Introducing Willmore Flow Into Level Set Segmentation of Spinal Vertebrae

Poay Hoon Lim*, Ulas Bagci, and Li Bai

***Abstract*—Segmentation of spinal vertebrae in 3-D space is a crucial step in the study of spinal related disease or disorders. However, the complexity of vertebrae shapes, with gaps in the cortical bone and boundaries, as well as noise, inhomogeneity, and incomplete information in images, has made spinal vertebrae segmentation a difficult task. In this paper, we introduce a new method for an accurate spinal vertebrae segmentation that is capable of dealing with noisy images with missing information. This is achieved by introducing an edge-mounted Willmore flow, as well as a prior shape kernel density estimator, to the level set segmentation framework. While the prior shape model provides much needed prior knowledge when information is missing from the image, and draws the level set function toward prior shapes, the edge-mounted Willmore flow helps to capture the local geometry and smoothes the evolving level set surface. Evaluation of the segmentation results with ground-truth validation demonstrates the effectiveness of the proposed approach: an overall accuracy of $89.32 \pm 1.70\%$ and 14.03 ± 1.40 mm are achieved based on the Dice similarity coefficient and Hausdorff distance, respectively, while the inter- and intraobserver variation agreements are $92.11 \pm 1.97\%$, $94.94 \pm 1.69\%$, 3.32 ± 0.46, and 3.80 ± 0.56 mm.**

***Index Terms*—Kernel density estimation (KDE), level set, vertebrae segmentation, Willmore flow.**

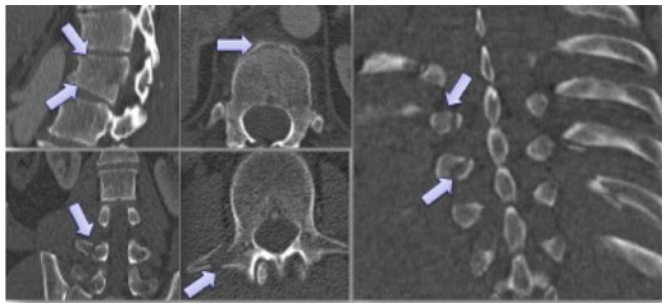

Fig. 1. Fractures of the vertebrae are indicated by arrows.

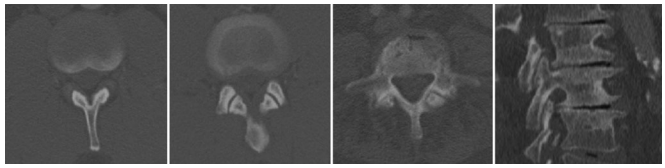

Fig. 2. Incomplete, missing or ambiguous information of the vertebrae in CT images.

I. Introduction

SPINE trauma is a devastating event with high morbidity and mortality, causing severe psychological, social, and financial burdens for patients, their families, and the society. Among the spine trauma incidents, motor vehicle accidents account for 42.1% of reported spinal cord injury cases, followed by falls (26.7%), acts of violence (15.1%) and sporting activities (7.6%) [1]. Compression fractures of the vertebral body in older adults are common, due to osteoporosis. Fractures of the thoracic and lumbar spine are more widespread than of the cervical spine, with highly affected area in thoracolumbar region (T11 to L4). Fractures and dislocations of spine may give rise to potentially devastating long-term consequences such as neural deficits, permanent disability, or even death. Fig. 1 shows examples of spinal disorders due to trauma and osteoporosis in computed tomography (CT) images. Identifying and grading severity of spine fractures and understanding their cause will help physicians determine the most effective pharmacological treatments and clinical management strategies for spinal disorders. One of the major challenges is to achieve *a firm diagnosis*. Detection and segmentation of spine from images are crucial steps in diagnostic imaging. Although these quantitative image analysis techniques have received increasing interest, accurate detection and segmentation methods are still lacking. Spine segmentation remains challenging due to the complexity of vertebrae shapes, gaps in the cortical bone and boundaries, as well as noise, inhomogeneity, and incomplete information in images, as shown in Fig. 2.

II. Previous Work

In recent years, CT imaging has replaced radiography as the primary modality for assessing bony structures in human anatomy, including the spine. Thus, this section will provide a summary of the existing methods on spinal vertebrae segmentation from CT images to highlight the need for accurate and automatic methods for 3-D vertebrae segmentation.

Manuscript received June 11, 2012; revised September 11, 2012; accepted October 15, 2012. Date of publication October 22, 2012; date of current version December 14, 2012. This work is supported in part by the University of Nottingham. *Asterisk indicates corresponding author.*

*P. H. Lim is with the School of Computer Science, University of Nottingham, Nottingham NG8 1BB, U.K. (e-mail: psxphl@nottingham.ac.uk; poayh11@yahoo.com).

U. Bagci is with the Radiology and Imaging Sciences, National Institutes of Health, Bethesda, MD 20892 USA (e-mail: ulas.bagci@nih.gov).

L. Bai is with the School of Computer Science, University of Nottingham, Nottingham NG8 1BB, U.K. (e-mail: Bai.Li@nottingham.ac.uk).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TBME.2012.2225833

Various attempts have been made at spine segmentation in recent years, but majority of them use 2-D images and/or require user intervention in the process. For example, Naegel [2] combined the watershed method and morphological approaches to segment vertebrae. Although the proposed method is promising in segmenting healthy bones from high-resolution images, manual refinement is necessary to obtain accurate segmentations, and the level of refinement is patient and resolution dependent. Ghebreab and Smeulders [3] constructed a deformable integral spine model to segment vertebrae. The method learns the appearance of vertebrae boundaries a priori from a set of training images. This model is then used to generate landmark points, in order to reduce the complexity of the segmentation process through point-based shape representation. However, it remains unclear if the landmark points correspond to the actual anatomical locations and whether they capture the biologically meaningful variations across different subjects. The method is also not fully automated and needs step-by-step inputs from the user, which makes the whole process tedious and time consuming. Ma *et al.* [4] presented an automatic vertebra segmentation and identification method on thoracic vertebra CT images. A learning-based bone structure edge detection algorithm was used and a hierarchical, coarse-to-fine deformable surface-based segmentation method was proposed based on the response maps from the learned edge detector. Though satisfactory results were obtained, the segmented vertebrae were only in 2-D and reproducibility of results in 3-D was not known.

