

Fast segmentation of kidney components using random forests and ferns

Chao Jin*, Fei Shi*, Dehui Xiang, Lichun Zhang, and Xinjian Chen^{a)}

School of Electronic and Information Engineering, Soochow University, Suzhou 215000, China

(Received 8 January 2017; revised 21 August 2017; accepted for publication 8 September 2017; published 7 November 2017)

Purpose: This paper studies the feasibility of developing a fast and accurate automatic kidney component segmentation method. The proposed method segments the kidney into four components: renal cortex, renal column, renal medulla, and renal pelvis.

Method: In this article, we have proposed a highly efficient approach which strategically combines random forests and random ferns methods to segment the kidney into four components: renal cortex, renal column, renal medulla, and renal pelvis. The proposed method is designed following a coarse-to-fine strategy. The initial segmentation applies random forests and random ferns with a variety of features, and combines their results to obtain a coarse renal cortex region. Then the fine segmentation of four kidney components is achieved using the weighted forests-ferns approach with the well-designed potential energy features which are calculated based on the initial segmentation result. The proposed method was validated on a dataset with 37 contrast-enhanced CT images. Evaluation indices including Dice similarity coefficient (DSC), true positive volume fraction (TPVF), and false positive volume fraction (FPVF) are used to assess the segmentation accuracy. The proposed method was implemented and tested on a 64-bit system computer (Intel Core i7-3770 CPU, 3.4 GHz and 8 GB RAM).

Results: The experimental results demonstrated the high accuracy and efficiency for segmenting the kidney components: the mean Dice similarity coefficients were 89.85%, 80.60%, 86.63%, and 77.75% for renal cortex, column, medulla, and pelvis, respectively, for right and left kidneys. The computational time of segmenting the whole kidney into four components was about 3 s.

Conclusions: The experimental results showed the feasibility and efficacy of the proposed automatic kidney component segmentation method. The proposed method applied an efficient weighted strategy to combine random forests and ferns, making full use of the advantages of both methods. The novel potential energy features help random forests effectively segment the kidney components and the background. The high accuracy and efficiency of our method make it practicable in clinical applications. © 2017 American Association of Physicists in Medicine [https://doi.org/10.1002/mp.12594]

Key words: kidney components, random ferns, random forests, segmentation

1. INTRODUCTION

The kidney is the most important organ in urinary system which participates in whole body homeostasis, regulating acid-base balance, electrolyte concentrations, extracellular fluid volume, and blood pressure. Kidney contains four different components with different functions, including renal cortex, renal column, renal medulla, and renal pelvis.¹ Different component of the kidney can be affected by different kidney diseases.^{2–5} Therefore, automatic, accurate, efficient, and robust segmentation of kidney components has great clinical values for renal function assessment, diagnosis, and treatment of kidney diseases.

However, segmentation of kidney components is still challenging due to the following reasons. Firstly, the anatomical structures of kidney are complex. Then, the renal cortex and renal column are connected and have similar intensity. Furthermore, the boundaries between kidney and adjacent organs such as liver and spleen are blurred. Figure 1 shows an illustration of the anatomy of kidney.

Several prior papers^{6–14} tackled the problem of automatic segmentation of kidney or renal cortex in different types of images, i.e., CT or MRI images. In these studies, cortex and

column were classified into one category since they are connected and have similar intensities. However, for clinical application, such as renal cortex thickness measurement, accurate separation of the renal cortex and renal column is needed. Chen et al.¹² proposed an automatic renal cortex segmentation method combining oriented active appearance model (OAAM) and graph cut (GC). In GC-OAAM, the OAAM method combines live wire¹⁵ and AAM to improve the segmentation performance. However, live wire is difficult to extend to 3D. Therefore, OAAM is a 2D method which cannot make use of the context information among slices. Recently, we proposed a 3D automatic segmentation method,¹⁶ which combines active appearance model (AAM) and random forests to segment kidney into four parts accurately. Nonetheless, there are still limitations in the work. The method worked well for kidneys whose volume and structures were not significantly altered by diseases. If diseases such as kidney failure cause kidney atrophy which dramatically decreases the volume of kidney, the AAM trained on the normal dataset may not perform well. Moreover, the computational cost of AAM was high due to its iterative nature.

In this paper, in order to overcome the problems described above, we introduce the state-of-the-art random ferns method

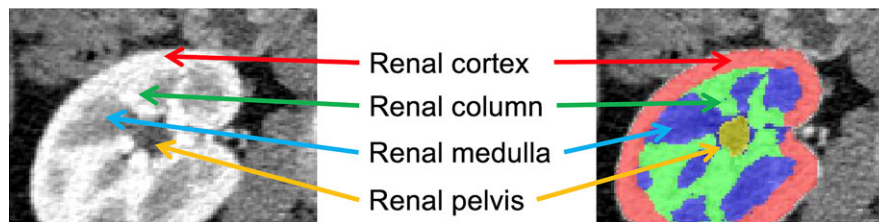


FIG. 1. Illustration of the anatomy of kidney. The renal cortex and renal column are connected and have very similar intensity. But they have different morphology and function. [Color figure can be viewed at wileyonlinelibrary.com]

into kidney components segmentation and combine it with the random forests. The random forests method is a learning based method and was first proposed by Breiman.¹⁷ Thanks to its advantage of dealing with multiclass problem and computational efficiency for processing a huge feature space, it has been widely applied in biomedical and computer vision domain.^{18–24} In recent years, the random forests method is also widely used for medical image segmentation and detection.^{11,16,25–28} The random ferns method proposed by Özuysal et al.²⁵ is an ensemble of constrained trees and was successfully applied in keypoints recognition²⁹ and organ detection.³⁰ Unlike random forests, random ferns applies a seminaive Bayesian scheme and has a simpler and more flat structure. Random forests and random ferns are both highly efficient and highly fault-tolerant. But they have different tree structures and use different features. Random forests can utilize a large amount of 2D and 3D features which reflect voxel intensity and local relationship information, while random ferns utilize fern features which consist of a series of binary features indicating the appearance of the patch surrounding the voxel of interest. By combination of these two methods, different kinds of image information can be exploited in classification. In this paper, we effectively combine random forests and random ferns methods. The proposed method consists of two phases: coarse segmentation of renal cortex and fine segmentation of kidney components. In the first phase, random forests and random ferns are applied independently, and the intersection of their results is obtained as the coarse segmentation of renal cortex. Then novel potential energy features which indicate spatial relationship between kidney tissues and adjacent organs are extracted based on the coarsely segmented renal cortex. In the second phase, a weighted forests-ferns method is proposed to segment kidney into four components simultaneously.

The contributions of this paper are summarized as follows: (a) We propose an efficient weighted strategy to combine random forests and ferns, which make full use of the advantages of both methods. The random forests applies a large number of features which indicate different kinds of local information, while the random ferns utilizes the fern features which describe the context information in a bigger neighborhood. (b) The novel potential energy features are proposed to indicate the spatial relationship of kidney components and adjacent organs, which can help random forests effectively segment the kidney components and the background.

2. METHOD

The proposed method consists of two phases: coarse segmentation of renal cortex and fine segmentation of kidney components. In the coarse segmentation phase, the center of kidney is first obtained by 3D Generalized Hough Transform. Then the coarse renal cortex is obtained by random forests and ferns, and the potential energy features are generated. In the fine segmentation phase, a weighted forests-ferns method incorporating the potential energy features is implemented to segment kidney into four components. Figure 2 shows the flowchart of the proposed method.

