An Automatic Method for Renal Cortex Segmentation on CT Images:

Evaluation on Kidney Donors

Xinjian Chen, PhD, Ronald M. Summers, MD, PhD, Monique Cho, MD, Ulas Bagci, PhD, Jianhua Yao, PhD

Rationale and Objectives: The aims of this study were to develop and validate an automated method to segment the renal cortex on contrast-enhanced abdominal computed tomographic images from kidney donors and to track cortex volume change after donation.

Materials and Methods: A three-dimensional fully automated renal cortex segmentation method was developed and validated on 37 arterial phase computed tomographic data sets (27 patients, 10 of whom underwent two computed tomographic scans before and after nephrectomy) using leave-one-out strategy. Two expert interpreters manually segmented the cortex slice by slice, and linear regression analysis and Bland-Altman plots were used to compare automated and manual segmentation. The true-positive and false-positive volume fractions were also calculated to evaluate the accuracy of the proposed method. Cortex volume changes in 10 subjects were also calculated.

Results: The linear regression analysis results showed that the automated and manual segmentation methods had strong correlations, with Pearson's correlations of 0.9529, 0.9309, 0.9283, and 0.9124 between intraobserver variation, interobserver variation, automated and user 1, and automated and user 2, respectively ($P < .001$ for all analyses). The Bland-Altman plots for cortex segmentation also showed that the automated and manual methods had agreeable segmentation. The mean volume increase of the cortex for the 10 subjects was $35.1 \pm 13.2\%$ ($P < .01$ by paired t test). The overall true-positive and false-positive volume fractions for cortex segmentation were $90.15 \pm 3.11\%$ and $0.85 \pm 0.05\%$. With the proposed automated method, the time for cortex segmentation was reduced from 20 minutes for manual segmentation to 2 minutes.

Conclusions: The proposed method was accurate and efficient and can replace the current subjective and time-consuming manual procedure. The computer measurement confirms the volume of renal cortex increases after kidney donation.

Key Words: Kidney; renal cortex; automatic renal cortex segmentation; kidney donors.

The renal cortex, the outer kidney layer consisting of renal corpuscles and convoluted tubules, has distinct morphology and function from the inner renal medulla. Because glomerular filtration, an important clinical assessment of renal function, is the main function of the renal cortex, there has been considerable interest in accurately assessing renal cortex size and volume. The current method for renal cortex segmentation in clinics, however, is mainly operated manually (1–6), which is subjective and tedious. There have been several prior investigations (7–14) of renal cortex segmentation on computed tomographic (CT) and magnetic resonance images, including both semiautomatic (8,10–12) and fully automatic (7,9,13,14) methods. However, most of these studies considered the renal cortex and renal column as one tissue type, although they are anatomically different. In this paper, we propose a method to precisely and automatically segment the renal cortex. To the best of our knowledge, this study is the first work that aims to automatically separate the renal cortex and renal column in renal segmentation.

In kidney transplantation, the ability to accurately and reliably measure renal volume may give clinicians a better understanding of the aftereffects of kidney donation and therefore improve the kidney donor selection process. Limited but available data suggest that larger renal volume is associated with better renal graft function in recipients 1 year after transplantation (15). A few studies (12–16) have estimated kidney volume change after donation by using image findings (such as computed tomography, magnetic resonance imaging, and ultrasound). In this investigation, we also tracked cortex volume change for the remaining kidneys of the donors. As for the volume change in the renal cortex, to the best of our knowledge, we are the first group to measure change after donation.

Acad Radiol 2012; 19:562–570

From the Radiology and Imaging Sciences Department, National Institutes of Health Clinical Center, Building 10, Room 1C515, Bethesda, MD 20892-1182 (X.C., R.M.S., U.B., J.Y.); and the Kidney Disease Branch, National Institute of Diabetes and Digestive and Kidney Diseases, Bethesda, Maryland (M.C.). Received September 26, 2011; accepted January 9, 2012. This paper was presented at the 2010 annual meeting of the Radiological Society of North America as a scientific paper. This research was supported by the Intramural Program of the National Institutes of Health Clinical Center. **Address correspondence to:** J.Y. e-mail: jyao@cc.nih.gov

doi:10.1016/j.acra.2012.01.005

MATERIALS AND METHODS

Donors

Approval for this retrospective study was obtained from our institutional review board, which waived the requirement for informed consent. This study was compliant with the Health Insurance Portability and Accountability Act.

Between 1999 and 2006, 108 patients had undergone kidney donation at our institution. To obtain comprehensive clinical evaluation, all donors were invited to participate in a follow-up study. Out of 108, 52 donors returned for clinical and laboratory exams and detailed interviews. Of the 52, we limited our study to those who had contrast-enhanced CT images and excluded those who underwent magnetic resonance imaging. This yielded 29 patients. Among these 29 patients, 27 donated the left kidney (the right kidney was imaged), and two donated the right kidney (we excluded these patients because the number was too small to build the anatomic model). Among the 27 subjects left in this study, there were seven men and 20 women, with a median age of 46 years (mean, 44 $\pm$ 11 years; range, 19–63 years) at the time of donation. Among these 27 subjects, 10 had contrast-enhanced CT images from both before and after nephrectomy. The median interval between the baseline and follow-up CT examinations for these 10 patients was 1564 days (range, 895–2181 days).

