segmentation, Riemannian geometry, Spine

1 Introduction

Graph cuts are one of the most widely used techniques for segmentation tasks in image analysis. The biggest advantage of the algorithm is that it can solve a segmentation problem as a global optimization problem without iterative calculation, and it guarantees a globally optimal solution [1]. The typical cost function minimized by the algorithm consists of 2 terms; a spatial coherency term and a data term, which are defined as follows:

$$E(f) = \sum_{p,q \in N} V_{p,q}(f_p . f_q) + \sum_{p \in P} D_p(f_p) \quad \textbf{(1)}$$

where D_p is the data energy, $V_{p,q}$ is the smoothness energy, N is the set of neighborhood pairs, f_p is the label assigned to the pixel p, and P are all pixels in the image [2]. Here, the 1st term (spatial coherency term) can be considered as a term which evaluates the length (in 2-D) or area (in 3-D) of the boundary, which is modulated with the contrast in the image. Therefore, minimizing the energy using this term causes a bias towards shorter boundaries. This behavior is known as

"shrinking bias" [3]. It causes severe problems especially when the target object has a long, thin spine-like process. Unfortunately, elongated structures are very common in anatomical objects. This fact significantly aggravates the segmentation of e.g. vertebrae in CT images and causes significant segmentation errors [4]. Though a number of methods have been reported to address the problem, such as [3] and [5], to the best of our knowledge, no simple way to avoid it has been presented.

In 2003, Boykov and Kolmogorov reported a method to construct a graph where cut metric approximates any given Riemannian metric, and utilized it for image segmentation [6]. The metrics were defined and calculated from the gradient information of the image to be segmented.

In this paper, we propose a novel method to avoid the shrinking problem by performing graph cuts in a Riemannian space. The Riemannian metrics is calculated not from the image itself but from a predefined shape template, or, "initial contour." Although the position and pose information of the target object is needed in advance, no other prior information like any "seed region" is required.

The basic idea is to compose a 3-D Riemannian space in which any surfaces parallel to the initial contour (an isosurface) has the same surface area. By performing graph cut segmentation in this 3-D Riemannian space, the inside and outside of the object can be considered and handled in the same way, so that the spatial coherency term serves as an evaluator of how the segmentation result differs from its closest isosurface. Therefore, a "shrinking problem" in the usual sense cannot occur.

2 Methods

The basic notion of this segmentation method is to perform graph cut-based segmentation in a Riemannian space which satisfies the following conditions:

1) It can be defined everywhere in the input image (excluding the points where the distance transformation of the initial contour, $dist(\mathbf{x})$, is not differentiable).
2) Any isosurface (level set surface) of $dist(\mathbf{x})$ — i.e. the set of points $\mathbf{x}$ with $dist(\mathbf{x}) = const$ — has the same surface area.

The "initial contour" approximates the object to be segmented (e.g. the mean shape of the target object). It is assumed that the given grayscale image and the initial contour is registered rigidly in advance. Then, the metrics of the Riemannian space are determined and calculated from the initial contour, so that any isosurface becomes parallel to each other and has the same surface area. Although such a Riemannian space has some singular points with indefinite metrics, our method can be performed in a stable manner even with them.

2.1. Calculation of the Riemannian Metrics

In order to apply the graph cut algorithm in Riemannian space, the Riemannian metrics at every grid point must be determined in advance. In this study, the metrics were calculated from the distance map of the initial contour $dist(\mathbf{x})$.

The metric tensor **G** at any point on the initial contour is defined to be isometric, thus, equal to an identity matrix. At any other point, the metrics are defined so that the sum of the area of any iso-surface (any level set of the distance map $dist(\mathbf{x}) = d$) will be identical. More detail on the calculation of the metrics will be given in Sections 2.1.1 and 2.1.2.

2.1.1. Metrics Calculation from Curvatures of the Distance Function. Fig. 1 illustrates the basic idea of the metrics calculation (in a 2-D space for explanatory usage.) Suppose that the metrics at the point $\mathbf{x}_d$ is to be calculated. Let the (outside-positive) signed distance of $\mathbf{x}_d$ from the initial contour be d.

$$dist(\mathbf{x}_d) = d \tag{2}$$

So, there must be a point $\mathbf{x}_0$, on the initial contour, whose distance from $\mathbf{x}_d$ equals d.