2.A. Coarse segmentation of renal cortex

2.A.1. ROI extraction

Before implementing random forests and ferns segmentation, in order to improve operating efficiency and reduce background interference, the center of kidney is located using 3D Generalized Hough Transform (GHT) and the region of interest is extracted.¹⁶ GHT can be used to detect instance of an object with arbitrary shapes, independent of scale and orientation. Its efficiency and robustness have been proved in many applications of different fields, especially in computer vision.³¹ In the training stage, the mean shape of training kidneys served as the template is stored in the R-table.³² In the testing stage, the normal direction for every voxel in the test image is obtained and used as indices to lookup in the R-table. Each voxel generates a vote to a candidate center of kidney according to the R-table. After traversing every voxel in the testing image, the candidate with the most votes is considered as the center of kidney. Then, a volume with constant size around the center of kidney is extracted, which is large enough to contain the whole kidney.

2.A.2. Feature extraction

Random forests method is capable of handling huge number of features. However, a significant number of these features are weak features which contribute little to the decision of the classifier. What is worse, the existence of such features may compromise the classifier and lead to poor classification results. This problem is usually solved by feature selection.²⁵ In this work, information gain ratio (IGR) is used as the metric to split a node in random trees. Therefore, the features

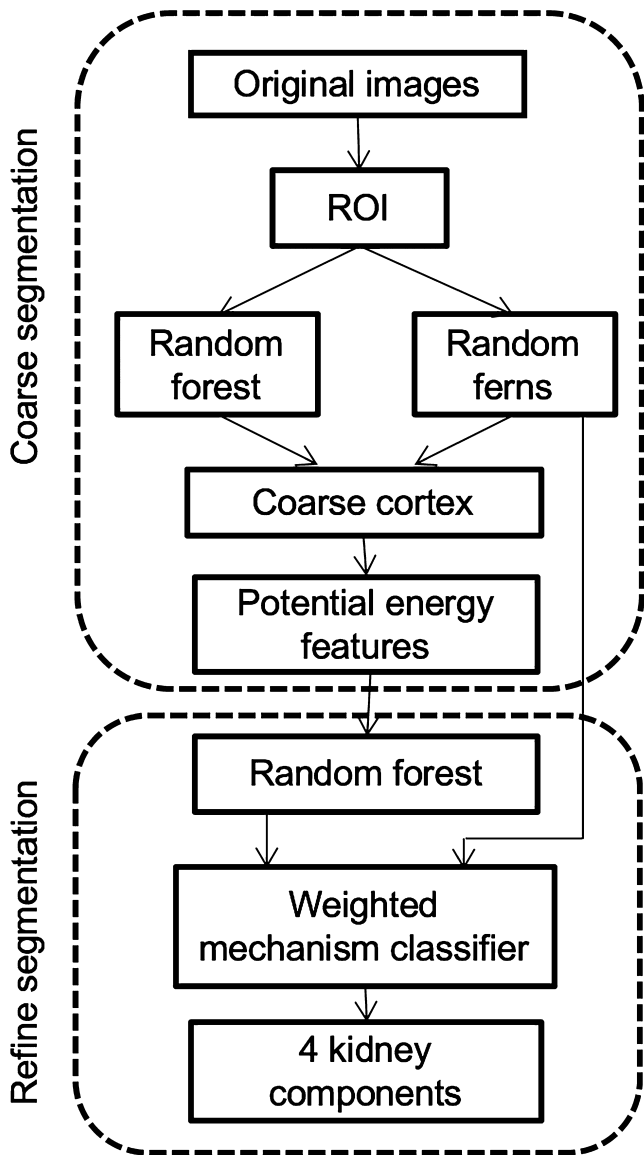


FIG. 2. Flowchart of the proposed method.

with highest IGR are chosen for random forests. For each feature, the IGR is calculated as follows:

$$\begin{aligned}
 IGR &= G(R)/I(D) \\
 G(R) &= \text{Info}(D) - \text{Info}_R(D) \\
 \text{Info}(D) &= -\sum_{i=1}^m p_i \log_2(p_i) \\
 \text{Info}_R(D) &= \sum_{j=1}^k \frac{|D_j|}{|D|} \times \text{Info}(D_j) \\
 I(D) &= -\sum_{j=1}^k \frac{|D_j|}{D} \times \log_2\left(\frac{|D_j|}{|D|}\right)
 \end{aligned} \tag{1}$$

In which, D is the dataset, R represents the attribute set, and p represents the probability for D . $G(R)$ is information gain. $I(D)$ is split information. Since leave-one-out strategy is applied to evaluate the proposed method, we calculated IGR

for each leave-one-out fold of the training dataset and chose 40 best features with the highest IGR to train the RF trees. We found that the features selected in each leave-one-out fold were almost the same.

In coarse segmentation of cortex, for random forests, both 2D and 3D features are used. The 2D features include: Hog features, Gabor features, Robert features and Hessian features. Hog features provide orientation information, Gabor features provide texture information, and Robert and Hessian features provide edge information of objects. 3D features include mean and variance of voxels in a block which contain local intensity information. These features indicate different kinds of information in a relatively small area around the voxel of interest. Random forests method follows the naive Bayes strategy, in which all the features are assumed independent of each other. Namely, random forests does not utilize the correlation between features. A pixel (v) can be classified by conditional probability with features (f_n):

$$C(v) = \underset{k}{\operatorname{argmax}} P(C_k) \prod_{n=1}^N P(f_n|C_k). \tag{2}$$

where C represents the class.

As will be shown in the following sections, with these features (including potential energy features), random forests can be successful in classifying most voxels. But it is difficult to identify voxels near blurry boundaries, especially for the boundary between renal cortex and column, and the boundary between renal medulla and pelvis. In order to achieve correct classification for voxels near the boundaries, features from a bigger neighborhood providing more context information will be helpful. The fern features describe such information using a series of binary features. In this article, as intensity is directly related to the underlying tissue distribution, each binary feature f_i is calculated from the intensities of two voxel I_{i1} and I_{i2} in a relatively big patch around the voxel of interest, defined as follows:

$$f_i = \begin{cases} 1, & \text{if } (I_{i1} \geq I_{i2}) \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

Since the binary feature is very simple, many such features are needed ($N \approx 400$) for accurate classification. Furthermore, the dependencies between these features are utilized. They are evenly divided into M groups of size $S = N/M$. These groups are defined as fern features utilized by random ferns.²⁹ A voxel can be classified using the conditional probability with fern features:

$$C(v) = \underset{k}{\operatorname{argmax}} P(C_k) \prod_{m=1}^M P(F_m|C_k). \tag{4}$$

where F_m is the m th fern feature consisted of S binary features (f_n) which are randomly sampled from N binary features with the sampling without replacement strategy. Once the order of binary features in a fern is set in the training stage, it will not be changed later. This strategy follows a seminaive Bayesian²⁹ approach by modeling only some of the

dependencies between features. In order to improve computational efficiency, binary features in each group are encoded into a decimal fern feature.

The process of generating a forest feature and a fern feature are shown in Fig. 3. The features described above are extracted from a cube or square around the voxel of interest.

2.A.3. Forests and ferns classification

The forest features indicate voxel intensity and local neighborhoods information, while fern features indicate the appearance of the relatively large area surrounding the voxel. In a test image, renal cortex is the most discriminative in intensity, geometry and position, while other kidney components have intensities similar to background or other adjacent organs and have no constant geometry. Therefore, detection of renal cortex is easier than other kidney components. Here, random forests and ferns are implemented to classify the ROI into the four components and the background, and their results are combined to find the renal cortex. Figures 4(a) and 4(b) show the results of random forests and random ferns respectively. One can find that in Fig. 4(a) only renal cortex and column are detected, and they are classified as one category. In Fig. 4(b), renal cortex is segmented roughly. But the result is

very noisy and there are many false positives in the background. However, since the segmentation error of the two results occurs at different positions, we can obtain most renal cortex structure by taking their intersection, as shown in Fig. 4(c).

