

CT Prostate Deformable Segmentation by Boundary Regression

Yeqin Shao^{1,2}, Yaozong Gao^{2,3}, Xin Yang¹, and Dinggang Shen²(✉)

¹ Institute of Image Processing and Pattern Recognition,
Shanghai Jiao Tong University, Shanghai 200240, China

² Department of Radiology and BRIC, University of North Carolina,
Chapel Hill, NC 27599, USA
dgshen@med.unc.edu

³ Department of Computer Science, University of North Carolina,
Chapel Hill, NC 27599, USA

Abstract. Automatic and accurate prostate segmentation from CT images is challenging due to low image contrast, uncertain organ motion, and variable organ appearance in different patient images. To deal with these challenges, we propose a new prostate boundary detection method with a boundary regression strategy for prostate deformable segmentation. Different from the previous regression-based segmentation methods, which train one regression forest for each specific point (e.g., each point on a shape model), our method learns a single global regression forest to predict the nearest boundary points from each voxel for enhancing the entire prostate boundary. The experimental results show that our proposed boundary regression method outperforms the conventional prostate classification method. Compared with other state-of-the-art methods, our method also shows a competitive performance.

1 Introduction

Prostate cancer is the second leading cause of male cancer death in USA [1]. As one of the major treatments to the prostate cancer, image-guided radiation treatment (IGRT) aims to deliver a high dose of X-ray to tumors, while limiting the dose exposed to the surrounding healthy organs. Inaccurate localization of the prostate could result in wrong dose delivery, and thus incur under-treatment or even serious side-effects (e.g. rectum bleeding). To ensure the high efficacy of treatment, accurate prostate segmentation from CT images is critical. On the other hand, traditionally, the prostate and surrounding organs are often manually segmented by physician(s). This process is time-consuming and also suffers from both intra- and inter-observer variations [2]. Therefore, automatic and accurate prostate segmentation is highly desired in IGRT.

Despite the importance in IGRT, automatic and accurate prostate segmentation from CT images is still a challenging task due to the following three reasons. First, the image contrast between the prostate and surrounding structures is low, as shown in

Y. Shao and Y. Gao—Co-first authors.

(a)–(c) of Fig. 1. Second, the motion of the prostate is unpredictable for different patients. Third, due to the uncertainty in the existence of bowel gas, the prostate appearance is highly variable, as can be seen by comparing (a) and (c) of Fig. 1.

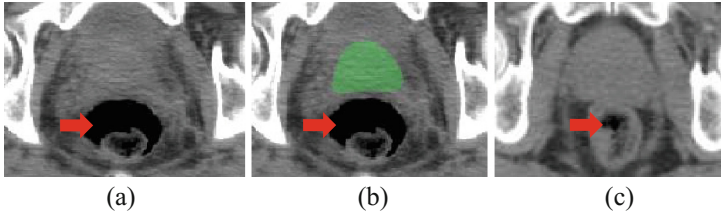


Fig. 1. Typical prostate CT images. The green area in (b) indicates the manual segmentation of prostate by a physician for the same image in (a). Image in (c) shows the prostate image of another patient with less bowel gas than both (a) and (b). Here, the red arrows indicate the bowel gas (Colour figure online).

To address these challenges, many prostate segmentation methods have been proposed for CT images. The methods in [3–5] use patient-specific information to localize the prostate. For these methods, images from the same patient are exploited to facilitate prostate segmentation. Feng et al. [3] leveraged the population and patient-specific image information for deformable segmentation of prostate. Liao et al. [4] collected the previous segmentations of the same patient to update the training images, under a hierarchical sparse label propagation framework, for accurate prostate segmentation. Gao et al. [5] employed the previous prostate segmentations of the same patient as patient-specific atlases to segment the prostate in CT images. Since no previous images from the same patient are available in planning stage, the methods in [3–5] cannot be directly applied to the prostate segmentation in planning CT images. Therefore, it is critical to develop a population-based segmentation method. Costa et al. [6] presented a non-overlapping constraint from nearby bladder on a coupled deformable models for prostate localization. Lu et al. [7] applied information theory to boundary inference process for pelvic organ segmentation. Chen et al. [8] adopted a Bayesian framework with anatomical constraints from surrounding bones to segment the prostate.