Let the initial surface (a curve for this 2-D example) be S_0, and the isosurface including $\mathbf{x}_d$ be S_d.

$$S_0 = \{\mathbf{x} \mid dist(\mathbf{x}) = 0\}, \quad S_d = \{\mathbf{x} \mid dist(\mathbf{x}) = d\} \tag{3}$$

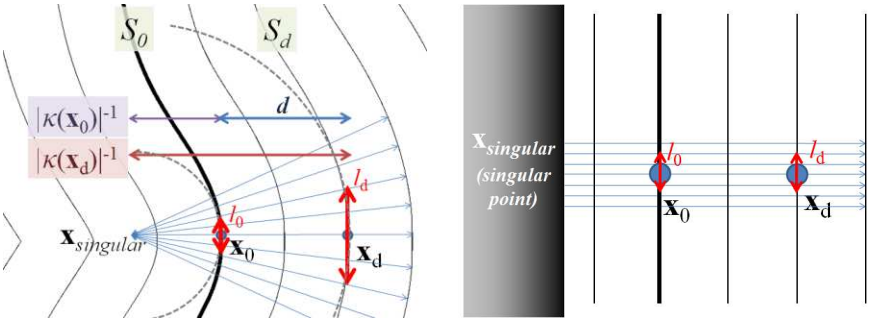

Fig. 1. Determination of the metrics. (left) The line segment l_0 and l_d are determined from the curvature radii, κ_0^{-1} and κ_d^{-1}, of the distance map. (right) Corresponding rectified presentation of the Riemannian space, in which the lengths of the two line segments are identical.

Let the curvature of the curve S_0 at $\mathbf{x}_0$ be κ_0, and that of S_d at $\mathbf{x}_d$ be κ_d (see also Fig. 1). Here, the curvatures are defined to be negative when the curve is convex. Under the assumption that both curvatures are negative (or both positive), there must be a point $\mathbf{x}_{singular}$ where the curvature becomes infinite. Note that both curves S_0 and S_d can be locally approximated as arcs whose center point is $\mathbf{x}_{singular}$, as shown in Fig. 1. Therefore, the curvature radii (the inverses of the curvatures) at $\mathbf{x}_0$ and $\mathbf{x}_d$ equal their distances from $\mathbf{x}_{singular}$. Thus, the curvatures satisfy the following equation:

$$-\kappa_0^{-1} + d = -\kappa_d^{-1} \tag{4}$$

Suppose that there are 2 minimal line segments, l_0 and l_d, which are parallel to the isosurfaces. Let further the Euclidean lengths of the two segments be l_0^E and l_d^E and the proportion of them be identical to the proportion of the curvature radii κ_0^{-1} and κ_d^{-1}. Using Formula (4), the proportion is calculated as:

$$\frac{{l_d}^E}{{l_0}^E} = \frac{{\kappa_d}^{-1}}{{\kappa_0}^{-1}} = \frac{1}{1+\kappa_d \cdot d} \tag{5}$$

In order to achieve the equalization of all iso-surface areas, the lengths of l_0 and l_d in a sense of Riemannian geometry, ${l_0}^R$ and ${l_d}^R$, must be identical (Fig. 1, right). According to the Riemannian geometry theory, the length of a minimal line segment l can be approximated as:

$$length(l) \cong \sqrt{\mathbf{v}^t \mathbf{G} \mathbf{v}} \tag{6}$$

where the length and direction of the vector **v** are those of the minimal line l, and a symmetric matrix **G** is the Riemannian metric tensor at that point. Furthermore, we have defined the metrics on the initial contour as Euclidean metrics, so that the metrics tensor on the initial contour, $\mathbf{G}_0$, is the identity matrix **I**. Thus, the length ${l_0}^R$ is equal to its Euclidean equivalent ${l_0}^E$. On the other hand, given the metric tensor $\mathbf{G}_d$ at the point $\mathbf{x}_d$, the length ${l_d}^R$ is calculated as:

$${l_d}^R \cong {l_d}^E \cdot \sqrt{\mathbf{w}^t \mathbf{G}_d \mathbf{w}} \tag{7}$$

where **w** is a unit vector with the same direction as l_d.