CT Imaging

Abdominal images were acquired during preoperative screening. All examinations were performed with one of two different types of CT scanners (LightSpeed Ultra, GE Medical Systems, Milwaukee, WI; or Mx8000 IDT 16, Philips Medical Systems, Andover, MA). Prior to image acquisition, the patients were injected with 130 mL of Isovue-300 contrast agent (Bracco Diagnostics, Milan, Italy). CT images were reviewed with a three-dimensional (3D) multiplanar reformatting interactive mode on an image-processing workstation (Advanced Workstation; GE Medical Systems). The integrated workstation allows the reader to draw contours and regions of interest and automatically calculates the area enclosed by a region of interest. The slice thickness ranged from 1 to 5 mm. The in-plane pixel size ranged from 0.63 to 0.88 mm. The distribution of the data was as follows: in a total of 37 images, there were 10, 12, and 15 images with in-plane pixel size by slice thickness of $0.63 \times 0.63 \times 1$, $0.71 \times 0.63 \times 2$, and $0.88 \times 0.88 \times 5$ mm^3, respectively.

Documented information included patient age at diagnosis and date of surgery. The examination date, section thickness, and section spacing were also recorded.

Manual Segmentation

Two independent trained observers (users 1 and 2, both radiologic CT technologists with >5 years of experience with kidney CT imaging) performed manual segmentation. User 1 performed manual segmentation for each patient twice with an interval of 3 months, and user 2 performed manual segmentation only once. To prevent bias, each observer was blinded to the results of the other. Segmentation was performed using GE Advanced Workstation thin-client server. The observers manually drew the contours on the cortex boundaries to obtain the outer-layer and inner-layer contours in a slice-by-slice mode. Then the tissues between the two layers were identified as cortex. The total volume of the cortex was obtained by summing all the voxel volumes lying within the boundaries. The total segmentation time was also recorded.

Automatic Segmentation

The segmentation of the renal cortex is not a trivial task. The main challenge in separating the cortex from the rest of the kidney is their quite similar intensity profiles; therefore, it is hard to separate them based only on intensity features. The shape of the kidney can provide necessary constraints. Hence, in this paper, we propose a 3D shape-constrained graph cut (GC) method (17,18) to solve this problem. GC methods have been used widely and have achieved good performance (18). Then, how to obtain the 3D shape constraint becomes essential for the whole method. We propose using the oriented active appearance model (OAAM) method to obtain the shape constraints. The OAAM effectively combines the active appearance model (AAM) (19–21) and live wire (LW) (22) methods. The proposed method consists of three main parts: model building (training), kidney initialization, and renal cortex segmentation, as shown in Figure 1. Details for each part are given as follows.

Model building and training. During the model building (training) step, there are three main steps: (1) manually specify landmarks on the training shapes, (2) construct an AAM, and (3) estimate parameters for the LW and GC methods on the basis of the training data set.

Landmark specification. —Because of the nature of the proposed method (slice by slice), we represent a 3D shape as a stack of two-dimensional contours and manually label the 3D shape slice by slice. For each slice, operators locate the shape visually and then identify prominent landmarks on that shape.

AAM construction. —The conventional AAM method (19,20) is used to construct the model. The model includes both shape and texture information.

Suppose M_j represents the AAM for slice level j and that the number of slice levels is n; the total model M can be represented as

$$M = (M_1,\ M_2,\ \ldots,\ M_n). \quad (1)$$

Although we use the pseudo-3D initialization strategy, we also build the real 3D AAM M_{3D} using the method described by Cootes et al (19). However, this 3D model M_{3D} is used only for providing the delineation constraints, as explained later.

Parameter estimation. —In the proposed method, the initialization is implemented using the multiobject OAAM method,

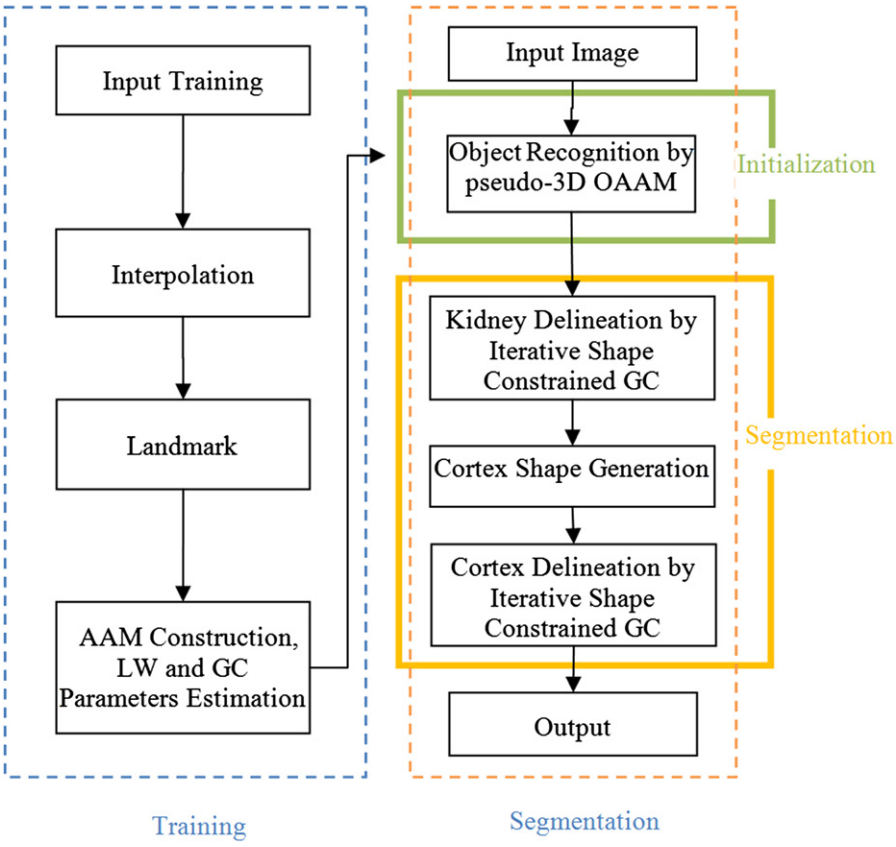

Figure 1. Flowchart of the proposed cortex segmentation system. AAM, active appearance model; GC, graph cut; LW, live wire; OAAM, oriented active appearance model; 3D, three-dimensional.

which combines the AAM and LW methods. The delineation is accomplished using the shape-constrained GC method. During the training stage, the parameters for LW are trained per the LW method (22) from the training data set. The parameter estimation for GC is given below.