2.A.4. Potential energy features

As shown in Fig. 1, renal cortex and renal column have very similar intensity and texture. Therefore, it is difficult to separate renal column from renal cortex. The intensity and texture of adjacent organs (background) such as spleen and liver are similar to renal medulla. Even worse, the volume of background tissue is much bigger than renal pelvis and renal medulla. Therefore, renal pelvis and medulla are easily misclassified to background by forests classifier with traditional features. On the other hand, for random ferns, organs near the kidney with similar appearance and texture may also be misclassified as kidney component.

However, one can find that there is a relatively fixed spatial relationship between the four kidney components and organs nearby. Renal pelvis is always in the center of the kidney, renal cortex is the surface layer of kidney, renal medulla and renal column are between the renal pelvis and renal cor-

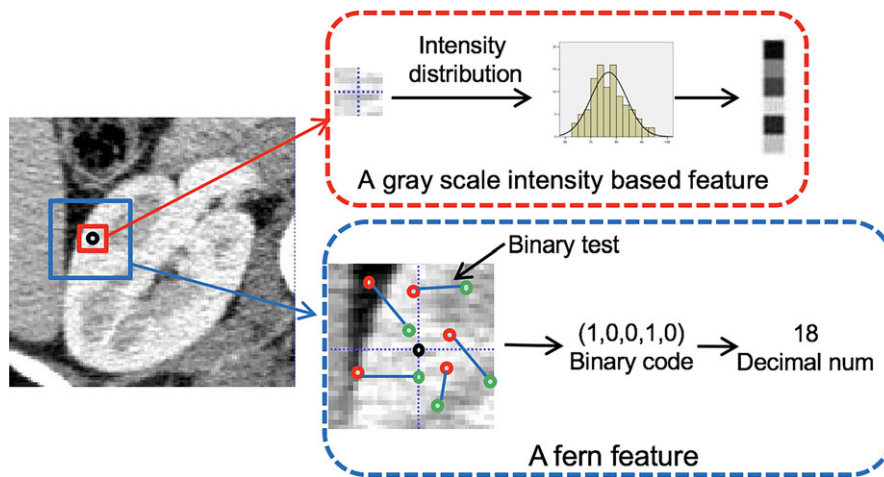


FIG. 3. Examples of feature extraction for random forests and ferns. Forest features indicate local information while fern features indicate appearance information of a relatively large area surrounding the voxel of interest. [Color figure can be viewed at wileyonlinelibrary.com]

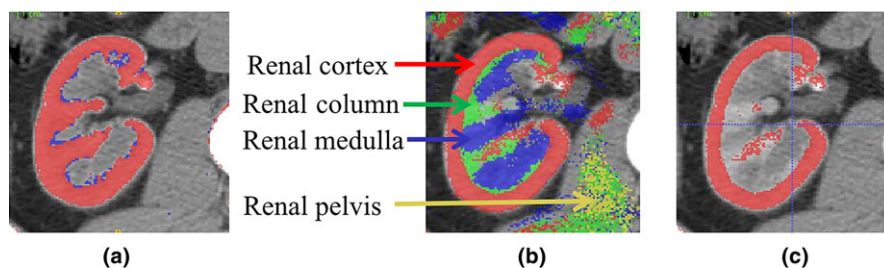


FIG. 4. (a), (b), (c) show the results of random forests, random ferns, and coarse renal cortex respectively. Red, green, blue, and yellow represent renal cortex, column, medulla, and pelvis respectively. [Color figure can be viewed at wileyonlinelibrary.com]

tex, while other organs are far away from the center of kidney, as shown in Fig. 5. Motivated by this, we propose the potential energy features (PEF), which indicate the spatial relationship information to improve the segmentation of kidney. After the center of kidney and coarse renal cortex are obtained, the shape of the kidney can be approximated by an ellipse with long axis L_{long} and short axis L_{short} . The long axis and short axis are defined as two maximum energy equipotential lines, from where the energy fall alone the normal direction. The energy is designed to fall slowly inside the kidney area while fall rapidly outside the kidney. Therefore, the potential energy features for two axes are defined as:

$$E_1(p) = \frac{w}{\frac{d_1^r(p)}{L_{short}} + K} \quad (5)$$

$$E_2(p) = \frac{w}{\frac{d_2^r(p)}{L_{long}} + K} \quad (6)$$

where p stands for a voxel, d_1 and d_2 are its distances to long axis and short axis, respectively. w is a constant, r controls the speed of the energy falling, and K represents a regularization term to avoid potential energy being infinity. K and r are set to 1 and 2 in all our experiments. Figures 6(b) and 6(c) show the two potential energy features extracted from Fig. 6(a).

2.B. Fine segmentation of kidney components

In this step, we add the potential energy features into the random forests feature pool, and implement random forests once again. For the two potential energy features, voxels

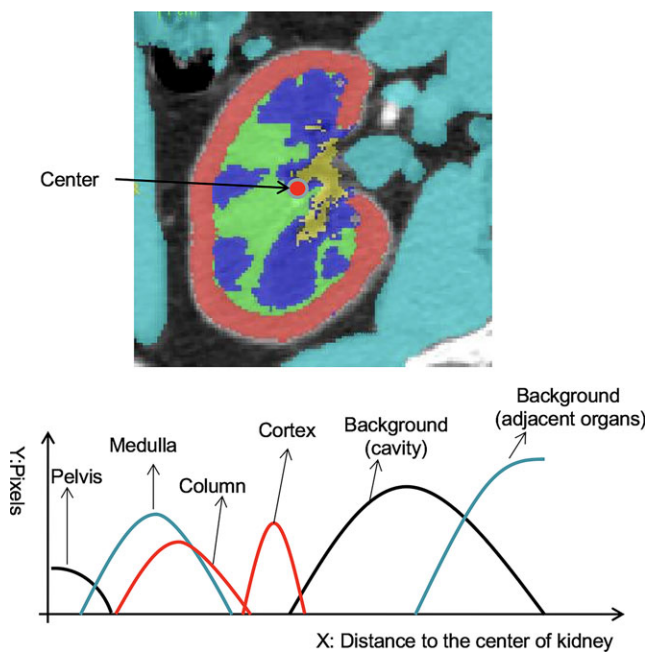


FIG. 5. Illustration of the spatial distribution of kidney components and background according to the distance to the kidney center. Curves with same color mean that they have similar intensity. [Color figure can be viewed at wileyonlinelibrary.com]

inside kidney have high and relatively stable energy. While outside kidney, with the increase in distance between the center of kidney and current voxel, the energy drops sharply. Inside the kidney, although renal cortex and renal column are similar in intensity, their potential energy values are different. For organs outside the kidney, the values of potential energy are very low, which separate them from tissues inside the kidney. Hence, with the potential energy features, the accuracy of random forests is improved. Renal column is separated from renal cortex, and renal pelvis and renal medulla are not misclassified to background, as shown in Figs. 7(d)–7(f).