Consequently, the metric tensor $\mathbf{G}_d$ must satisfy the formula below:

$$\mathbf{w}^t \mathbf{G}_d \mathbf{w} = (1+\kappa_d \cdot d)^2 \tag{8}$$

To satisfy (8), we defined the tensor $\mathbf{G}_d$ as follows:

$$\mathbf{G}_d = \mathbf{R} \cdot \mathrm{diag}\left((1+\kappa_d \cdot d)^2, 1\right) \cdot \mathbf{R}^t \tag{9}$$

$$\mathbf{R} = (\mathbf{w} \quad \mathbf{n}), \quad |\mathbf{w}| = |\mathbf{n}| = 1, \quad \mathbf{w} \cdot \mathbf{n} = 0, \quad \mathbf{n} = \frac{\nabla dist(\mathbf{x})}{|\nabla dist(\mathbf{x})|}$$

Here, **R** is an orthogonal matrix whose column vectors **w** and **n** are unit vectors perpendicular and parallel to the gradient vector of the distance map, respectively.

2.1.2. Metrics Calculation for 3-D Volumes. In order to extend the metric calculation to a 3-D space, the metrics must be calculated from two principal curvatures of the given isosurface, instead of only one curvature in 2-D. When the 1st and 2nd curvatures are given as κ_1 and κ_2, and the corresponding principal directions are given as unit vectors w_1 and w_2, the metric tensor can be defined as:

$$\mathbf{G}_d = \mathbf{R} \cdot \mathrm{diag}\left((1+\kappa_1 \cdot d)^2, (1+\kappa_2 \cdot d)^2, 1\right) \cdot \mathbf{R}^t \tag{10}$$

$$\mathbf{R} = (\mathbf{w}_1 \quad \mathbf{w}_2 \quad \mathbf{n}), \quad \mathbf{n} = \frac{\nabla dist(\mathbf{x})}{|\nabla dist(\mathbf{x})|}$$

Using these metrics, the areas of all isosurfaces become equal. (The proof is omitted due to space limiting.)

2.2. Modification of the Edge Weights

In order to perform graph cut in the Riemannian space, both the data term and the spatial coherency term must be adequately modified. The former can be defined by the graph's edge weights between each image grid point and the *s*- (source) or *t*- (sink) node. The latter is defined by the edge weights between 2 adjacent grid points.

The theoretical framework to perform graph cut with Riemannian metrics was firstly presented by Boykov et al [6] based on integral geometry. However, their original approach is for finding the minimal surface in a Riemannian space, without considering any apparent equivalent of the data term (aside from hard constraints for seed regions). Therefore, we chose another, much simpler solution to modify the terms under the assumptions that:

1) The weights of edges between adjacent image grid points have to be proportional to the intersectional area of their border (in the Riemannian space).
2) The weights of edges to the s/t node have to be proportional to the volume of the space occupied by the grid point (again in the Riemannian space).

The first assumption is derived from the fact that these weights compose the spatial coherency term, which minimizes the surface area. The second assumption is determined in order to apply the same s/t weight to any unit volume.

In determining the weights, there is a difficulty due to singularity of the distance map. Our definition of metrics depends completely upon differentiability of the distance map. However, it is not differentiable at any point which has multiple nearest points on the initial contour. In our model, any singular point can be considered having infinite metrics, so that it has infinite volume (i.e., an infinite s/t-edge weight). Moreover, the finite difference approximation for the differentials is problematic near the singular points.

To avoid this problem, we restrict the area of interest to a band-like region $|dist(\mathbf{x})| < d_{\max}$. The constant d_{max} is determined as the maximal expected dislocation between the initial contour and the true contour to be segmented. The restriction is performed as follows:

1) Grid points close to any singular point, or *singular-including* points, are detected.
2) All *singular-including* grid points are removed from the graph, as well as any grid points with $|dist(\mathbf{x})| \geq d_{\max}$.
3) Grid points *adjacent to singular* (= adjacent to any *singular-including* point) have to be treated specifically: these grid points are considered to have finite 'depth' d_{max} towards the singular point (Fig. 2).

At each grid points, the data term is multiplied by the volume occupied by the point in the Riemannian space. Given the grid size is δ, and the metric tensor at the grid center point is $\mathbf{G}$, the volume occupied by the grid point can be approximated as

$$V = \delta^3 \cdot \sqrt{\det \mathbf{G}} \quad \textbf{(11)}$$

If the grid point is *adjacent to singular*, the occupied volume is considered as the volume within the proximal border of the grid and the border defined by the pre-defined depth d_{max}. It is simply approximated by multiplying the volume V by $\left(d_{\max} - |dist(\mathbf{x})|\right)$.