Kidney initialization. The initialization step plays an important role in the overall framework. It not only provides shape constraints to the later GC segmentation but also makes the whole method fully automated. The proposed initialization method includes three main steps. First, a slice localization method is applied to detect the top and bottom slices of the kidney. After the localization of the top and bottom slices, all slices in between are linearly interpolated into 32 slices, as in the AAM model construction process . Second, the organ is recognized slice by slice via a multiresolution OAAM method. A multiobject strategy (23) is used to further assist in kidney initialization. Finally, after objects are recognized in all slices, the recognized shapes are stacked together to form the 3D initialization result. In the initialization method, the second step is the key for the success of the whole initialization. More details are given as follows.

Localization of top and bottom slices. —The aim here is to locate the top and bottom slices of the kidney. Because we have already trained a model for each organ slice, we can use this model for slice localization. The proposed method is based on the similarity to the top and bottom slices' OAAM model. For top slice localization in a given image, the top slice model is applied to each slice in the image using the recognition method detailed in the following subsection, and the respective similarity metric is evaluate (equation 2). Then the slice corresponding to the maximum similarity is taken as the top slice of the organ. For the bottom slice detection, a similar method is used.

OAAM. —The conventional AAM matching method for object recognition is based on the root mean square difference between the appearance model instance and the target image. Such a strategy is better suited for matching appearances than for the detailed segmentation of target images. This is because the AAM is optimized on global appearance and is thus less sensitive to local structures and boundary information. Conversely, the LW delineates the boundary very well (22), but it needs a good initialization of landmarks and is an interactive method. Here, we integrate the AAM with the LW method (termed OAAM) to combine their complementary strengths. That is, the AAM provides the landmarks to the LW method, and in return, LW improves the shape model of the AAM. The LW method is fully integrated with AAM in two aspects: (1) LW is used to refine the shape model in AAM, and (2) the LW boundary cost is integrated into cost computation during the AAM optimization method.

In the conventional AAM optimization method, the optimization is based only on the difference between the appearance model instance and the target image. The

boundary cost is not taken into consideration. By combining the boundary cost, the performance of AAM matching can be considerably improved. In the proposed method, the LW technique is integrated into the cost computation during the optimization process. Given a current estimate of the model parameters b, the post t, the texture transformation u, and the image sample at the current estimate g_{im}, our optimization method is as follows:

1. Project the texture sample into the texture model frame using $g_s = T_u^{-1}(g_{\text{im}})$.
2. Evaluate the errors:
   Texture error $E_{\text{AAM}} = |r(\varphi)|^2$, where $\varphi^T = (b^T|t'^T|u^T)$ and $r(\varphi) = g_s' - g_{\text{im}}'$.
   Total error by combining E_{AAM} and LW cost along the shape boundary E_{LW},

$$E_{\text{total}} = \alpha_1 \times E_{\text{AAM}} + \alpha_2 \times E_{\text{LW}}. \tag{2}$$

3. Compute the predicted displacements $\delta_\varphi = -Rr(\varphi)$, where $R = \left(\frac{\partial r^T}{\partial \varphi}\frac{\partial r}{\partial \varphi}\right)^{-1}\frac{\partial r^T}{\partial \varphi}$.
4. Set $k = 1$.
5. Update the parameters $\varphi \leftarrow \varphi + k\delta_\varphi$.
6. On the basis of the new parameters φ, repeat steps 1 to 3, and obtain the new error E'_{total}.
7. If $E'_{\text{total}} < E_{\text{total}}$, accept the new parameters and proceed to step 8; otherwise, try $k = 0.5$, $k = 0.25$, and so on, and return to step 5.
8. Repeat starting with step 1 until no improvement is made to the total error.

Renal cortex segmentation. The shape constrained GC methods is the building block of the proposed renal cortex segmentation algorithm, we presented it first.

Shape-constrained GC. —The segmentation problem can be formulated as an energy minimization problem such that for a set of pixels P and a set of labels L, the goal is to find a labeling f: $P \rightarrow L$ that minimizes the energy function

$$\text{En}(f) = \sum_{p\in P} R_p\left(f_p\right) + \sum_{p\in P, q\in N_p} B_{p,q}\left(f_p, f_q\right), \tag{3}$$

where N_p is the set of pixels in the neighborhood of p, $R_p(f_p)$ is the cost of assigning label $f_p \in L$ to p, and $B_{p,q}(f_p, f_q)$ is the cost of assigning labels $f_p, f_p \in L$ to p and q.

In our framework, the unary cost $R_p(f_p)$ is the sum of a data penalty $D_p(f_p)$ and a shape penalty $S_p(f_p)$ term. The data term is defined on the basis of the image intensity and can be considered as a log likelihood of the image intensity for the target object. The shape prior term is independent of image information, and the boundary term $B_{p,q}$ is based on the gradient of the image intensity.