Although the added potential energy features are effective, the boundary of renal cortex and renal column is still difficult to define. In order to improve the accuracy of segmenting renal cortex and renal column, not only voxel intensity and local relationship information, but also all appearance in the block is needed. Therefore, random ferns method is also needed in the fine segmentation step. Here, we use the results of ferns from coarse segmentation step again and combine it with the updated results of random forests. However, contributions of forests and ferns for each part of kidney are not equal. For example, it is easy to separate renal medulla from cortex and column just depending on pixel intensity difference (with forests), while separating renal column from cortex relies more on local structure appearance (with ferns) and position information. Hence, we propose a weighted mechanism to combine forests and ferns classifiers:

$$C_i = \operatorname{argmax}_j (k_j P_j + r_j Q_j), \quad j = 1, 2, 3, 4, 5. \quad (7)$$

where P_j is the likelihood for each class obtained by random forests, Q_j is the likelihood obtained by random ferns, k_j and r_j are the weights of random forests and random ferns, respectively, j is class label, and C_i is the label assigned to current voxel. The values of j (1, 2, 3, 4, 5) indicate renal cortex, column, medulla, pelvis and background respectively. In addition, we define $k_j + r_j = 1$. In our experiment, we set weights of forests for medulla and background bigger than ferns, while weights of forests for cortex and column smaller than ferns. Figures 7(g)–7(i) shows the final result of the proposed method.

3. EXPERIMENT AND RESULTS

3.A. Image dataset and ground truth

Abdominal images were acquired during preoperative screening from 27 subjects before kidney donation. Among these 27 subjects, seven had only right kidney, and the rest had both sides of kidney, in which 10 also had contrast-enhanced CT images after nephrectomy. Therefore, there were 20 images with both kidneys, and 17 images with only right kidney. In total, we have 37 image volumes with 57 kidneys.

The dataset was acquired from two different types of CT scanner (Light-Speed Ultra, GE Medical Systems, Little Chalfont, UK, and Mx8000 IDT 16, Philips, Amsterdam, Netherlands). The in-plane pixel size ranged from 0.55 to 1 mm. The slices thickness ranged from 1 to 5 mm. Among

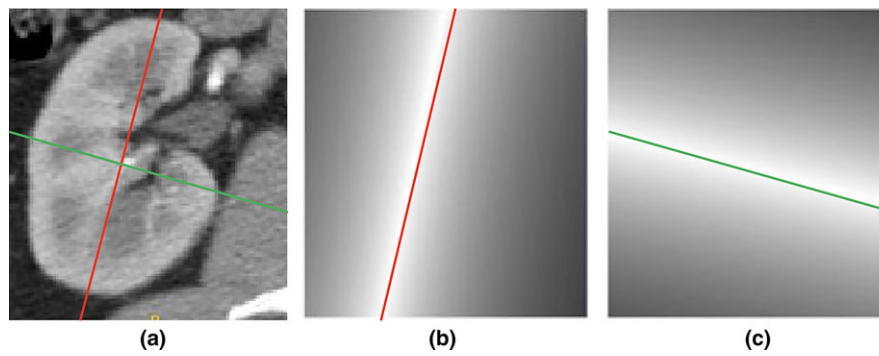


FIG. 6. Potential energy features of long axis and short axis. In (b) and (c), the higher intensity means higher energy. [Color figure can be viewed at wileyonlinelibrary.com]

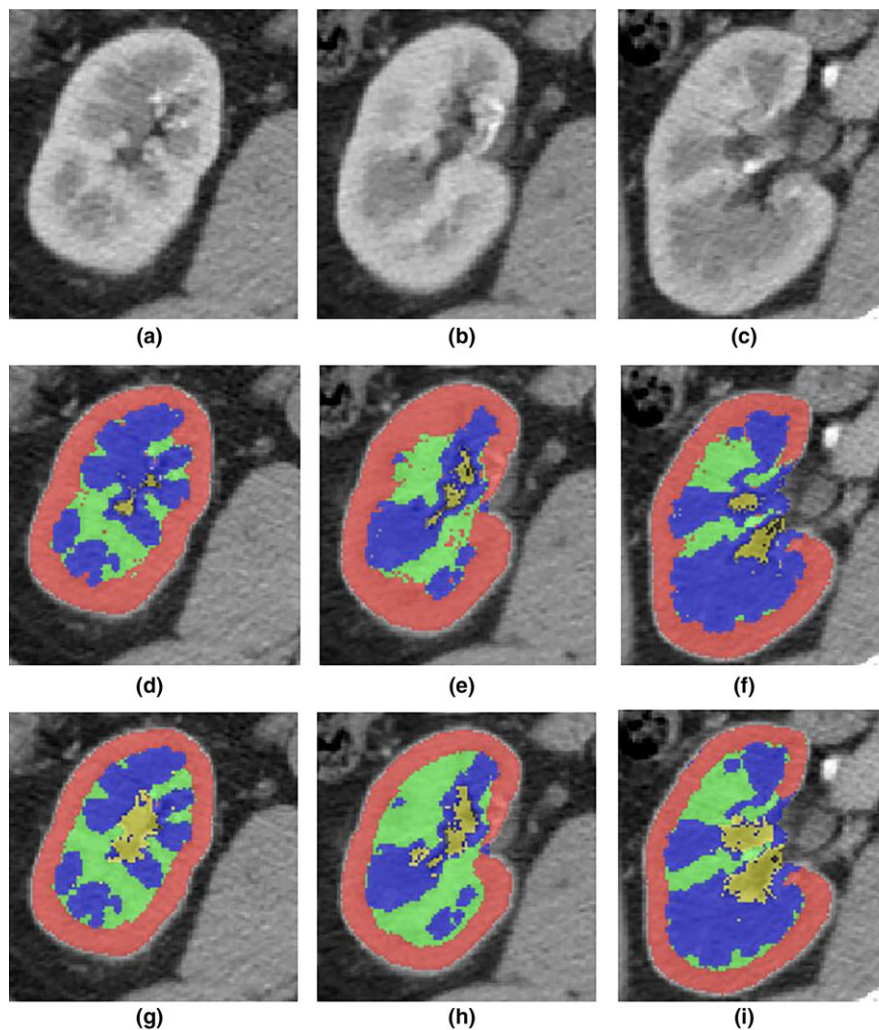


FIG. 7. (a–c) are three slices of the ROI of original image; (d–f) show different slices of the result by random forests with potential energy features. (g–i) are the final result of the proposed method. [Color figure can be viewed at wileyonlinelibrary.com]

these images, 22 images had slice thickness less than 2.5 mm, 15 images had slice thickness greater than 2.5 mm. The slice number of original abdominal images ranged from 51 to 525. The slice number of kidney area in each abdominal image ranged from 20 to 220.

Two independent trained observers (user1 and user2) performed manual segmentation as ground truth. Both the observers were experienced radiologic technologists in hospital. Each observer was blinded to the results of the other. The observers manually drew the kidney components in a slice-

by-slice mode using ITK-snap software (<http://www.itksnap.org>). The experimental results showed that proposed method has a strong correlation with both user1 and user2, and here the user1 is chosen as ground truth.

Since the dataset was not big enough to be partitioned into training and testing sets without losing significant testing capability, we applied leave-one-out strategy as a fair way to validate the proposed method in this paper. Leave-one-out cross-validation involves using one image as the validation set and the remaining images in the dataset as the training set. This is repeated until each image in the dataset is tested.

The proposed method was validated on right and left kidneys separately. Note that for right kidney, one subject might have two scans, taken before and after donation. Two training strategies were used for validation. For the first strategy (Proposed-I) is normal leave-one-out validation, for which kidneys before and after donation were considered different. In order to avoid possible bias, the second strategy (Proposed-II) is leave-one-patient-out validation. For each test data with repeated scans, the other scan from the same subject was excluded from the training data.