The non-s/t edge weights are modified by multiplying the cross-sectional area A. A is approximated as follows:

$$A = \delta^2 \cdot \sqrt{\left(\mathbf{u}_1^{\,t} \cdot \mathbf{G} \cdot \mathbf{u}_1\right) \cdot \left(\mathbf{u}_2^{\,t} \cdot \mathbf{G} \cdot \mathbf{u}_2\right) - \left(\mathbf{u}_1^{\,t} \cdot \mathbf{G} \cdot \mathbf{u}_2\right)^2} \tag{12}$$

where $\mathbf{u}_1$ and $\mathbf{u}_2$ are unit vectors perpendicular to the line segment connecting the two grid points and perpendicular to each other.

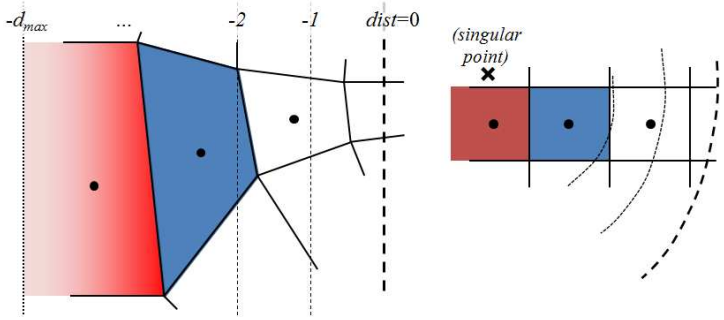

Fig. 2. Schematic views of image grids, (left) in a rectified presentation of the Riemannian space, and (right) in the original image space. The two painted boxes represent the areas occupied by two grid points. The left grid point is referred to as *adjacent to singular*.

2.3. Evaluation

The method was evaluated on 220 human vertebral bones in 10 datasets of clinical computed tomography (CT) images. The 1st and 2nd cervical vertebrae were excluded because of their unique shapes. The vertebral bones were divided into 4 groups: cervical, upper thoracic, lower thoracic and lumbar ones. In each group, the mean shape was calculated and used as initial contour. Before the proposed method was applied, each vertebra was cropped and rigidly registered in the same manner as described in [4]. Though 209 vertebrae were correctly identified and cropped by this full automatic process, 11 were failed and manually corrected in this study. Therefore, the following experiments were performed using these 220 pose-compensated, cropped volume images of vertebrae. The voxel size of 1×1×1 mm and $d_{\max}$=24 mm were used in this study.

The similarity index, Hausdorff distance and mean distance compared to the manual segmentation results were calculated for each vertebra. The results were compared to a graph cut segmentation without Riemannian metrics. Whenever possible, the parameters of the non-Riemannian version were set to the same values as the Riemannian graph cut version.

The data term used in this study (before modified with metrics) is a binary function whose value is -1 and 1 within and out of the initial contour, respectively. The spatial coherency term is a modified version of frequently used term introduced by Boykov et al. in [7]. In the modification only inward-positive image gradients are considered and outward-positive ones are regarded as 0 to detect outline of the bones.

3 Experimental Results and Discussions

The results of vertebral bone segmentation are shown in Table 1. An example result is shown in Fig. 3. The results of the proposed method turned out to be superior to the conventional graph cut approach in all criteria and in all 4 vertebral groups. The overall mean distance error (±s. d.) of the proposed method and the conventional method was 1.28 ±0.65 mm and 3.76 ±2.67 mm, respectively.

The result was comparable to another method reported by the authors [4] based on shape-intensity combined statistical models, in which the overall mean distance was 1.28 ±1.52 mm. It was also comparable to the study reported by Klinder et al. [8], in which the overall mean distance was 1.12 ±1.04 mm. On the other hand, the proposed method was less precise in thoracic vertebrae, mainly due to incorrect segmentation of the region adjacent to the ribs. For thoracic vertebrae, another study by Ma et al. [9] reported a better result with 0.95 ±0.91 mm of the mean distance.

One of limitations of the proposed method is that the initial contour must be given in advance. Therefore, it is possible that the method is less effective for targets with larger shape variation. On the contrary, it might be especially effective for objects with less variable but more complex shapes (e.g., with many thin or protruding parts), because most of known segmentation methods are not good at segmenting such complex objects reliably.

Table 1. The segmentation results. (mean ± s.d.)