The proposed shape-integrated energy function is defined as follows:

$$\begin{aligned}\text{En}(f) = &\sum_{p\in P}\left[\alpha \times D_p\left(f_p\right) + \beta \times S_p\left(f_p, x_o\right)\right] \\ &+ \sum_{p\in P, q\in N_p} \gamma \times B_{p,q}\left(f_p, f_q\right),\end{aligned} \tag{4}$$

where α, β, and γ are the weights for the data term, shape term S_p, and boundary term, respectively, satisfying $\alpha + \beta + \gamma = 1$. These components are defined as follows:

$$\begin{aligned}D_p\left(f_p\right) = \Big\{ &-\ln P\left(I_p|O\right) \text{if } f_p = \text{object label} \\ &-\ln P\left(I_p|B\right) \text{if } f_p = \text{background label},\end{aligned} \tag{5}$$

$$B_{p,q}\left(f_p, f_q\right) = \exp\left[-\frac{\left(I_p - I_q\right)^2}{2\sigma^2}\right] \times \frac{1}{d(p,q)} \times \delta\left(f_p, f_q\right), \tag{6}$$

and

$$\delta\left(f_p, f_q\right) = \begin{cases} 1 & \text{if } f_p \neq f_q \\ 0 & \text{otherwise} \end{cases}, \tag{7}$$

where I_p is the intensity of pixel p; object label is the label of the object (foreground); $P(I_p \mid O)$ and $P(I_p \mid B)$ are the probability of intensity of pixel p belonging to object and background, respectively, which are estimated from object and background intensity histograms during the training phase (details given below); $d(p, q)$ is the Euclidian distance between pixels p and q; and σ is the standard deviation of the intensity differences of neighboring voxels along the boundary.

$$S_p\left(f_p, x_o\right) = 1 - \exp\left[-\frac{d_{f_p}(p, x_o)}{r_o}\right] \tag{8}$$

and

$$d_{f_p}(p, x_o) = \begin{cases} d(p, x_o) & \text{if } f_p = \text{source label} \\ d(p, \overline{x}_o) & \text{otherwise} \end{cases},$$

where $d(p, \mathbf{x}_O)$ is the distance from pixel p to the set of pixels that constitute the interior of the current shape $\mathbf{x}_o$ of object O (note that if p is in the interior of $\mathbf{x}_o$, then $d[p, \mathbf{x}_O] = 0$); $d(p, \overline{x}_o)$ is the distance from voxel p to the complementary of the shape x_o; and r_O is the radius of the sphere that roughly encloses $\mathbf{x}_o$. The linear time method of reference (24) was used in this study for computing this distance.

During the training stage, the intensity histogram of each object is estimated from the training images. On the basis of this, $P(I_p \mid O)$ and $P(I_p \mid B)$ can be computed. As for parameters α, β, and γ in equation 4, because $\alpha + \beta + \gamma = 1$, we estimate only α and β by optimizing accuracy as a function of α and

β and set $\gamma = 1 - \alpha - \beta$. We use the gradient descent method for the optimization. Accu(α, β) represents the algorithm's accuracy (here we use the true-positive volume fraction (25)). Parameters α and β are initialized to 0.35 each, and Accu(α, β) is optimized over the training data set to determine the best α and β.

Renal Cortex Segmentation

The proposed renal cortex segmentation method consists of three main steps. First, the kidney is segmented using the iterative shape constrained GC method, which incorporated the kidney shape information generated from the initialization step. Second, after getting the kidney contour, morphologic operations (five times erosion followed by image subtraction) are applied to it to obtain the initial cortex shape. Finally, the renal cortex is segmented by using the iterative shape-constrained GC method. At this time, the initial cortex shape result is used as a constraint. The iterative shape-constrained GC method is given as follows:

1. Perform GC segmentation minimizing equation 3 using the available shape constraints.
2. Compute the new position of the landmarks by moving each landmark in the shape to the point closest on the GC boundary.
3. If no landmarks moved or the distance between two shapes is less than predefined threshold *thr*, then convergence is assumed, and we proceed to step 4; otherwise, subject the shape result to the constraint of model M_{3D} and update it as the new shape constraints, and return to step 1.
4. Perform one final GC segmentation on the basis of the latest shape constraints, and obtain the associated object boundaries.

In our implementation, *thr* is set as 1 unit, and for shape update constraints, we limit the distance each voxel on the surface can move within any iteration to 6 units, which will make the shape change smoother.

Statistical Analyses and Segmentation Evaluation

Linear regression analysis (26) and Bland-Altman plots (27,28) were used to evaluate the correlation and agreement between the automated and manual segmentation methods. Analysis of intraobserver variation (two segmentations by user 1), interobserver variation (between user 1 and user 2), variation between the automated method and user 1, and variation between the automated method and user 2 were performed. The median, mean, and standard deviation of the cortex volume change for the remaining kidney before and after the donation were also calculated.

The true-positive volume fraction (TPVF) and false-positive volume fraction (FPVF) (25) were also calculated to evaluate the accuracy of the proposed segmentation method. The TPVF indicates the fraction of the total amount of tissue in the true delineation, and the FPVF denotes the amount of tissue falsely identified, defined as follows:

$$\mathrm{TPVF} = \frac{|C_{\mathrm{TP}}|}{|C_{\mathrm{td}}|}, \tag{9}$$

and

$$\mathrm{FPVF} = \frac{|C_{\mathrm{FP}}|}{|U_{\mathrm{d}} - C_{\mathrm{td}}|}, \tag{10}$$

where C_{TP} is the set of voxels in the true-positive delineation, C_{td} is the set of voxels in the ground truth, C_{FP} is the set of voxels falsely identified, U_{d} is assumed to be a binary scene with all voxels in the scene domain, and $|\cdot|$ denotes volume.

The proposed method was evaluated using the "leave-one-out" strategy, where "one" refers to one patient. In total, we had 37 images from 27 patients (10 of whom had two CT scans before and after nephrectomy). When the testing image was of a patient who had only one image, the remaining 36 images were used for training, and when the testing image was of a patient who had two images, the remaining 35 images were used for training.