Another limitation is the complexity and/or inaccuracy of current segmentation methods. For example, Lorenze and Krahnstoever [5] proposed a statistical shape model whereby the mean shape was constructed from a set of training samples. The initialization of the shape model for segmentation was done manually and is highly sensitive to dislocation. If the model is not located in the proximity of vertebrae, segmentation may fail. More recently Klinder *et al.* [6] used a mesh-based method to extract spine curves, and then generalized Hough transform and curved planar reformation to detect the vertebrae. The proposed approach has a further identification step to the detected vertebrae via rigid registration of appearance model. Although they achieved very competitive identification rates for vertebrae, their algorithm depends heavily on spatial registration of the model, which is computationally very expensive. In a paper by Mastmeyer *et al.* [7], a hierarchical 3-D technique was developed to segment the vertebral bodies in order to measure bone mineral density. The proposed framework needs excessive user intervention to precisely locate seed points to facilitate region growing segmentation. This process is time consuming and impractical for unhealthy bone segmentation. A similar approach integrating region growing segmentation with local shape and intensity refinement for delineating vertebrae was proposed by Kang *et al.* [8]. First, locally adaptive thresholds were used to facilitate region growing segmentations globally, followed by 3-D morphological operations to refine the segmented surfaces. This method still required a site specific separation of individual bones, which remains a challenge for vertebrae segmentation.

Due to the aforementioned drawbacks of the existing spinal vertebrae segmentation methods, we have developed a new method capable of segmenting spinal accurately from noisy images with missing information. The method is developed by introducing an edge-mounted Willmore flow, as well as a prior shape kernel density estimator, to the level set segmentation framework. While the prior shape model provides much needed prior knowledge when information is missing from the image, the edge-mounted Willmore flow helps to capture the local geometry and smoothes the evolving level set surface.

III. Contribution and Overview

We focus on 3-D segmentation of individual vertebrae with the aim to help physicians better visualize the shapes and internal structures of the vertebrae, as well as measure quantitatively the vertebrae volume and size of fracture for surgical procedures, e.g., spinal implant. The contribution of this study is twofold, clinically and technically: 1) providing a precise individual vertebrae segmentation system for clinical use; 2) introducing Willmore flow into the level set segmentation framework. To the best of our knowledge, this is the first work on 3-D segmentation of spinal vertebrae using Willmore flow within the level set method. We have previously successfully applied Willmore flow to 2-D vertebrae segmentation [9]. In the following sections, we shall describe in detail the segmentation method, followed by experimental results, ground-truth validation, comparison of the approach with the classical level set, region growing, and graph-cut approaches.

IV. Vertebra Segmentation

A. Level Set

The level set method [10] has been widely used for image segmentation [11]. For highly challenging segmentation tasks, such as segmenting occluded objects from medical images, level set methods have achieved good results when coupled with prior knowledge or prior shape models [12]–[15]. The level set method embeds an interface in a higher dimensional function ϕ (the signed distance function) as a level set $\phi = 0$. The equation that governs the evolution of the level set function $\phi(t)$ is $\frac{\partial \phi}{\partial t} + F|\nabla \phi| = 0$, where F represents the speed function.

In more recent development, the variational framework is often considered. Under the variational framework, an energy $E(\phi)$ is defined in relation to the speed function, and minimization of the energy generates the Euler–Lagrange equation and, hence, providing the evolution equation through the calculus of variation

$$\frac{\partial \phi}{\partial t} = -\frac{\partial E(\phi)}{\partial \phi}. \tag{1}$$

In this paper, we consider the fusion of energy, i.e., using a shape prior distribution estimator E_s with an edge-mounted Willmore energy E_{w_0}

$$E(\phi) = \lambda E_s + E_{w_0} \tag{2}$$

where $\lambda (0 < \lambda \leq 1)$ is the weight parameter. Details on E_s and E_{w_0} will be described in the following sections.

In order to incorporate a prior dataset $\{\phi_1, \phi_2, \ldots, \phi_N\}$ into the level set segmentation framework, we adopt a shape

dissimilarity measure based on the Kernel density estimation (KDE) discussed by Cremers *et al.* [13]. This nonparametric distribution estimator overcomes the two shortcomings of existing algorithms: 1) the assumption that the shapes are Gaussian distributed, which is generally inappropriate when the number of training set is small, and not practical for modeling shapes with high complexity and structure; 2) the shapes are represented by signed distance functions, which constitute a nonlinear space that does not include the mean.

B. Kernel Density Estimation

KDE is a nonparametric approach in statistics for estimating the probability density function of a random variable. The underlying theory of KDE states that data with unknown statistical distribution converge to its actual distribution as the number of samples approaches infinity. In practice, KDE provides a fundamental smoothing estimator even with a small number of data samples. In application with N samples of shape models, the density estimation is formulated as a sum of Gaussian of shape dissimilarity measures $d^2(H(\phi), H(\phi_i)), i = 1, 2, \ldots, N$

$$P(\phi) \propto \frac{1}{N} \sum_{i=1}^{N} e^{-\frac{d^2(H(\phi), H(\phi_i))}{2\sigma^2}} \tag{3}$$

where $H(\phi)$ is the *Heaviside* function, the shape dissimilarity measure [16]–[18] is

$$d^2(H(\phi), H(\phi_i)) = \int_\Omega \frac{1}{2} \left(H(\phi) - H(\phi_i)\right)^2 dx \tag{4}$$

and σ^2 is the mean squared nearest neighbor distance

$$\sigma^2 = \frac{1}{N} \sum_{i,j=1}^{N} \min_{j \neq i} d^2(H(\phi_i), H(\phi_j)). \tag{5}$$

Note that the L^2-norm is invariant under translation and scaling with respect to the principal axis of the shape. Hence, the shape dissimilarity measure d^2 is also invariant under these transformations when the prior shapes are normalized with respect to translation and scaling accordingly [13].