3.B. Implementation details

As mentioned in details of the data, the slice number may vary a lot between different images. Interpolation was performed before the ROI extraction, and all the images were interpolated into 32 slices.

The image size in axial plane was 512×512 . We chose the ROI size as $128 \times 128 \times 32$, which was large enough to contain the whole kidney.

For random forests, the number of trees was set as 15, and the maximum tree depth was set as 12. In practice, larger number of trees and depth of tree brought little improvement to segmentation accuracy. The number of selected features was 40. For random ferns, the block size was set as $32 \times 32 \times 3$. For each fern feature, the number of binary test per fern and number of ferns were both empirically set as 18. For the weighted combination of random forests and random ferns, k_j/r_j ($j = 1, 2, 3, 4, 5$) were set as 0.4/0.6, 0.4/0.6, 0.6/0.4, 0.5/0.5, and 0.8/0.2 respectively. The proposed method was implemented and tested on a 64-bit system computer (Intel Core i7-3770 CPU, 3.4G Hz and 8 GB RAM).

The training time of random forests on ROI was about 10 min while the training time of random ferns was only about 20 s. This is because random ferns method has a more concise and flat structure, and uses simpler features than random forests. The testing time on each image for random forests and random ferns were both less than 1 s. In total, the proposed method could segment the whole kidney into four components within 3 s, including the time for coarse segmentation, potential energy features extraction, and fine segmentation.

In order to assess the performance of the proposed method, we compared our approach with three other methods:

- (1) Random forests with potential energy features (RF with PEF). For fair comparison, the feature group and the coarse cortex segmentation step are the same as the proposed method. While in fine segmentation step, only random forests method is used instead of weighted random forests and ferns.
- (2) Random forests with AAM initialization (AAM-RF).¹⁶ This method applied AAM to initialize the renal cortex. In kidney components segmentation step, random forests method is applied with cortex information obtained in AAM initialization. This method used a large amount of 2D and 3D features but did not use potential energy features.
- (3) GC-OAAM¹² (graph cut and oriented active appearance model) based method. In this method, live wire and AAM are combined to serve as the initialization for graph cut. GC-OAAM is used to obtain renal cortex precisely. Then thresholding is applied to segment renal column, renal medulla and renal pelvis.

These methods for comparison were evaluated on right kidneys with the first validation strategy.

3.C. Results and evaluation

Qualitative and quantitative comparisons for the proposed method are provided in this section. As shown in Fig. 8, both sides of the kidneys are segmented into four parts by the proposed method. Figure 9 shows the segmentation results of right kidney by different methods on three different patients.

To objectively evaluate the results of the four components of kidney: renal cortex, renal column, renal medulla, and renal pelvis, the accuracy in terms of dice similarity coefficient (DSC), true positive volume fraction (TPVF), and false positive volume fraction (FPVF) were calculated. DSC is used for comparing the similarity between the automated segmentation results and the ground truth. TPVF indicates the rate of correctly detected volume compared with the ground truth. FPVF denotes the fraction of incorrectly detected volume in the true negative volume.³³

As shown in Table I, the average DSC for the segmentation of renal cortex, renal column, renal medulla, and renal pelvis in right kidney were 90.16%, 81.01%, 86.81%, and 77.60% respectively. The average running time for segmenting one image is shown in Table II. One can find that the performance of the proposed method is much better than the method RF with PEF. This is mainly because the introduction of random ferns allowed more accurate segmentation on the boundary of renal cortex and column, as shown in Figs. 7(d)–7(i) and the third and fourth column of Fig. 9. The performance of the proposed method is comparable to that of AAM-RF, but the proposed method is much more efficient. As shown in the fifth column in Table I, the proposed method can also segment the left kidney into four parts accurately. Comparing the first and last columns of Table I, the difference between results of the two validation strategies

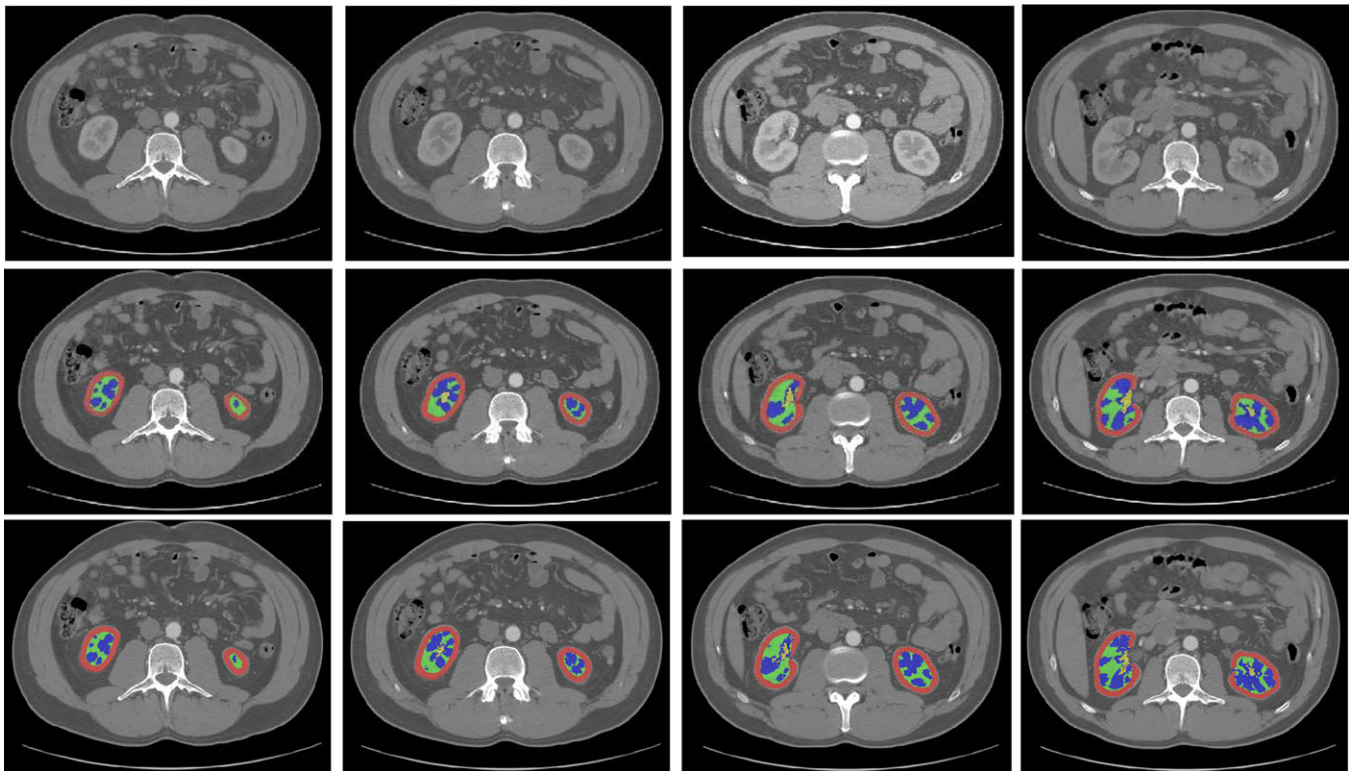


FIG. 8. The segmentation result of both left and right kidneys of one patient in different slices. The first row shows the original image. The second row shows the ground truth and third row shows the result of the proposed method. [Color figure can be viewed at wileyonlinelibrary.com]