Recently, motivated by random forest [9], the regression-based voting strategy achieves promising results in medical image segmentation. For example, Criminisi et al. [10] employed the regression forest to vote the centers of organs' bounding boxes. Lindner et al. [11] adopted the regression forest to predict the optimal positions of global and local models for accurate proximal femur segmentation. These works all train one regression forest for each specific point (e.g. center of the bounding box or point of the deformable model). However, it is difficult to extend this schema to boundary detection, especially in 3D case, since there might be a large number of points on the boundary.

In this paper, we present a new voting strategy to detect the weak boundary for prostate segmentation. Different from the previous methods, our method learns only a single global regression forest to estimate and vote the nearest boundary points for enhancing the entire prostate boundary, then guiding the later deformable

segmentation. The advantages of our method include: (1) our method does not require point-to-point correspondences for learning the regression forest, thus avoiding the difficulty in capturing the correspondences of the 3D prostate boundary points from different subjects in the boundary regression; (2) our method does not require the effort to train one regression forest for each boundary point of 3D object, thus avoiding the training of a large number of regression forests.

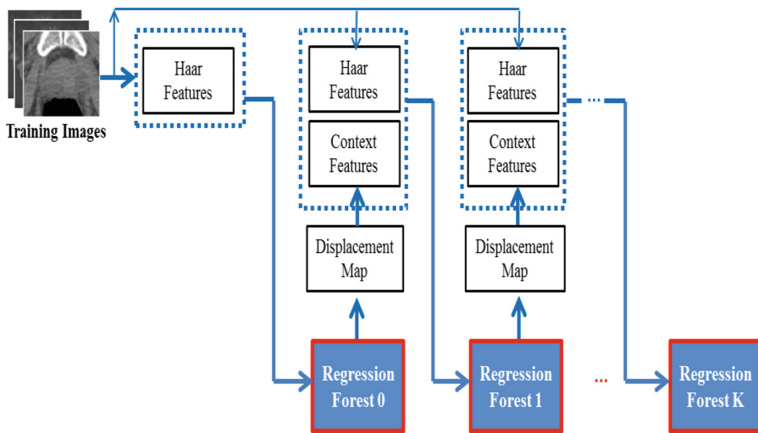


Fig. 2. Training a sequence of regression forests in our proposed method.

2 Method

To accurately segment the prostate from CT images, we first propose a boundary regression method, based on the regression forest, to estimate and vote the nearest prostate boundary points from each image point according to its local image appearance. A boundary voting map is thus obtained to enhance the whole prostate boundary in each CT image (as demonstrated in Fig. 3). Then, to further boost the performance of our boundary regression method, we combine the regression forest with the auto-context model [12] to achieve a more accurate boundary voting map. Specifically, in the training stage (Fig. 2), we train a sequence of regression forests, each of which can estimate a 3D displacement vector from each image point to its nearest prostate boundary, using the local appearance (of this image point). Different from the previous works [10, 11], which often use only the image appearance features, our method further utilizes the context features from the output displacement map of the previous regression forest to train next regression forest. Iteratively, our method improves the estimation of displacement vectors in the whole image, and finally obtains an improved boundary voting map. In the testing stage, given a testing image, the learned regression forests could be sequentially applied to estimate the 3D displacement vector for each point. As more regression forests are applied, the prediction of the 3D displacement vectors in the whole image can be more accurate, and thus can be used in the regression voting strategy to generate a better boundary voting map. Finally, a deformable model,

which has been trained in the training stage, can be applied to the obtained boundary voting map for final prostate segmentation.

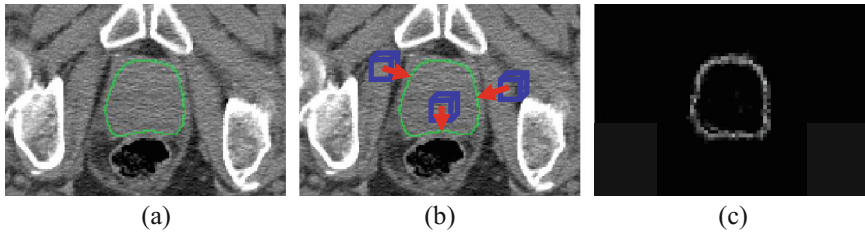


Fig. 3. Demonstration of our boundary regression and voting for prostate. (a) shows the original image and the manual contour (green). (b) demonstrates our boundary regression by local image patches (blue cubes). Red arrows indicate the displacement vectors (regression target) from the centers of local image patches to their corresponding nearest boundary points. (c) shows the boundary voting map of our method using only image appearance features (Colour figure online).