The segmentation is obtained by maximizing the conditional probability of ϕ given image I

$$P(\phi|I) = \frac{P(I|\phi)P(\phi)}{P(I)}. \tag{6}$$

Since the negative logarithmic scale of the probability distribution $P(\phi|I)$ nicely defines an energy that associates the evolution of ϕ with the minimization problem, the shape energy is formulated as

$$E_s(\phi) = -\log P(\phi|I). \tag{7}$$

Hence, the variational with respect to ϕ becomes

$$\begin{aligned} \frac{\partial E_s}{\partial \phi} &= \frac{\sum_{i=1}^{N} \alpha_i \frac{\partial}{\partial \phi} d^2(H(\phi), H(\phi_i))}{2\sigma^2 \sum_{i=1}^{N} \alpha_i} \\ &= \sum_{i=1}^{N} \frac{e^{-\frac{d^2(H(\phi), H(\phi_i))}{2\sigma^2}}}{2\sigma^2 \sum_{i=1}^{N} \alpha_i} \left(2\delta(\phi) \Big[H(\phi) - H(\phi_i(x - \mu_\phi)) \Big] \right. \\ &\quad + \int \Big[H(\phi(\xi)) - H(\phi_i(\xi - \mu_\phi)) \Big] \\ &\quad \left. \times \delta(\phi(\xi)) \frac{(x - \mu_\phi)^T \nabla \phi(\xi)}{\int H(\phi) dx} d\xi \right) \end{aligned} \tag{8}$$

where μ_ϕ is the centroid of $\{\phi > 0\}$ and $\alpha_i = \exp\left(-\frac{1}{2\sigma^2} d^2(H(\phi), H(\phi_i))\right)$ is the weight factor for $i = 1, 2, \ldots, N$.

C. Willmore Flow

Willmore energy is a function of mean curvature, which is a quantitative measure of how much a given surface deviates from a round sphere. It has been applied to image inpainting, restoration of implicit surfaces [19], [20], and to studies of the bending energy of biological cell membranes as these cell membranes tend to position themselves to minimize Willmore energy [21]. Willmore flow is the gradient flow of Willmore energy. Willmore flow of a surface is the evolution of the surface in time to follow variations of the Willmore energy. Willmore energy was defined after the British Geometer T. Willmore [22] and is formulated as

$$E_w = \frac{1}{2} \int_M h^2 dA \tag{9}$$

where M is a d-dimensional surface embedded in $\mathbb{R}^{d+1}$ and h the *mean curvature* on M.

In this paper, we integrate Willmore flow into the level set segmentation framework as a geometric functional. Willmore energy is defined on the collection of level sets, and Willmore flow is enabled by defining a suitable metric, the Frobenius norm, on the space of the level sets. The Frobenius norm of an arbitrary matrix $A = (a_{ij})_{m \times n}$, which is defined as $\|A\|_F = (\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2)^{1/2}$, coincides with the calculation for the gradient decent. It is equivalent to the l^2-norm (the Euclidean norm) of a matrix, i.e., $\|A\|_2 \leq \|A\|_F \leq \sqrt{r}\|A\|_2$, where r is the rank of A. More importantly, it is computationally attainable comparing to l^2-norm. As Frobenius norm is an inner-product norm, the optimization in the variational method comes naturally. Based on the formulation by Droske and Rumpf [23], Willmore flow or the variational form for the Willmore energy with respect to ϕ is

$$\frac{\partial E_w}{\partial \phi} = -\|\nabla \phi\| \left(\Delta_M h + h(t) \left(\|S(t)\|_2^2 - \frac{1}{2} h(t)^2 \right) \right) \tag{10}$$

where $\Delta_M h = \Delta h - h \frac{\partial h}{\partial n} - \frac{\partial^2 h}{\partial n^2}$ is the Laplacian Beltrami operator on h with $n = \frac{\nabla \phi}{\|\nabla \phi\|}, S = (I - n \otimes n)(\nabla \times \nabla)\phi$ is the shape operator on ϕ, and $\|S\|_2$ is the Frobenius norm of S.

In order to ensure that the smoothing effect of Willmore energy acts around the constructed surface and does not affect adversely the edge of vertebrae, we propose to multiply the edge indicator function $g(I) = \frac{1}{1 + |\nabla G_\sigma * I|^2}$ to the level set evolution

$$\frac{\partial E_{w_0}}{\partial \phi} = -g(I)\|\nabla \phi\| \left(\Delta_M h + h(t) \left(\|S(t)\|_2^2 - \frac{1}{2} h(t)^2 \right) \right) \tag{11}$$

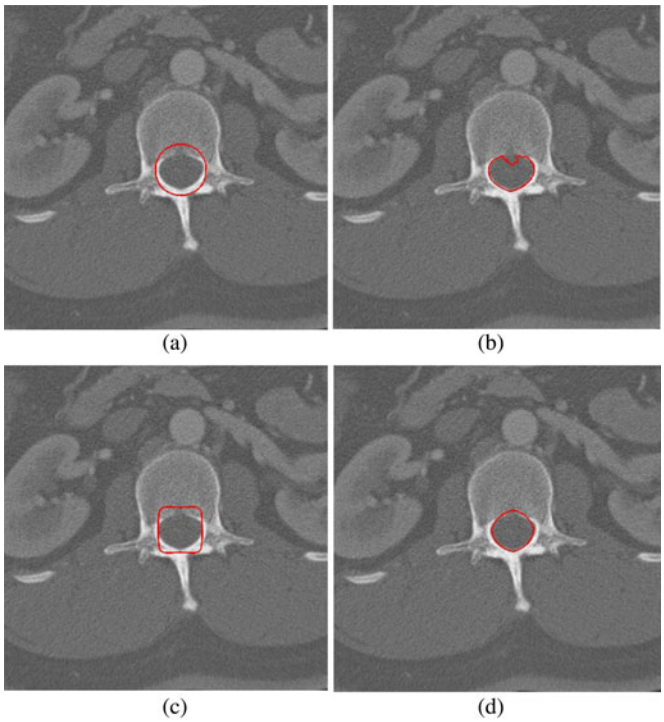

Fig. 3. (a) Initial contour. Segmented samples obtained by (b) Caselles, (c) Willmore, (d) edge-mounted Willmore approach.