Similarity index	Cervical	Upper th.	Lower th.	Lumbar
with proposed metrics	0.77±0.03	0.79±0.06	0.86±0.03	0.87±0.04
with Euclidean metrics	0.45±0.21	0.48±0.27	0.68±0.23	0.82±0.05

Hausdorff dist. (mm)	Cervical	Upper th.	Lower th.	Lumbar
with proposed metrics	9.03±2.15	17.10±7.71	11.13±5.61	11.80±4.26
with Euclidean metrics	17.62±4.16	22.08±6.84	18.05±7.34	16.93±5.07

Mean dist. (mm)	Cervical	Upper th.	Lower th.	Lumbar
with proposed metrics	1.11±0.24	1.75±0.90	1.11±0.51	1.11±0.40
with Euclidean metrics	4.16±2.00	4.90±3.1	3.53±2.89	2.14±1.01

4 Conclusion

An approach of graph cut segmentation based on a newly introduced Riemannian metrics was presented. The experimental result suggested its advantage in segmenting thin, spine-like structures in which conventional graph cut methods are affected by the "shrinking problem." Despite the simplicity of the approach, it achieves in some situations even comparable segmentation quality as more complex model-based methods. In future work, we aim at integrating the Riemannian graph cut and model-based approaches in order to develop a more powerful and accurate segmentation scheme.

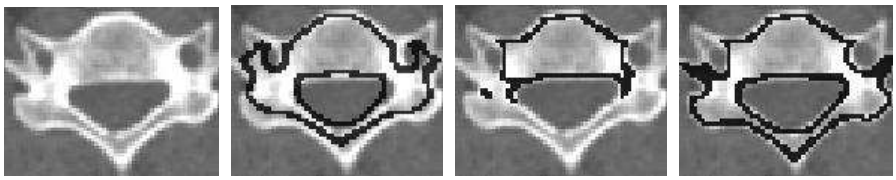

Fig. 3. An example of segmentation result for a cervical vertebra. (From left) The original grayscale image, rigidly-registered initial contour and two graph cut results with Euclidean and proposed metrics, respectively.

Acknowledgement

This study is a part of the research project "Computational Anatomy for Computer-aided Diagnosis and therapy: Frontiers of Medical Image Sciences", which is financially supported by the grant-in-aid for scientific research on innovative areas MEXT, Japan.

References

1. Peng, B., and Veksler, O.: Parameter selection for graph cut based image segmentation. In: Proc. of British Machine Vis. Conf. (2008)
2. Vineet, V., and Narayanan, P.: CudaCuts: Fast graph cuts on the GPU. In: CVPR'08 Workshop on Visual Computer Vision on GPUs (CVGPU'08), pp.1-8 (2008)
3. Vicente, S., Kolmogorov, V., Rother, C.: Graph cut based image segmentation with connectivity priors. In: CVPR. (2008)
4. Hanaoka, S., Fritscher, K. D., Schuler, B., Masutani, Y., Hayashi, N., Ohtomo, K., Schubert, R.: Whole vertebral bone segmentation method with a statistical intensity-shape model based approach. In: SPIE medical imaging (2011)
5. Kolmogorov, V. and Boykov, Y.: What metrics can be approximated by Geo-Cuts, or global optimization of length/area and flux. In: ICCV, volume I, pp.564-571 (2005)
6. Boykov, Y. and Kolmogorov, V.: Computing geodesics and minimal surfaces via graph cuts. In: ICCV, volume I, pp. 26–33 (2003)
7. Boykov, Y. and Jolly, M.-P.: Interactive graph cuts for optimal boundary and region segmentation of objects in N-D images. In: ICCV, Volume I, pp. 105–112 (2001)
8. Klinder T., Ostermann J., Ehm M., Franz, A., Kneser, R., Lorenz, C.: Automated model-based vertebra detection, identification, and segmentation in CT images. Medical Image Analysis 13(3), 471-482 (2009)
9. Ma, J., Lu, L., Zhan, Y., Zhou, X., Salganicoff, M., Krishnan, A.: Hierarchical Segmentation and Identification of Thoracic Vertebra Using Learning-Based Edge Detection and Coarse-to-Fine Deformable Model. In: Jiang, T., Navab, N., Pluim, J. P. W., Viregever, M. A. (eds.) MICCAI 2010, Part I. LNCS, vol. 6361, pp. 19–27. Springer, Heidelberg (2010)