2.1 Boundary Regression and Voting

Motivated by [10], we propose to employ the regression forest for voting the prostate boundary. A regression forest is used to learn a non-linear mapping from the local patch of each image point to its nearest prostate boundary point, obtaining a 3D displacement. Note that, for voting prostate boundary, this 3D displacement is defined as a 3D vector from each image point to its nearest prostate boundary point. Specifically, in the training stage, from each training image, we can first randomly sample a large number of image points $\mathbf{p}_i (i = 1, 2, \dots, N)$ around the manually-delineated prostate boundary. Each sampled point \mathbf{p}_i is characterized by the extended Haar features \mathbf{f} [13], extracted from its $w \times w \times w$ local image patch. Then, the displacement $\mathbf{d}_i = (\Delta x_i, \Delta y_i, \Delta z_i)$ between point \mathbf{p}_i and its nearest prostate boundary point is considered as the regression target. Based on all pairs of Haar features and displacements, $\langle \mathbf{f}(\mathbf{p}_i), \mathbf{d}_i \rangle$, from all training images, we can learn a regression forest R_0 . For our regression forest, the split node in each decision tree is determined by the best combination of feature and threshold, which can achieve the maximum information gain [10] from splitting. Each leaf node in a decision tree stores the mean displacement of training samples falling into this node. In the testing stage, given a testing image, we use all image points in a region of interest (ROI) for boundary regression and voting. Specifically, for each image point $\hat{\mathbf{p}}$, its extended Haar features $\mathbf{f}(\hat{\mathbf{p}})$ can be first extracted using the same way as described in the training stage. Then, the respective displacement $\hat{\mathbf{d}} = R_0(\mathbf{f}(\hat{\mathbf{p}}))$ can be predicted by the trained regression forest R_0 . Finally, a vote will be accumulated on the position $\hat{\mathbf{p}} + \hat{\mathbf{d}}$ in the boundary voting map. By visiting all possible points in the testing image, we can get a boundary voting map for the prostate, with a typical example shown in Fig. 3.

2.2 Refinement of Boundary Voting Map by the Auto-Context Model

To refine the boundary voting map, we adopt the auto-context model [12] to iteratively train a sequence of regression forests by integrating both image appearance features

and the context features extracted from the intermediate displacement map of the previous regression forest. Specifically, we can learn a sequence of regression forests $R_i (i = 0, 1, \dots, K)$ by using the same technique described in Sect. 2.1. The regression forest R_0 (as detailed in Sect. 2.1) is trained by using only the Haar features extracted from the training images, while the latter regression forests $R_i (i = 1, 2, \dots, K)$ are trained by using both the Haar features extracted from the training images and the context features extracted from the respective intermediate displacement maps of the previous regression forest R_{i-1} . The context features used here are again the extended Haar features extracted from the local patch of the intermediate displacement map, instead of the radiation-like features as used in the previous work [12]. This is because the traditional radiation-like auto-context features are the voxel-wise values in the displacement map, which are sensitive to noises in the displacement map. In contrast, the extended Haar features are computed based on local patches, thus more robust to the wrong predictions of displacements produced by the previous regression forest R_{i-1} . Moreover, the extended Haar features provide much richer feature representations than the voxel-wise values for learning the regression forest. Empirically, for the auto-context model, Haar context features achieve faster convergence rate than the traditional radiation-like context features.

Based on the trained regression forests, displacement maps can be sequentially estimated for a testing image. Specifically, the regression forest R_0 is first employed to predict the 3D displacement vectors of the first displacement map, using only the local appearance features from the testing image. Then, by combining the local appearance features (from the testing image) with the context information (from the displacement map of previous regression forest), the latter regression forests $R_i (i = 1, 2, \dots, K)$ could iteratively refine the prediction of the 3D displacement vector for each point in the testing image, and obtain more and more accurate boundary voting maps, as shown in Fig. 4.

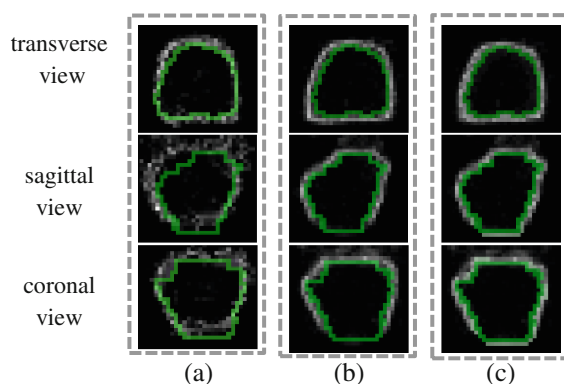


Fig. 4. Boundary voting maps of a patient generated by the auto-context model at the 1st, 3rd and 5th iterations (a-c), respectively. The green curves indicate the manual segmentations (Colour figure online).