RESULTS

Correlations between Manual and Automatic Segmentation

The linear regression analysis (Figs 2 and 3) showed that Pearson's correlation were 0.9529, 0.9309, 0.9283, and 0.9124 for intraobserver variation, interobserver variation, the automated method and user 1, and the automated method and user 2, respectively ($P < .001$ for all analyses). The Bland-Altman plots (Figs 2 and 3) showed that the 95% limits of agreement were −10.31 to 18.19, −4.81 to 31.90, −16.26 to 21.28, and −8.53 to 30.62 mL for intraobserver variation, interobserver variation, the automated method and user 1, and the automated method and user 2, respectively. The correlation analysis results showed that automated segmentation had strong correlations with both manual segmentation results. Figure 4 shows examples of cortex segmentation results of user 1, user 2, and the automatic method, as well as the 3D visualization results.

Accuracy of Localization of Top and Bottom Slices

The OAAM models of the top and bottom slices were used to identify the locations of the top and bottom slices of the kidney. The kidney was manually checked by an expert to generate the reference standard of top and bottom positions. The average localization error for the top and bottom slices was 7.3 ± 5.2 and 6.1 ± 4.5 mm, respectively.

Segmentation Accuracy Measurement

Because the correlation analysis results showed that automated segmentation had a strong correlation with both

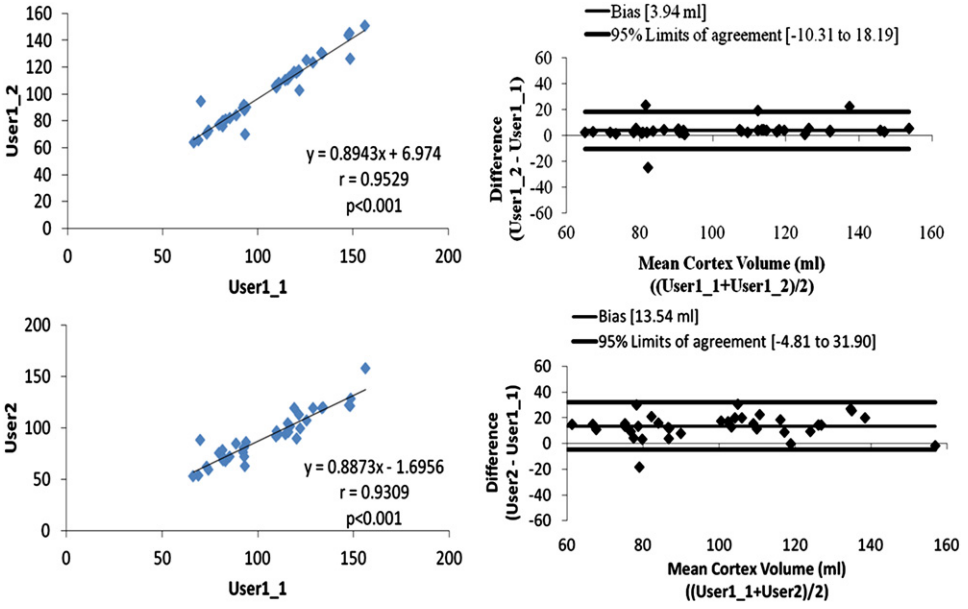

Figure 2. Linear regression analysis and Bland-Altman plots for intraobserver and interobserver assessment of manual segmentation results.

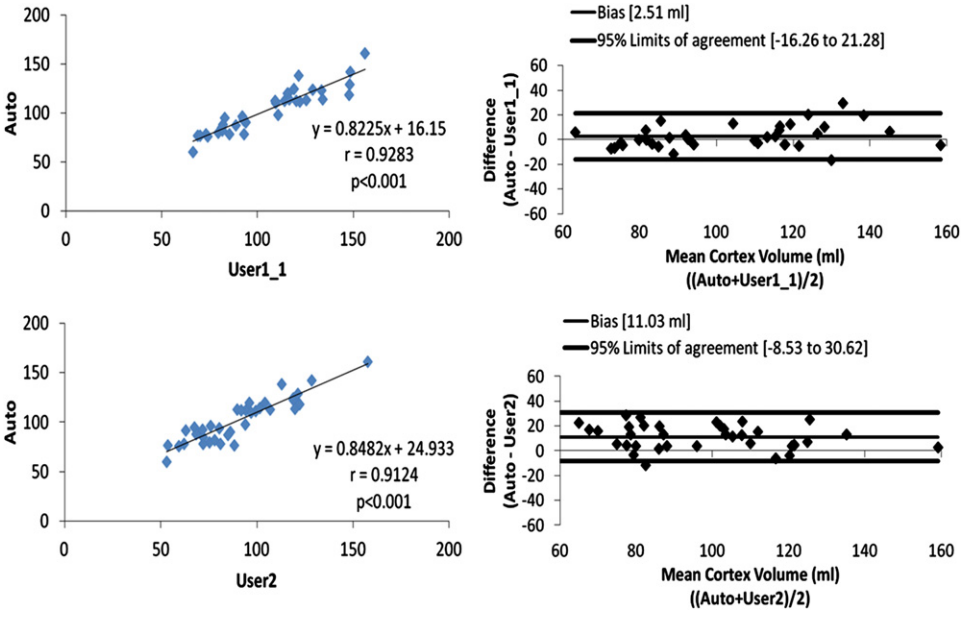

Figure 3. Linear regression analysis and Bland-Altman plots for automated and manual segmentation results.

manual segmentation results, the user 1 first segmentations (User1_1 in Fig 2) were chosen as references. Figure 5 shows the experimental results for one slice level of kidney and cortex segmentation. The TPVF and FPVF for the kidney segmentation over the whole data set were 96.32 ± 6.12% and 0.29 ± 0.05%, respectively. For cortex segmentation, the TPVF and FPVF over the whole data set were 90.15 ± 3.11% and 0.85 ± 0.05%, respectively. The performance on the training data set was as follows: the TPVF and FPVF for kidney segmentation were 97.85 ± 5.06% and 0.18 ± 0.05%, respectively, and for cortex segmentation were 92.89 ± 2.25% and 0.68 ± 0.04%, respectively.