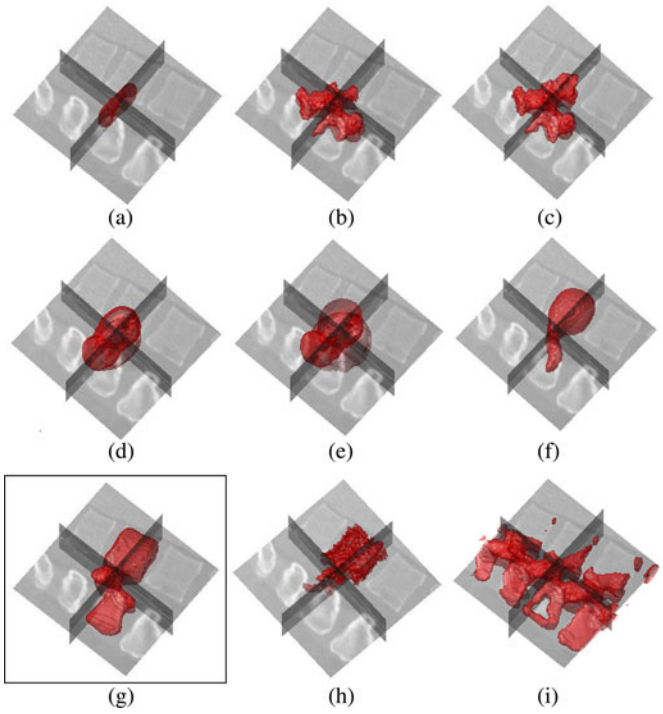

Fig. 4. (a) Initial contour for level set segmentation. Segmentation results obtained by (b) Chan–Vese,(c) Chan–Vese with shape prior, (d) Caselles, (e) Caselles with shape prior, (f) edge-mounted Willmore, (g) edge-mounted Willmore with shape prior energies, (h) region growing, and (i) graph cut, respectively.

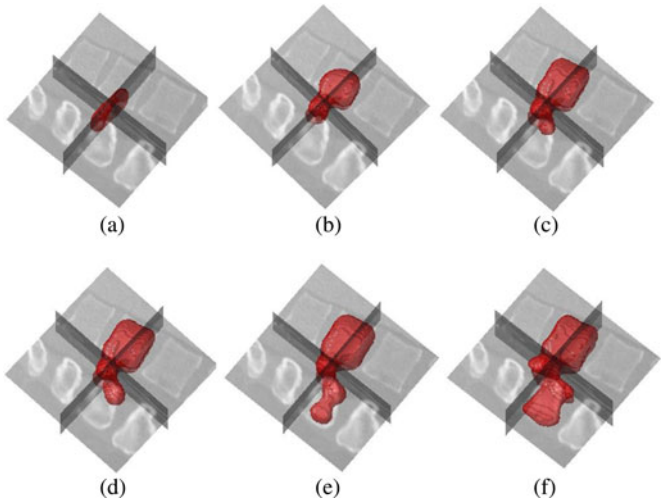

Fig. 5. (a)–(f): Various stages of level set evolution for a sample of lumbar vertebrae (L1) using the edge-mounted Willmore with shape prior energies.

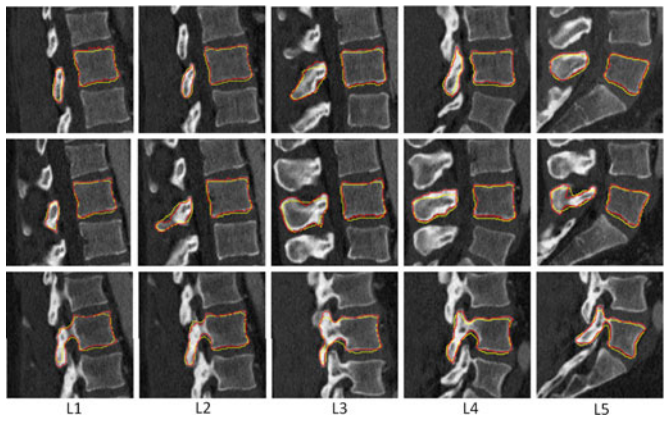

Fig. 6. Two-dimensional projections of 3-D segmentation results for lumbar vertebrae L1 to L5 shown on three selected slices (red) with ground truth (yellow).

where G_σ is the Gaussian function with standard deviation σ.

V. Experimental Results

A. Clinical Data and Ground-Truth Construction

The dataset consists of 20 CT images of normal spinal vertebrae images of patients aged 18 to 66. The patients are carefully selected by radiologists to form a representative group. These images are acquired from various CT scanners such as 32-detector row Siemens definition, 64-detector row Philips Brilliance, and 320-detector row Toshiba Aquilion. The in-plane resolutions for these images range from 0.88 to 1.14 mm, with consistent slice thickness of 2 mm. Original images have fixed sizes of 512×512, with slices varying from 45 to 98. The ground truths are obtained by manual delineations using TURTLESEG, an interactive 3-D image segmentation software [24] and verified by radiologists. Although 3-D manual delineation is time consuming, the benefits are twofold: 1) it allows validation of segmentation results; 2) it serves as shape models in the segmentation framework.