2.3 Deformable Segmentation Based on the Boundary Voting Map

Up to this stage, the boundary voting map for prostate is achieved by the regression-based voting with the auto-context model. Based on the boundary voting map, the prostate can be readily segmented by a deformable model [14]. Specifically, to apply the deformable model, we need to build a shape model, often with thousands of vertices on the surface. To accomplish that, in the training stage, we first use the marching cubes to extract surfaces from all manually-segmented prostates in the training images. Then, a template surface is selected and warped to all other surfaces to establish voxel correspondences [14]. With the established correspondences, each training prostate surface can be affine aligned onto a template surface space. PCA is then used to build a prostate shape subspace by capturing the major shape variations from all aligned training prostate surfaces. In the testing stage, the mean prostate shape is first transformed onto the testing image as the initial shape for the deformable model, by a similarity transform. Here, the similarity transform is evaluated by minimizing the least square distance between six detected landmarks (superior, inferior, left, right, anterior, posterior) and their counterparts on the mean shape. Note that those six landmarks are automatically detected using the landmark detector described in [10], which is learned on six manually-annotated landmarks in all training images. Based on this shape initialization, each vertex in the shape model can be independently deformed on the boundary voting map, along its respective normal direction, to a position with the maximum boundary votes. By adopting the landmark-guided initialization, we can achieve a robust initialization (the DSC between initial shape and the manual segmentation is about 0.78 for our dataset), which largely decreases the chances of falling into bad local minima for deformable model segmentation. In the meanwhile, the deformed shape is also constrained by the learned PCA shape model. By alternating the model deformation and the shape refinement, the shape model can be gradually driven onto the prostate boundary under the guidance of both boundary voting map and the PCA shape subspace.

3 Experiments

To evaluate the performance of our proposed method, we conduct experiments on a prostate dataset with 70 planning CT images from 70 different patients with prostate cancer. Each image has voxel size $0.938 \times 0.938 \times 3.0 \text{ mm}^3$ which was isotropically resampled to $2.0 \times 2.0 \times 2.0 \text{ mm}^3$ for the experiment. A clinical expert manually delineated the prostates in all 70 images, which we use as the ground truth in our experiment.

In the experiments, we use four-fold cross-validation to evaluate the performance of our method. The parameters adopted in our method are as follows: the number of trees in each regression forest is set as 10; the maximum depth of each tree is 15; the number of candidate features for node split is 1000; the minimum number of samples in each leaf is 5; the patch size w for extracting Haar features is 30; the number of samples drawn around the prostate boundary in each training image is $N = 10000$; the number of iterations in the auto-context model is 5 (i.e., $K = 4$); the PCA shape subspace captures 98 % shape variation, regarding about 18 eigen-modes; and the number of

iterations used for deformable model is 20. The ROI for boundary regression and voting is determined by the tightest bounding box that covers the initial shape of the deformable model.

3.1 Boundary Regression Vs Prostate Classification

Since the image contrast between prostate and the surrounding structures is low, prostate boundary is not clear and even ambiguous in the CT images, which renders difficulty for accurate prostate segmentation. In the literature [5], to address this tough problem, classification-based method has been proposed to distinguish prostate from the background by assigning each image point (voxel) a prostate likelihood value. Specifically, a classifier is trained by using the positive samples from prostate and the negative samples from background. In the testing stage, the learned classifier can be used to voxel-wisely classify the new testing image for producing a classification response map, which is then utilized by the deformable model to finally segment the prostate. To evaluate the effectiveness of our proposed boundary regression method, we conduct a comparative experiment between the prostate classification method and our proposed boundary regression method. Specifically, in the prostate classification method, we use classification forest as the classifier with the same setting (e.g., number of trees, number of features and thresholds, splitting stop criterion) as boundary regression to estimate the posterior probability of each voxel belonging to the prostate. Then, the generated classification response map is used to guide the deformable segmentation by finding the voxel along the normal with the maximum gradient. In contrast, our boundary regression method uses the obtained boundary voting map to guide the deformable segmentation by searching the voxel along the normal with the maximum boundary votes. Note that both methods use the same sampling strategy, features and shape models, as well as the same auto-context model. Table 1 shows the quantitative segmentation results for the two methods, where Dice Similarity Coefficient (DSC) measures the overlap between automated and manual segmentations. ASD denotes Average Surface Distance between automated and manual segmentations.