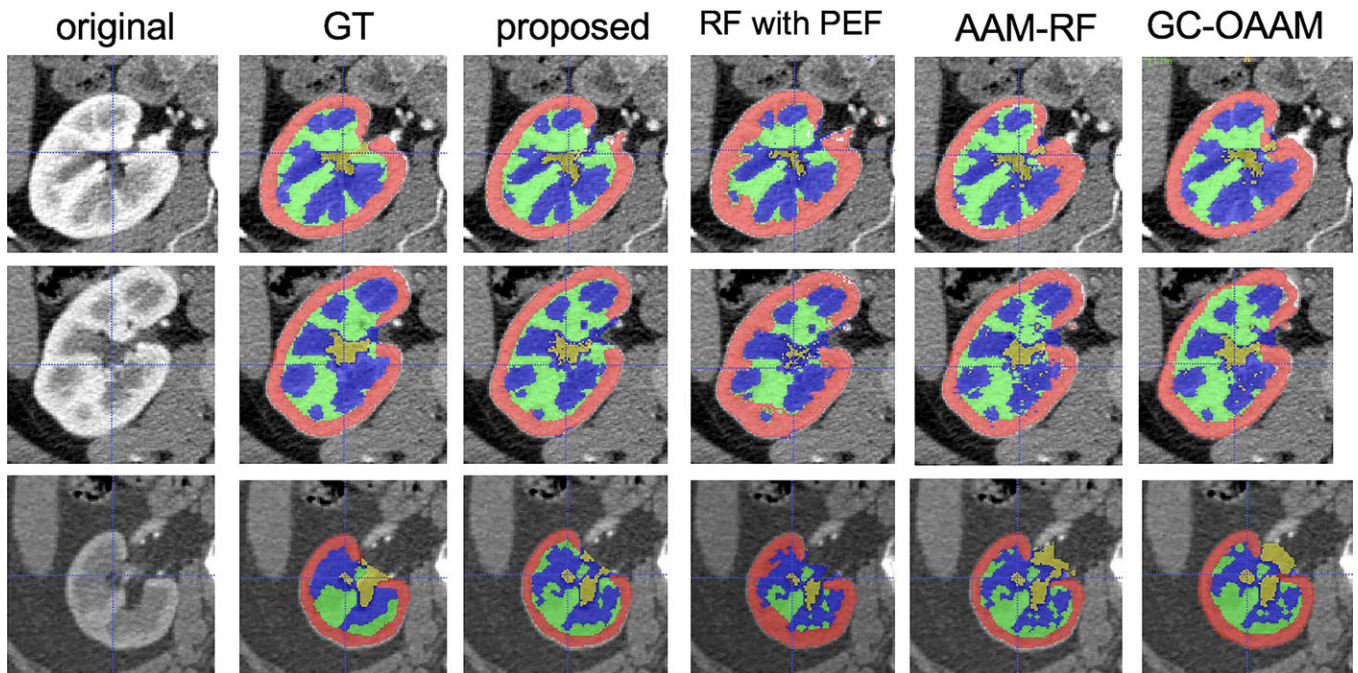


FIG. 9. The segmentation results of right kidney. The performance of the proposed method is better than RF with PEF especially on the boundary of renal cortex and column, and is comparable to that of AAM-RF. [Color figure can be viewed at wileyonlinelibrary.com]

is very small. This is because the right kidney changed a lot before and after donation. As analyzed in our previous work,¹⁶ the volume of four kidney components increased in

most cases. The mean volume change in renal cortex, medulla and pelvis were over 30%. Therefore, the right kidney before and after operation can be considered different,

TABLE I. The segmentation results as DSC, TPVF, and FPVF for different methods: the proposed method, RF with PEF, AAM-RF, and GC-OAAM. Proposed-I and Proposed-II refer to two validation strategies of the proposed method (mean \pm std).

	Proposed-I	RF with PEF	AAM-RF	GC-OAAM	Left kidney	Proposed-II
DSC(%)						
Cortex	90.16 \pm 1.46	84.93 \pm 2.19	91.61 \pm 1.45	84.65 \pm 2.27	89.76 \pm 2.68	89.92 \pm 1.78
Column	81.01 \pm 3.20	71.65 \pm 5.35	83.92 \pm 2.91	73.69 \pm 4.16	80.33 \pm 4.17	80.82 \pm 3.98
Medulla	86.81 \pm 3.64	83.92 \pm 4.07	82.22 \pm 2.22	77.29 \pm 8.05	86.48 \pm 3.11	86.75 \pm 3.57
Pelvis	77.60 \pm 4.90	61.07 \pm 11.04	75.19 \pm 6.93	72.07 \pm 8.60	78.97 \pm 7.06	76.74 \pm 5.18
Overall	83.90 \pm 3.30	75.39 \pm 5.66	83.23 \pm 3.38	76.92 \pm 5.77	83.88 \pm 4.25	83.56 \pm 3.62
TPVF(%)						
Cortex	87.77 \pm 3.16	88.96 \pm 4.49	93.15 \pm 2.21	89.42 \pm 2.96	85.24 \pm 4.06	87.88 \pm 4.16
Column	84.05 \pm 8.08	63.47 \pm 8.11	83.09 \pm 4.64	80.67 \pm 5.11	80.17 \pm 7.53	83.71 \pm 6.99
Medulla	89.26 \pm 3.69	88.79 \pm 3.63	81.92 \pm 9.88	72.16 \pm 13.29	91.05 \pm 6.64	88.58 \pm 4.73
Pelvis	78.88 \pm 4.89	53.86 \pm 12.81	80.28 \pm 7.69	73.18 \pm 9.72	69.39 \pm 10.01	78.04 \pm 6.15
Overall	84.99 \pm 4.95	73.77 \pm 7.26	84.61 \pm 6.10	78.86 \pm 7.77	81.46 \pm 7.06	84.55 \pm 5.50
FPVF(%)						
Cortex	0.31 \pm 0.17	1.85 \pm 0.49	0.37 \pm 0.16	1.34 \pm 0.89	0.37 \pm 0.14	0.32 \pm 0.20
Column	1.04 \pm 0.27	0.59 \pm 0.14	0.97 \pm 0.49	1.85 \pm 0.99	0.79 \pm 0.23	1.07 \pm 0.32
Medulla	0.98 \pm 0.62	1.38 \pm 0.73	0.55 \pm 0.15	0.71 \pm 0.53	1.11 \pm 0.46	0.93 \pm 0.55
Pelvis	0.08 \pm 0.04	0.04 \pm 0.03	0.30 \pm 0.20	0.15 \pm 0.13	0.09 \pm 0.10	0.08 \pm 0.04
Overall	0.60 \pm 0.27	0.96 \pm 0.35	0.55 \pm 0.25	1.01 \pm 0.63	0.59 \pm 0.23	0.60 \pm 0.28

and leave-one-out validation will not lead to much biased results compared to leave-one-patient-out validation.

Although the manual segmentation of user 1 is chosen as ground truth to evaluate the performance in Table I, the performance indices as compared to the ground truth of user 2 are also shown in Table III.

It can be noticed that the highest accuracy is achieved for the renal cortex segmentation, which is due to the fact that the renal cortex has more discriminative intensity, geometry, and position. While the renal pelvis shows the lowest accuracy, which is mainly due to the fact that the shape of renal pelvis varies a lot among subjects and the renal pelvis extends out of the kidney.