Operation Time Evaluation

The running time for each step was recorded, which consisted of three parts: initialization time (ie, the time required to obtain the initialization results), computing time for kidney segmentation (ie, the time required to automatically delineate the kidney contours), and computing time for cortex segmentation (ie, the time required to automatically delineate the cortex contours). The operation times (expressed as mean ± standard deviation) for two manual and automatic segmentations are reported in Table 1. For automatic segmentation, the total time is divided into the time required for kidney initialization, kidney segmentation, and cortex segmentation. For

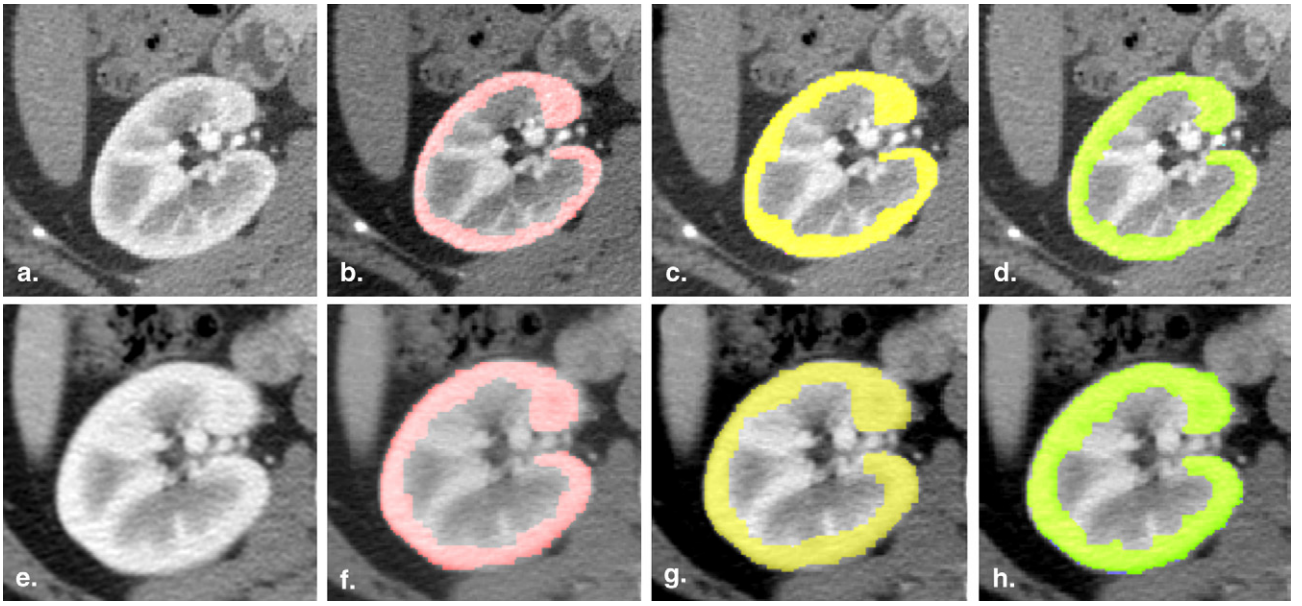

Figure 4. Examples of segmentation results for cortex segmentation. The *top and bottom rows* show the corresponding slices before and after nephrectomy, respectively. **(a)** One slice image before nephrectomy, **(b)** user 1's segmentation results on **(a)**, **(c)** user 2's segmentation results on **(a)**, **(d)** automated segmentation results on **(a)**, **(e)** corresponding slice image after nephrectomy of **(a)**, **(f)** user 1's segmentation results on **(e)**, **(g)** user 2's segmentation results on **(e)**, and **(h)** automated segmentation results on **(e)**.