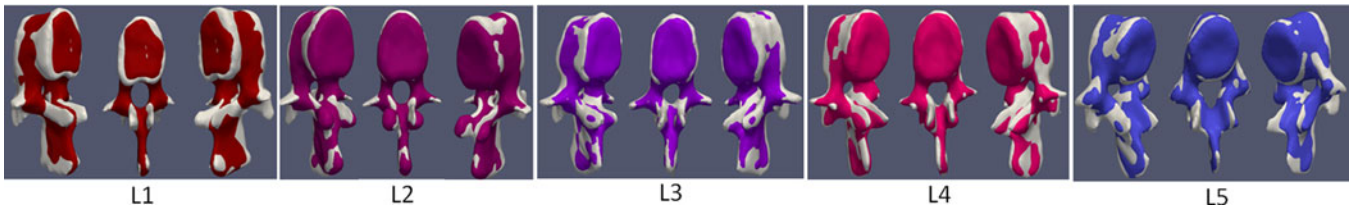

Fig. 7. Three-dimensional segmentation results (colors) of lumbar vertebrae L1 to L5 overlaid with ground truth (white).

With a torus as the initial contour set manually, the level set method is implemented around a narrow band [25] with reinitialization algorithm [26]. The classical Caselles [27] and Chan–Vese models [28] are combined with prior shape energies to obtain various segmentation results. The popular segmentation methods such as graph cut and region growing are used for comparison with the proposed method. The inter- and intraobserver variation estimation for manual delineations of ground truths are acquired to verify the difficulty of segmentation tasks in 3-D using the dataset. The leave-one-out approach is applied for cross validation and results are validated with the ground truth.

B. Segmentation Results and Comparisons

First, we compare the segmentation results obtained by the classical level set model driven by mean curvature flow as proposed by Caselles [27] and level set evolution using the edge-mounted Willmore energy as shown in Fig. 3. As can be seen in Fig. 3(b), the Caselles energy fails to fully segment the spinal canal since partial edge information is missing. It is clear from Fig. 3(c) and (d) that the Willmore energy maintains steady flow from initial contour with smooth level set evolution. However, this evolution is highly dependent on the initial contour, thus the segmentation must be accomplished with proper guidance to achieve desired results. For example, with the edge-mounted Willmore energy, segmentation clearly moves toward the spinal canal as shown in Fig. 3(d).

Fig. 4 demonstrates the 3-D segmentation results using various energy combinations in the level set methods, region growing and graph cut. The results show that Chan–Vese energy suffers from edge leaking to other vertebras, and the approaches fail to capture the full vertebrae shape when image inhomogeneity is present. The leaking problem is not resolved even when Chan–Vese energy is combined with shape prior energy, since the nearby vertebras are closely connected and share similar image intensities. Similarly, the leaking problem happens when the Caselles model is used due to incomplete edges in the image, and adding prior shape energy does not help the situation. The region growing algorithm fails to segment the vertebrae, and the graph-cut algorithm suffers badly from leaking. The edge-mounted Willmore energy is able to overcome problems such as incomplete edges and leaking, with added advantage of smoothing the segmentation. However, it can be halted by noisy patches before reaching the edges. Thus, a prior shape energy is incorporated to resolve the problem.

Fig. 5 illustrates the 3-D level set evolution and segmentation of a vertebra sample using the proposed method. Overall results of the proposed method are demonstrated in Figs. 6, 7, and Fig. 8, respectively. Fig. 6 demonstrates three selected 2-D projections of 3-D segmentation results, while Figs. 7 and 8 show the 3-D segmentation results overlaid with the ground truth for vertebrae L1 to L5, respectively. The vertebrae shape is completely captured by the proposed framework.

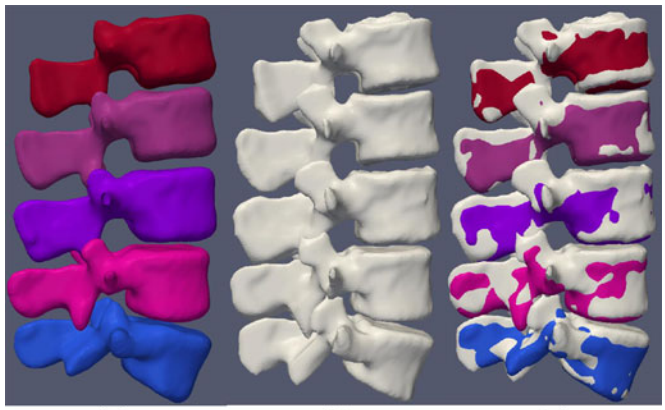

Fig. 8. (a) Three-dimensional segmentation results (colors), (b) ground truth (white), and (c) results overlaid with ground truth for lumbar vertebrae L1 to L5.

TABLE I
AVERAGE MEASUREMENT (MM) WITH STANDARD DEVIATION FOR UPPER VERTEBRAL WIDTH AND DEPTH (UVW, UVD), LOWER VERTEBRAL WIDTH AND DEPTH (LVW, LVD), VERTEBRAL BODY HEIGHT IN POSTERIOR AND ANTERIOR (VBHP, VBHA) OF LUMBAR VERTEBRAE L3 TO L5

Vertebra	L3	L4	L5
UVW	43.2±4.3	48.5±4.7	52.2±5.1
UVD	32.3±3.3	34.6±3.6	35.7±3.7
LVW	51.7±4.8	52.5±4.7	53.1±6.0
LVD	35.3±3.6	36.2±3.7	36.0±4.0
VBHp	29.6±2.4	28.7±2.3	25.9±2.0
VBHa	30.2±2.1	30.1±2.4	30.8±2.5