In order to demonstrate the robustness of the proposed method, the result for a kidney with a longer and narrower shape than others is shown in Fig. 10. Thanks to the spatial relationship information described by potential energy feature and context information provided by ferns features, the proposed method performed better than AAM-RF which was trained on data with round shape, especially for renal cortex. The renal cortex result of the proposed method [Fig. 10(d)] has more uniform thickness and smoother inner surface than the result of AAM-RF [Fig. 10(e)]. The DSCs for renal

cortex, column, medulla, and pelvis for this case by the proposed method (88.81%, 82.57%, 88.62%, 76.65%) were a little better than AAM-RF (88.61%, 83.57%, 83.99%, 71.39%) and much better than OC-OAAM (83.79%, 76.24%, 85.41%, 80.18%).

In another case, the kidney is more like the shape of a triangle rather than oval. But the proposed method can still segment the kidney into four components accurately as shown in Fig. 11(c). The Dice coefficients for renal cortex, column, medulla and pelvis for this case by the proposed method were 89.01%, 82.43%, 80.73%, and 80.78%. Although the result seems good, there are still segmentation errors on the corner of renal cortex, indicated by yellow arrows in Fig. 11(b). This problem may be caused by the potential energy features which are not flexible enough to fit the shape of the kidney.

TABLE III. The performance indices of the proposed method as compared to the ground truth of user 2 with leave-one-out validation strategy (mean \pm std).

		DSC(%)	TPVF(%)	FPVF(%)
Right kidney	Cortex	89.97 \pm 1.22	89.06 \pm 2.95	0.42 \pm 0.21
	Column	81.22 \pm 2.63	82.04 \pm 6.51	1.00 \pm 0.23
	Medulla	85.24 \pm 5.80	87.67 \pm 5.64	1.03 \pm 0.56
	Pelvis	77.30 \pm 7.09	76.70 \pm 4.50	0.13 \pm 0.06
	Overall	83.43 \pm 4.18	83.86 \pm 4.90	0.64 \pm 0.26
Left kidney	Cortex	89.69 \pm 2.62	85.55 \pm 4.16	0.41 \pm 0.15
	Column	80.04 \pm 4.14	80.00 \pm 7.22	0.81 \pm 0.25
	Medulla	86.03 \pm 3.03	90.32 \pm 2.72	1.10 \pm 0.64
	Pelvis	78.66 \pm 4.87	70.79 \pm 7.60	0.12 \pm 0.16
	Overall	83.60 \pm 3.66	81.66 \pm 5.42	0.61 \pm 0.30

TABLE II. Average computational time for testing.

	Segmentation
Proposed method	3 sec
RF with PEF	3 sec
AAM-RF	20 sec
GC-OAAM	More than 1 min

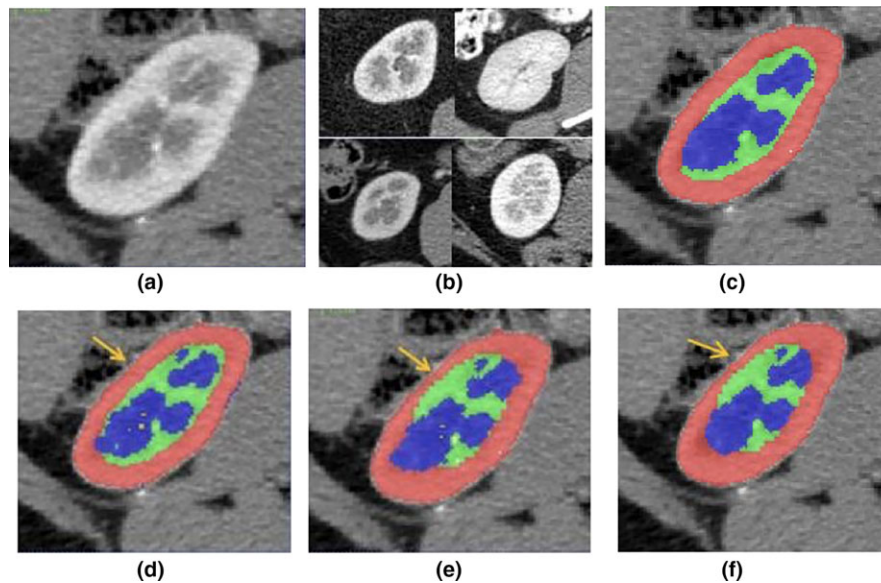


FIG. 10. A case where the kidney has a narrower shape (a) than other cases (b). (b) shows four typical kidneys in the same position. (a) and (b) show the same position of different kidneys. (c) is the ground truth. (d) is the result of the proposed method. (e) and (f) are the results of AAM-RF and GC-OAAM. The arrows show that the renal cortex result of the proposed method (d) has more uniform thickness and smoother inner surface than the results of other methods. [Color figure can be viewed at wileyonlinelibrary.com]

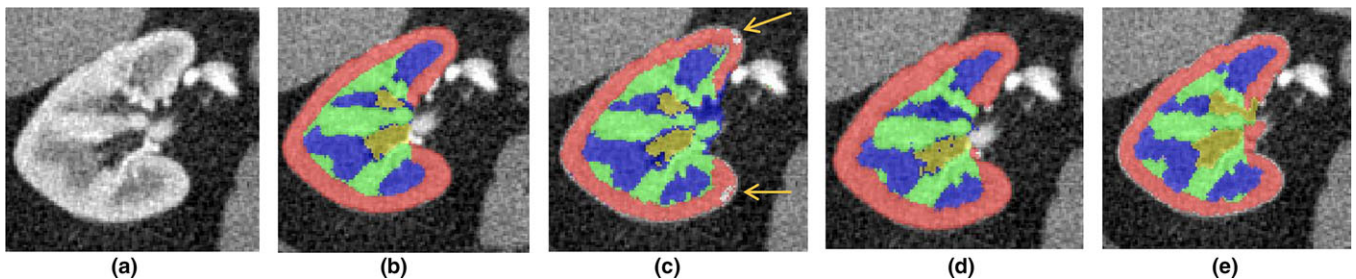


FIG. 11. A case where the kidney shape is different from others, but the segmentation result is good. (a) the ROI of the original image. (b) is the ground truth. (c), (d), and (e) show the segmentation results of the proposed method, AAM-RF, and OC-OAAM respectively. The arrows in (c) show that there are still segmentation errors on the corner of renal cortex by the proposed method. This problem is caused by the potential energy features which are not flexible enough to fit the shape of the kidney. [Color figure can be viewed at wileyonlinelibrary.com]

4. CONCLUSION AND DISCUSSION

This paper presents a fast and accurate automatic kidney component segmentation method. The proposed method consists of two steps: coarse segmentation of renal cortex and fine segmentation of kidney components. In coarse segmentation, results of random forests and ferns using 2D and 3D features are combined in a simple way to get a coarse renal cortex. Then the potential energy features are extracted from the coarse renal cortex. In fine segmentation, the random forests are updated using the potential energy features, and a weighted forests-ferns method is applied to segment the kidney into four components. Experiments conducted on a dataset of 37 images showed the high accuracy and efficiency of the proposed method.

We compared the proposed method to the method RF with PEF to verify the effectiveness of the proposed weighted forests-ferns strategy. The proposed method was

also compared to AAM-RF to prove the high efficiency of our method. It takes only 3 s for the proposed method to segment kidney into four parts, while AAM-RF need about seven times as long to complete the same segmentation task. Meanwhile, the proposed method is also better than GC-OAAM.