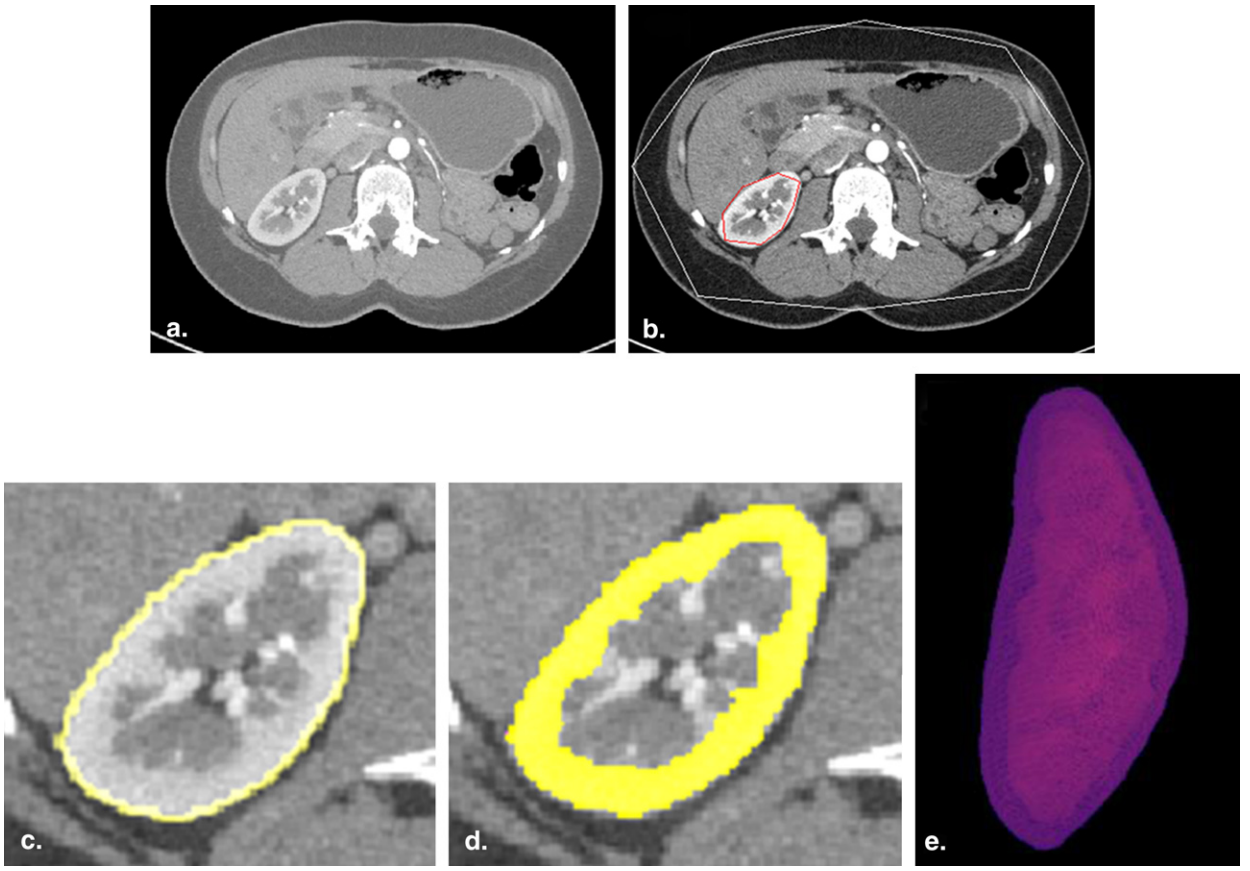

Figure 5. Experimental results of kidney and cortex segmentation on one slice by the proposed method. **(a)** Original slice image, **(b)** initialization results, **(c)** kidney segmentation results, **(d)** cortex segmentation results, and **(e)** three-dimensional visualization of cortex segmentation results.

TABLE 1. Running Time in the Segmentation Procedures

Procedure	Time (s)
Automatic segmentation	
Step 1: initialization	40 ± 5
Step 2: kidney segmentation	35 ± 6
Step 3: cortex segmentation	30 ± 3
Total	
Kidney segmentation (step 1 + step 2)	75 ± 7
Cortex segmentation (step 1 + step 2 + step 3)	105 ± 8
Manual segmentation: user 1	
Kidney segmentation	435 ± 45
Cortex segmentation	1152 ± 60
Manual segmentation: user 2	
Kidney segmentation	486 ± 35
Cortex segmentation	1209 ± 50

Data are expressed as mean ± standard deviation.

kidney segmentation, the time was reduced from an average of 8 minutes for manual segmentation to about 1.3 minutes for automatic segmentation. For cortex segmentation, the time was reduced from an average of 20 minutes for manual segmentation to <2 minutes for automatic segmentation.

Volume Change before and after Donation

The median cortex volume of the remaining kidney was 83.4 mL (mean, 82.9 ± 10.7 mL; range, 60.1–96.1 mL) before donation and 113.1 mL (mean, 112.5 ± 14.8 mL; range, 76.8–128.7 mL) after donation on the basis of automated segmentation results. Figure 4 shows one case of cortex segmentation (kidney #2 in Fig 6) on one corresponding slice before and after nephrectomy. Figure 6 shows the quantitative results: the median volume increase for the renal cortex was 34.1% (mean, 35.1 ± 13.2%; range: 14.6%–57.3%) on the basis of automated segmentation results. The increase was significant ($P < .01$ by paired t test).

DISCUSSION

Figures 2 and 3 show that, first, the intraobserver variation had the highest correlation ($r = 0.9529$), and interobserver variation also shows a high correlation ($r = 0.9309$). Compared to the correlations for manual segmentations, automated segmentation also achieved high correlations with manual segmentation ($r = 0.9283$ for automated segmentation and user 1 and $r = 0.9124$ for automated segmentation and user 2). Second, on the Bland-Altman plots, the 95% limits of agreement were −10.31 to 18.19 and −4.81 to 31.90 for intraobserver and interobserver variation, respectively. Compared to the manual results, automated and manual segmentation had very similar agreement, with 95% limits of agreement of −16.26 to 21.28 and −8.53 to 30.62, respectively, between automated segmentation and user 1 and automated segmentation and user 2. The proposed automated method achieved high correlations and agreements with manual segmentation.

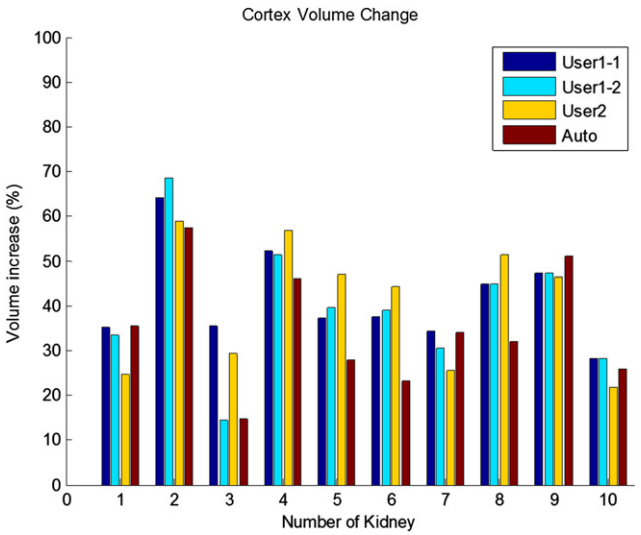

Figure 6. Renal cortex volume change by user 1, user 2, and automated segmentation results.