C. Quantitative Evaluation of Segmentation

The Dice similarity coefficient (DSC) [29] and Hausdorff distance (HD) [30] are used to measure the intrinsic index of segmentation performance. While DSC measures the relative superposition between two enclosed volumes, HD measures the relative differences between boundaries of the segmented objects. The DSC is formulated as

$$D(\Omega_S, \Omega_G) = \frac{2(|\Omega_S \bigcap \Omega_G|)}{|\Omega_S| + |\Omega_G|} \tag{12}$$

TABLE II
AVERAGE DSC (%) AND HD (MM) WITH STANDARD DEVIATION FOR SEGMENTATIONS OF LUMBAR VERTEBRAE L1 TO L5 USING CHAN–VESE (CV), CHAN–VESE WITH PRIOR SHAPE (CV + S), CASELLES (CA), CASELLES WITH PRIOR SHAPE (CA + S), EDGE-MOUNTED WILLMORE (W_0), EDGE-MOUNTED WILLMORE WITH PRIOR SHAPE (W_0 + S) ENERGIES, REGION GROWING (RG), AND GRAPH-CUT (GC) APPROACH

Vertebra	L1		L2		L3		L4		L5	
	DSC(%)	HD(mm)	DSC(%)	HD(mm)	DSC(%)	HD(mm)	DSC(%)	HD(mm)	DSC(%)	HD(mm)
CV	42.19±8.32	23.10±2.31	37.03±7.80	25.37±2.52	36.79±6.32	27.14±2.52	36.22±6.11	28.02±1.81	36.17±6.8	29.78±1.71
CV+S	50.78±8.52	21.49±2.38	44.97±7.81	23.89±2.47	45.23±6.87	25.61±2.71	42.78±6.99	26.72±2.52	41.71±7.50	28.82±1.81
Ca	53.52±8.28	21.05±1.21	53.40±9.47	21.04±0.89	56.39±8.73	21.72±1.02	56.88±7.43	22.54±1.55	58.55±6.80	24.73±3.17
Ca+S	69.46±3.27	17.49±0.61	69.71±2.96	17.68±0.65	70.98±2.89	18.17±0.93	72.10±2.25	18.28±1.10	73.33±2.24	20.31±2.44
W_0	73.35±2.91	17.77±1.22	74.79±2.54	18.05±1.75	75.77±3.22	19.90±1.35	76.33±2.25	20.29±1.79	78.86±3.11	20.29±1.41
W_0+S	**89.11±2.24**	**12.68±1.03**	**89.62±1.27**	**12.96±1.77**	**89.47±1.32**	**14.57±1.89**	**88.98±2.18**	**14.81±1.31**	**89.43±2.18**	**15.12±0.99**
RG	42.72±11.22	23.60±1.31	42.27±11.51	24.13±2.90	43.33±9.97	24.50±1.50	42.21±10.84	24.82±2.34	40.97±13.62	28.95±3.44
GC	11.51±11.73	63.91±15.92	12.54±12.93	64.74±16.06	14.56±12.97	61.07±16.81	17.58±14.33	57.90±19.18	9.94±3.60	68.49±6.12

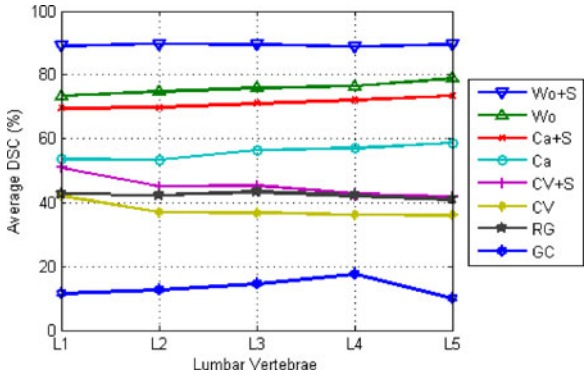

Fig. 9. Average DSC (%) for segmentations of lumbar vertebrae L1 to L5 using Chan–Vese, Chan–Vese with prior shape, Caselles, Caselles with prior shape, edge-mounted Willmore, edge-mounted Willmore with prior shape energies, region growing, and graph cut.

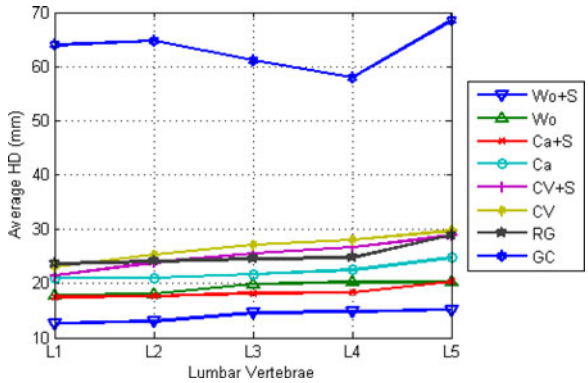

Fig. 10. Average HD (mm) for segmentation of lumbar vertebrae L1 to L5 using Chan–Vese, Chan–Vese with prior shape, Caselles, Caselles with prior shape, edge-mounted Willmore, edge-mounted Willmore with prior shape energies, region growing, and graph cut.

where $|\Omega_S|$ and $|\Omega_G|$ represent the volumes of segmented object Ω_S and the ground-truth Ω_G, respectively. The measurement (varies from 0 to 1) indicates the correspondence between two volumes, i.e., 0 means the two volumes do not overlap and 1 shows they are perfectly matched.

On the other hand, the HD is defined as

$$d_H(X,Y) = \max\{\sup_{x\in X}\inf_{y\in Y} d(x,y), \sup_{y\in Y}\inf_{x\in X} d(x,y)\} \quad (13)$$

where X and Y are the boundaries of two different segmented volumes, respectively. It measures the distance between the farthest point of a set to the nearest point of the other and vice versa. The measurement (varies from 0 to ∞ theoretically) signifies the difference between two closed surfaces, e.g., 0 indicates that both volumes share exactly the same boundaries, and larger HD values mean larger distances between the boundaries. In summary, a high DSC and a low HD are desirable for good segmentation.