Table 1. Quantitative comparison between classification and boundary regression.

Method	DSC	ASD
Classification	0.82 ± 0.07	2.47 ± 1.06
Boundary regression	0.85 ± 0.06	2.01 ± 0.81

From Table 1, we can see that our boundary regression method allows for better segmentations, in terms of higher DSC and lower ASD, than the prostate classification method. Also, the performance improvement of our method regarding DSC and ASD is statistically significant ($p < 0.05$). This result proves that our proposed method is more effective to produce a guidance map for steering the deformable segmentation.

3.2 Effectiveness of Using the Auto-Context Model

To show the effectiveness of using the auto-context model for iteratively refining the boundary voting map, in Fig. 4, we have shown three prostate boundary voting maps estimated at the 1st, 3rd and 5th iterations with the auto-context model. As we can see, with the increase of iteration, the prostate boundary becomes clearer and closer to the manual ground-truth. For quantitative evaluation on the final segmentation results, we perform deformable segmentation based on each intermediate boundary voting map, which is generated at each iteration in the auto-context model. As shown in Fig. 5, the accuracy of prostate segmentation increases with iterations (the DSC increases and the ASD decreases). This result shows the effectiveness of the auto-context model in both enhancing the boundary voting map and facilitating the final prostate segmentation.

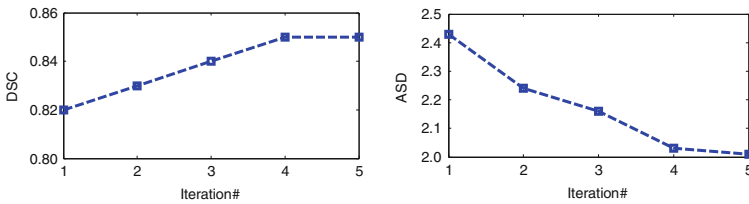


Fig. 5. Iterative improvement of segmentation accuracy with the auto-context model. The left panel shows DSC and the right shows ASD.

3.3 Comparison with Other State-of-the-Art Methods

Due to the unavailability of either source codes or the datasets used by other prostate segmentation methods, it is difficult to directly compare them with the proposed method quantitatively. In order to get a rough understanding of the status in CT prostate segmentation, we list the segmentation accuracies reported by other works in Table 2. To quantitatively evaluate the methods, except the aforementioned DSC and ASD, we also employ other three metrics: sensitivity (SEN), positive predictive value (PPV), and false positive ratio (FPR).

$$\text{SEN} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (1) \quad \text{PPV} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2) \quad \text{FPR} = \frac{\text{FP}}{\text{TP} + \text{FP}} \quad (3)$$

where TP is the number of correctly labeled prostate voxels, FP is the number of falsely labeled organ voxels (i.e., labeling background voxels as prostate voxels), and FN is the number of falsely labeled background voxels (i.e., labeling prostate voxels as background voxels).

As can be seen in Table 2, our method achieves competitive segmentation accuracy to the state-of-the-art methods under comparison, although we utilize only one organ (prostate), without incorporating the constraints from the nearby structures as Costa et al. [6] and Chen et al. [8] do.

Table 2. Quantitative comparison between our method and other methods. *NA* indicates that the respective metric is not reported in the publication.

Method	Mean SEN	Median SEN	Mean ASD	Median ASD	Mean PPV	Median FPR
Costa et al. [6]	0.75	<i>NA</i>	<i>NA</i>	<i>NA</i>	0.80	<i>NA</i>
Lu et al. [7]	<i>NA</i>	<i>NA</i>	2.37	2.15	<i>NA</i>	<i>NA</i>
Chen et al. [8]	<i>NA</i>	0.84	<i>NA</i>	1.10	<i>NA</i>	0.13
Our method	0.84	0.87	2.01	1.85	0.86	0.10

4 Conclusion and Discussion

We have presented a new boundary voting method for CT prostate segmentation. Compared with the previous point regression methods that train one regression forest for each specific point, we learn a single global regression forest for the detection of the entire prostate boundary. To boost the boundary regression performance, we further combine the regression forest with the auto-context model for iteratively refining the boundary voting map of the prostate. Finally, the deformable model is also adopted to segment the prostate under the guidance of both the boundary voting map and the learned prostate shape subspace. Validated on 70 CT images from 70 different patients, our proposed method achieves better segmentation accuracy than the traditional prostate classification method, as well as competitive performance to several state-of-the-art methods under comparison. In this study, due to the relatively stable shapes of the prostates, we use PCA for shape modeling. However, for other organs with complex shape variations (e.g., rectum), a recently proposed shape modeling technique [15, 16], namely sparse shape composition, might be better. This will be our future work.