It is important to notice that in the generation of potential energy features, we assume that the kidney can be approximated as an ellipse. If diseases such as kidney tumor cause dramatic change in kidney morphology, the proposed method with potential energy features may not perform well. More flexible potential energy features will be developed in the near future to characterize the spatial distribution of kidney tissues and other organs more accurately.

Besides, the small size of data set used in this paper is another limitation. It potentially can lead to positively biased results. In the near future, we will validate the proposed method on data sets both large in number and in variety, as

well as use an independent test set, to verify its effectiveness and robustness.

ACKNOWLEDGMENTS

This work has been supported in part by the National Natural Science Foundation of China (NSFC) under Grant 61622114, 61401294, 61401293, 81371629, 81401472, in part by the National Basic Research Program of China (973 Program) under Grant 2014CB748600, and in part by Natural Science Foundation of the Jiangsu Province under Grant BK20140052.

*These authors contribute equally.

^{a)}Author to whom correspondence should be addressed. Electronic mail: xjchen@suda.edu.cn.

REFERENCES

- Clapp WL. Renal anatomy. In: Zhou XJ, Laszik Z, Nadasdy T, D'Agati VD, Silva FG, eds. *Silva's Diagnostic Renal Pathology*. New York, NY: Cambridge University Press; 2009:1–10.
- Siemer S, Lahme S, Altziebler S, et al. Efficacy and safety of tacho-silas haemostatic treatment versus standard suturing in kidney tumour resection: a randomised prospective study. *Eur Urol*. 2007;52:1156–1163.
- Jun L, Xiaodong Z, Erping L. Study on differential diagnosis of renal column hypertrophy and renal tumors by pulsed subtraction contrast-enhanced ultrasonography, Chinese. *J Ultras Med*. 2006;22:308–310.
- Hart T, Gorry M, Hart P, et al. Mutations of the UMOD gene are responsible for medullary cystic kidney disease 2 and familial juvenile hyperuricaemic nephropathy. *J Med Genet*. 2002;39:882–892.
- Bennington J, Beckwith J. *Tumors of the Kidney, Renal Pelvis, and Ureter*. Washington, DC: Armed Forces Institute of Pathology; 1975.
- Freiman M, Kronman A, Esses SJ, Joskowicz L, Sosna J. Non-parametric iterative model constraint graph min-cut for automatic kidney segmentation. In Proc. of the 13th Int. Conf. of Medical Image Computing and Computed Aided Interventions, MICCAI 2010, Part III, LNCS 6363, pp. 73–80; 2010.
- Lin DT, Lei CC, Hung SW. Computer-aided kidney segmentation on abdominal CT images. *IEEE Trans Inf Technol Biomed*. 2006;10:59–65.
- Ali AM, Farag AA, El-Baz AS. Graph Cuts Framework for kidney segmentation with prior shape constraints. MICCAI, 384–392; 2007.
- Tang Y, Jackson HA, De Filippo RE, Nelson MD, Moats RA. Automatic renal segmentation applied in pediatric MR Urography. *IJHIP*. 2010;1:12–19.
- Xie J, Jiang Y, Tsui H. Segmentation of kidney from ultrasound images based on texture and shape priors. *IEEE Trans Med Imag*. 2005;24:45–57.
- Cuingnet R, Prevost R, Lesage D, et al. Automatic detection and segmentation of kidneys in 3D CT images using random forests. *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2012*. Springer Berlin Heidelberg, 66–74; 2012.
- Chen X, Summers MR, Cho M, Bagci U, Yao J. An automatic method for renal cortex segmentation on CT images. *Acad Radiol*. 2012;19:562–570.
- Will S, Martirosian P, Würslin C, Schick F. Automated segmentation and volumetric analysis of renal cortex, medulla, and pelvis based on non-contrast-enhanced T1- and T2-weighted MR images. *Magma*. 2014;27:445–454.
- Yang X, Le Minh H, Cheng T, Sung KH, Liu W. Automatic segmentation of renal compartments in DCE-MRI images. In: *Medical Image Computing and Computer-Assisted Intervention 2015*. Berlin, Heidelberg: Springer; 2015:3–11.
- Falcao A, Udupa J, Samarasekera S, Sharma S, Hirsch B, Lotufo R. User-steered imagesegmentation paradigms: live wire and live lane. *Graph Models Image Process*. 1998;60:233–260.
- Jin C, Shi F, Xiang D, Chen X. 3D fast automatic segmentation of kidney based on modified AAM and random forest. *IEEE Trans Med Imag*. 2016;35:1395–1407.
- Breiman L. Random forests. *Mach Learn*. 2001;45:5–32.
- Deng H, Runger G. Gene selection with guided regularized random forest. *Pattern Recognit*. 2013;46:3483–3489.
- Gall J, Yao A, Ravazi N, Van Gool L, Lempitsky V. Hough forests for object detection, tracking action recognition. *IEEE Trans Pattern Anal Mach Intell*. 2011;33:2188–2202.
- Shotton J, Johnson M, Cipolla R. Semantic texton forests for image categorization and segmentation. *IEEE conf. CVPR*, 2008, 1–8; 2008.
- Lepetit V, Fua P. Keypoint recognition using randomized trees. *IEEE Trans Pattern Anal Mach Intell*. 2006;28:1465–1479.
- Díaz-Uriarte R, De Andres SA. Gene selection and classification of microarray data using random forest. *BMC Bioinform*. 2006;7:3.
- Chen X, Liu M. Prediction of protein–protein interactions using random decision forest framework. *Bioinformatics*. 2005;21:4394–4400.
- Kandaswamy K, Chou K, Martinetz T, et al. AFP-Pred: a random forest approach for predicting antifreeze proteins from sequence-derived properties. *J Theoret Biol*. 2011;270:56–62.
- Yaqub M, Javaid M, Cooper C, Noble J. Investigation of the role of feature selection and weighted voting in random forests for 3-D volumetric segmentation. *IEEE Trans Med Imag*. 2014;33:258–271.
- Lindner C, Thiagarajah S, Wilkinson J, Wallis G, Cootes T. Fully automatic segmentation of the proximal femur using random forest regression voting. *IEEE Trans Med Imag*. 2013;32:1462–1472.
- Mualla F, Schöll S, Sommerfeldt B, Maier A, Hornegger J. Automatic cell detection in bright-field microscope images using SIFT, random forests, and hierarchical clustering. *IEEE Trans Med Imag*. 2013;32:2274–2286.
- Iglesias J, Liu C, Thompson P, Tu Z. Robust brain extraction across datasets and comparison with publicly available methods. *IEEE Trans Med Imag*. 2011;30:1617–1634.
- Özuyul M, Calonder M, Lepetit V, Fua P. Fast keypoint recognition using random ferns. *IEEE Trans Pattern Anal Mach Intell*. 2010;32:448–461.
- Pauly O, Glocker B, Criminisi A, et al. Fast multiple organ detection and localization in whole-body MR Dixon sequences. In *Medical Image Computing and Computer-Assisted Intervention*. Springer Berlin Heidelberg, 239–247; 2011.
- Ballard DH. Generalizing the hough transform to detect arbitrary shapes. *Pattern Recognit*. 1981;13:111–122.
- Khoshelham K. Extending Generalized Hough Transform to detect 3D objects in laser range data. *ISPRS Workshop on Laser Scanning 2007 and SilviLaser 2007*, Espoo, Sep 12–14; 2007, Finland.
- Ju W, Xiang D, Zhang B, Wang L, Kopriva I, Chen X. Random walk and graph cut for co-segmentation of lung tumor on PET-CT images. *IEEE Trans Image Process*. 2015;24:5854–5867.