In this study, the DSC are converted into 100 percentile for general physical interpretation. For more meaningful comparison using HD, we summarize the geometrical dimensions of lumbar vertebrae provided by Zhou *et al.* [31] in Table I.

The accuracy of segmentation results in comparison with the ground truths is indicated in Table II, Figs. 9, and 10. In addition, the box-plot diagram of DSC and HD distribution for vertebrae segmentation results using the proposed method is demonstrated in Fig. 11. As can be seen, even

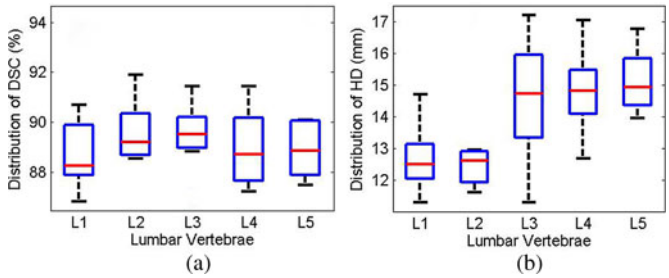

Fig. 11. Box-plot diagram of (a) DSC (%) and (b) HD (mm) distribution for segmentation results of lumbar vertebrae L1 to L5 using the edge-mounted Willmore with shape prior approach.

without integrating prior shape into the segmentation framework, the edge-mounted Willmore flow leads to superior segmentation results (DSC $75.82 \pm 2.81\%$, HD 19.26 ± 1.51 mm) over other methods, e.g., Caselles with prior shape models (DSC $71.12 \pm 2.72\%$, HD 18.39 ± 1.15 mm), Chan–Vese with prior shape models ($45.09 \pm 7.54\%$, HD 25.31 ± 2.38 mm), region growing (DSC $42.30 \pm 11.43\%$, HD 25.20 ± 2.30 mm), graph cut (DSC $13.23 \pm 11.11\%$, HD 63.37 ± 14.42 mm), as well as Caselles and Chan–Vese models without prior shape models. By integrating both edge-mounted Willmore flow and prior shape models, the proposed approach has achieved an overall accuracy of $89.32 \pm 1.70\%$ (DSC) and 14.03 ± 1.40 mm (HD), clearly outperforming the other methods.

The level of difficulty of the segmentation tasks on the dataset is determined by inter- and intraobserver agreements. For an interobserver agreement, manual delineation on the same dataset is performed by two independent observers. Whereas for an intraobserver agreement, two manual delineations are performed by an observer, with the second annotation done one week after the first. Based on the DSC and HD evaluation schemes, the inter- and intraobserver agreements for the dataset are $92.11 \pm 1.97\%, 94.94 \pm 1.69\%$ (DSC), 3.32 ± 0.46 mm, and 3.80 ± 0.56 mm (HD), respectively.

D. Computational Complexity Analysis

The experimental results in Section V show that the performance of our algorithm is superior in terms of accuracy over others such as region-growing and graph-cut algorithms. This improvement in accuracy, however, introduces an increased computational cost. The complexity of our algorithm is $O(N^2)$, where N is the image size. This is comparable to that of the algorithm proposed by Cremer *et al.* [13]. The computational cost is related to the number of operators involved in the algorithm. As the gradient descent evolution algorithm is governed by a combination of regional difference, dissimilarity measure and Willmore energy, the number of operators increases accordingly. The main advantage of our algorithm is its accuracy segmenting highly complex anatomical objects.

VI. Conclusion and Future Work

We have presented a new segmentation framework based on the level set for spinal vertebrae segmentation. The framework incorporates not only prior shape knowledge through the KDE method, but also local geometrical features through Willmore flow, into the level set segmentation. Experimental results on CT images, validated with ground truth, demonstrate the effectiveness of proposed framework. The framework has achieved much higher segmentation accuracy than existing methods, such as Chan–Vese and Caselles, region growing, and graph cut. Ongoing research includes integrating the segmentation framework into a system for detection and quantification of vertebrae fractures and other diseases of the spine.

Acknowledgment

The authors would like to thank Prof. J. K. Udupa of Department of Radiology, University of Pennsylvania for the discussion on a related project using MR images. Special thanks to Dr. O. Aras of NCI for his rigorous discussions on the medical problem and Dr. J. Yao of NIH for his useful comment.

References

[1] S. Looby and A. Flanders. (2011, Jan.). Spine trauma. *Radiol. Clin. North. Am.* [Online]. *49(1)*, pp. 129–163, Available: http://linkinghub.elsevier.com/retrieve/pii/S0033838910001454

[2] B. Naegel, "Using mathematical morphology for the anatomical labeling of vertebrae from 3-D CT-scan images," *Comput. Med. Imaging Grap.*, vol. 31, no. 3, pp. 141–156, 2007.

[3] S. Ghebreab and A. Smeulders. (2004, Oct.). Combining strings and necklaces for interactive three-dimensional segmentation of spinal images using an integral deformable spine model. *IEEE Trans. Biomed. Eng.* [Online]. *51(10)*, pp. 1821–1829, Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1337150

[4] J. Ma, L. Lu, Y. Zhan, X. Zhou, M. Salganicoff, and A. Krishnan, "Hierarchical segmentation and identification of thoracic vertebra using learning-based edge detection and coarse-to-fine deformable model," in *Proc. Int. Conf. Med. Imag. Comput. Comput. Aided Intervention*, 2010, pp. 19–27.

[5] C. Lorenz and N. Krahnstoever, "3D statistical shape models for medical image segmentation," in *Proc. 2nd Int. Conf. 3-D Dig. Imag. Model.*, 1999, pp. 4–8.

[6] T. Klinder, J. Ostermann, M. Ehm, A. Franz, R. Kneser, and C. Lorenz. (2009, Jun.). Automated model-based vertebra detection, identification, and segmentation in ct images. *Med. Imag. Anal.* [Online]. *13(3)*, pp. 471–482, Available: http://linkinghub.elsevier.com/retrieve/pii/S1361841509000085

[7] A. Mastmeyer, K. Engelke, C. Fuchs, and W. A. Kalender, "A hierarchical 3-d segmentation method and the definition of vertebral body coordinate systems for qct of the lumbar spine," *Med. Image Anal.*, vol. 10, pp. 560–577, 2006.