References

1. A Snapshot of Prostate Cancer. <http://www.cancer.gov/researchandfunding/snapshots/pdf/Prostate-Snapshot.pdf>
2. Foskey, M., Davis, B., Goyal, L., Chang, S., Chaney, E., Strehl, N., Tomei, S., Rosenman, J., Joshi, S.: Large deformation three-dimensional image registration in image-guided radiation therapy. *Phys. Med. Biol.* **50**, 5869 (2005)
3. Feng, Q., Foskey, M., Chen, W., Shen, D.: Segmenting CT prostate images using population and patient-specific statistics for radiotherapy. *Med. Phys.* **37**, 4121–4132 (2010)
4. Liao, S., Gao, Y., Lian, J., Shen, D.: Sparse patch-based label propagation for accurate prostate localization in CT images. *IEEE Trans. Med. Imaging* **32**, 419–434 (2013)
5. Gao, Y., Liao, S., Shen, D.: Prostate segmentation by sparse representation based classification. *Med. Phys.* **39**, 6372–6387 (2012)
6. Costa, M.J., Delingette, H., Novellas, S., Ayache, N.: Automatic segmentation of bladder and prostate using coupled 3D deformable models. In: Ayache, N., Ourselin, S., Maeder, A. (eds.) *MICCAI 2007, Part I. LNCS*, vol. 4791, pp. 252–260. Springer, Heidelberg (2007)

7. Lu, C., Zheng, Y., Birkbeck, N., Zhang, J., Kohlberger, T., Tietjen, C., Boettger, T., Duncan, J.S., Zhou, S.: Precise segmentation of multiple organs in CT volumes using learning-based approach and information theory. In: Ayache, N., Delingette, H., Golland, P., Mori, K. (eds.) MICCAI 2012, Part II. LNCS, vol. 7511, pp. 462–469. Springer, Heidelberg (2012)
8. Chen, S., Lovelock, D.M., Radke, R.J.: Segmenting the prostate and rectum in CT imagery using anatomical constraints. *Med. Image Anal.* **15**, 1–11 (2011)
9. Breiman, L.: Random forests. *Mach. Learn.* **45**, 5–32 (2001)
10. Criminisi, A., Shotton, J., Robertson, D., Konukoglu, E.: Regression Forests for Efficient Anatomy Detection and Localization in CT Studies. In: Menze, B., Langs, G., Tu, Z., Criminisi, A. (eds.) MICCAI 2010. LNCS, vol. 6533, pp. 106–117. Springer, Heidelberg (2011)
11. Lindner, C., Thiagarajah, S., Wilkinson, J.M., Consortium, T., Wallis, G., Cootes, T.: Fully automatic segmentation of the proximal femur using random forest regression voting. *IEEE Trans. Med. Imaging* **32**, 1462–1472 (2013)
12. Tu, Z., Bai, X.: Auto-context and its application to high-level vision tasks and 3d brain image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **32**, 1744–1757 (2010)
13. Zhan, Y., Dewan, M., Harder, M., Krishnan, A., Zhou, X.S.: Robust automatic knee mr slice positioning through redundant and hierarchical anatomy detection. *IEEE Trans. Med. Imaging* **30**, 2087–2100 (2011)
14. Shen, D., Herskovits, E.H., Davatzikos, C.: An adaptive-focus statistical shape model for segmentation and shape modeling of 3-D brain structures. *IEEE Trans. Med. Imaging* **20**, 257–270 (2001)
15. Zhang, S., Zhan, Y., Dewan, M., Huang, J., Metaxas, D.N., Zhou, X.S.: Towards robust and effective shape modeling: sparse shape composition. *Med. Image Anal.* **16**, 265–277 (2012)
16. Zhang, S., Zhan, Y., Metaxas, D.N.: Deformable segmentation via sparse representation and dictionary learning. *Med. Image Anal.* **16**, 1385–1396 (2012)