[8] Y. Kang, K. Engelke, and W. A. Kalender, "A new accurate and precise 3d segmentation method for skeletal structures in volumetric ct data," *IEEE Trans. Med. Imag.*, vol. 22, no. 5, pp. 586–598, May 2003.

[9] P. H. Lim, U. Bagci, O. Aras, Y. Wang, and L. Bai, "A novel spinal vertebrae segmentation framework combining geometric flow and shape prior with level set method," in *Proc. IEEE Int. Symp. Biomed. Imag.*, Barcelona, Spain, May 2012, pp. 1703–1706.

[10] S. Osher and J. Sethian, "Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations," *J. Comput. Phys.*, vol. 79, pp. 12–49, 1988.

[11] D. Feltell and L. Bai, "Level set image segmentation refined by intelligent agent swarms," presented at the World Congr. Computational Intelligence, Barcelona, Spain, 2010.

[12] A. Tsai, A. Yezzy, W. Wells, C. Tempany, D. Tucker, A. Fan, W. E. Grimson, and A. Willsky, "A shape-based approach to the segmentation of medical imagery using level sets," *IEEE Trans. Med. Imag.*, vol. 22, no. 2, pp. 137–154, Feb. 2003.

[13] D. Cremers, S. J. Osher, and S. Soatto, "Kernel density estimation and intrinsic alignment for shape priors in level set segmentation," *Int. J. Comput. Vis.*, vol. 69, no. 3, pp. 335–351, 2006.

[14] M. Rousson and N. Paragios. (2008, Mar.). Prior knowledge, level set representations & visual grouping. *Int. J. Comput. Vis.* [Online]. *76(3)*, pp. 231–243, Available: http://www.springerlink.com/index/10.1007/s11263-007-0054-z

[15] P. H. Lim, U. Bagci, and L. Bai, "A new prior shape model for level set segmentation," in *Proc. Iberoamer. Congr. Conf. Progress Pattern Recog., Image Anal., Comput. Vis., and Appl.*, 2011, ch. 14, pp. 125–132.

[16] T. Chan and W. Zhu, "Level set based shape prior segmentation," Comput. Appl. Math., Univ. California, Los Angeles, Tech. Rep. 03–66, 2003.

[17] T. Riklin-Raviv, N. Kiryati, and N. Sochen, "Unlevel sets: Geometry and prior-based segmentation," in *Proc. Eur. Conf. Comput. Vis.*, 2004, pp. 50–61.

[18] G. Charpiat, O. Faugeras, and R. Keriven, "Approximations shape metrics and application to shape warping and empirical shape statistics," *Found. Comput. Math.*, vol. 5, no. 1, pp. 1–58, Feb. 2005.

[19] R. Schneider and L. Kobbelt, "Generating fair meshes with G1 boundary conditions," in *Proc. Geom. Model. Process. Conf.*, 2000, pp. 251–261.

[20] S. Yoshizawa and A. G. Belyaev, "Fair triangle mesh generation with discrete elastica," in *Proc. Geom. Model. Process.*, 2002, pp. 119–123.

[21] J. W. Barrett, H. Garcke, and R. Nrnberg. (2008, Jan.). Parametric approximation of willmore flow and related geometric evolution equations. *SIAM J. Sci. Comput.* [Online]. *31(1)*, pp. 225–253, Available: http://epubs.siam.org/doi/abs/10.1137/070700231

[22] T. J. Willmore, "Note on embedded surfaces," *Analele Ştiinţifice ale Universităţii Al. I. Cuza din Iaşi. Serie Nouă*, vol. Ia 11B, pp. 493–496, 1965.

[23] M. Droske and M. Rumpf. (2004). A level set formulation for Willmore flow. *Interfac. Free Boundar.* [Online]. *6(3)*, pp. 361–378, Available: http://www.ems-ph.org/doi/10.4171/IFB/105

[24] A. Top, G. Hamarneh, and R. Abugharbieh, "Spotlight: Automated confidence-based user guidance for increasing efficiency in interactive 3D image segmentation," in *Proc. Med. Imag. Comput. Comput.-Assist. Interv. Workshop Med. Comput. Vis.*, 2010, pp. 204–213.

[25] D. Adalsteinsson and J. A. Sethian, "A fast level set method for propagating interfaces," *J. Comput. Phys.*, vol. 118, pp. 269–277, 1995.

[26] M. Sussman, P. Smereka, and S. Osher, "A level set approach for computing solutions to incompressible 2-phase flow," *J. Comput. Phys.*, vol. 114, no. 1, pp. 146–159, 1994.

[27] V. Caselles, R. Kimmel, and G. Sapiro. (1997). Geodesic active contours. *Int. J. Comput. Vis.* [Online]. *22(1)*, pp. 61–79, Available: http://www.springerlink.com/openurl.asp?id=doi:10.1023/A:1007979827043

[28] T. Chan and L. Vese, "Active contour without edges," *IEEE Trans. Imag. Process.*, vol. 10, no. 2, pp. 266–277, Feb. 2001.

[29] L. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, pp. 297–302, 1945.

[30] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*. New York: Springer-Verlag, 2005.

[31] S. H. Zhou, I. D. McCarthy, A. H. McGregor, R. R. H. Coombs, and S. P. F. Hughes. (2000, Jun.). Geometrical dimensions of the lower lumbar vertebrae—analysis of data from digitised CT images. *Eur. Spine J.* [Online]. *9(3)*, pp. 242–248, Available: http://www.springerlink.com/openurl.asp?genre=article&id=doi:10.1007/s0 05860000140

Authors' photographs and biographies not available at the time of publication.