For the accuracy analysis of the proposed segmentation method, user 1's first segmentations were chosen as references, and high TPVFs of 96.32% and 89.15% and low FPVFs of 0.29% and 0.85% were achieved for kidney and cortex segmentation, respectively. In addition, we also evaluated the proposed segmentation method on the basis of the references of user 1's second and user 2's segmentations to determine the performance differences. The TPVFs were 96.21% and 90.32% (96.02% and 90.01%) and the FPVFs were 0.36% and 0.92% (0.26% and 0.81%) for kidney and cortex segmentation on the basis of the references of user 1's second (user 2's) segmentations, respectively. We can conclude from these results that no matter which reference was used, high TPVFs of >96% and >90% and low FPVFs of <0.5% and <1% were achieved for kidney and cortex segmentation, respectively. The TPVF (FPVF) for cortex segmentation was lower (higher) than for kidney segmentation, which could be due to the greater difficulty in locating the cortex inner wall.

Use of the automatic segmentation algorithm greatly reduced the time needed for the renal cortex segmentation. The reduction of the mean segmentation time from about 20 minutes to <2 minutes makes cortex segmentation on CT imaging more practical in the clinic. The segmentation process consists of three steps: initialization, kidney segmentation, and cortex segmentation. The initialization time might be further reduced by using a hierarchical, multiresolution strategy, which can also improve the initialization accuracy. For kidney and cortex segmentation, we used the iterative shape-constrained GC method. Here, we set the iteration number to 3. With better initialization, the iteration number can also be decreased, which can save additional time.

For the cortex volume change of the remaining kidney before and after donation, all cortex volumes increased on the basis of user 1's two segmentations and user 2's and the automated segmentations, as shown in Figure 6. Because automated segmentation achieved high correlations with

manual segmentation, only the cortex volume change results on the basis of automated segmentation are given here. On the basis of automated segmentation, the average increase of the renal cortex volume was about 35.1%. To the best of our knowledge, this is the first report of the volume change of the renal cortex for kidney donors.

Although one limitation of our study was the size of the study sample, we hope the results presented here will serve as a pilot study prompting further studies with larger patient samples to validate the results.

CONCLUSIONS

We developed a fully automated technique to segment the renal cortex on abdominal contrast-enhanced CT scans. The method was validated on a clinical data set with 37 contrast-enhanced CT images from 27 patients and achieved high correlation with manual segmentation. The use of automated cortex segmentation saved a substantial amount of time. The proposed technique provides an automated, objective, and accurate segmentation of the renal cortex, and it can replace the current subjective and time-consuming manual procedure. Other potential further improvements of the technique include the segmentation and measurement of renal medulla and pelvic volumes.

REFERENCES

1. Lee VS, Rusinek H, Noz ME, et al. Dynamic three-dimensional MR renography for the measurement of single kidney function: initial experience. Radiology 2003; 227:289–294.
2. Van den Dool SW, Wasser MN, de Fijter JW, et al. Functional renal volume: quantitative analysis at gadolinium-enhanced MR angiography-feasibility study in healthy potential kidney donors. Radiology 2005; 236:189–195.
3. Holden A, Smith A, Dukes P, et al. Assessment of 100 live potential renal donors for laparoscopic nephrectomy with multi-detector row helical CT. Radiology 2005; 237:973–980.
4. Halpern EJ, Mitchell DG, Wechsler RJ, et al. Preoperative evaluation of living renal donors: comparison of CT angiography and MR angiography. Radiology 2000; 216:434–439.
5. Sahani DV, Rastogi N, Greenfield AC, et al. Multi-detector row CT in evaluation of 94 living renal donors by readers with varied experience. Radiology 2005; 235:905–910.
6. Low RN, Martinez AG, Steinberg SM, et al. Potential renal transplant donors: evaluation with gadolinium-enhanced MR angiography and MR urography. Radiology 1998; 207:165–172.
7. Shen W, Kassim AA, Koh HK, et al. Segmentation of kidney cortex in MRI studies: a constrained morphological 3D h-maxima transform approach. Int J Med Eng Inform 2009; 1:330–341.
8. Chevaillier B, Ponvianne Y, Collette JL, et al. Functional semi-automated segmentation of renal DCE-MRI sequences. In: Proceedings of the International Conference on Acoustics, Speech, and Signal Processing; 2008:525–528.
9. Lin DT, Lei DT, Hung SW. Computer-aided kidney segmentation on abdominal CT images. IEEE Trans Inform Technol Biomed 2006; 10: 59–65.
10. Li X, Chen X, Yao J, et al. Renal cortex segmentation using optimal surface search with novel graph construction. Lecture Notes Comput Sci 2011; 6893:387–394.
11. de Priester JA, Kessels AG, Giele EL, et al. MR renography by semiautomated image analysis: performance in renal transplant recipients. J Magn Reson Imaging 2001; 14:134–140.
12. Shim H, Chang S, Tao C, et al. Semiautomated segmentation of kidney from high-resolution multidetector computed tomography images using a graph-cuts technique. J Comput Assist Tomogr 2009; 33:893–901.
13. Ali AM, Farag AA, El-Baz AS. Graph cuts framework for kidney segmentation with prior shape constraints. In: Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention; 2007:384–392.
14. Tang Y, Jackson HA, De Filippo RE, et al. Automatic renal segmentation applied in pediatric MR urography. Int J Intel Inform Process 2010; 1: 12–19.
15. Hugen CM, Polcari AJ, Farooq AV, et al. Size does matter: donor renal volume predicts recipient function following live donor renal transplantation. J Urol 2011; 185:605–609.
16. Jeon HG, Lee SR, Joo DJ, et al. Predictors of kidney volume change and delayed kidney function recovery after donor nephrectomy. J Urol 2010; 184:1057–1063.
17. Boykov Y, Kolmogorov V. An experimental comparison of min-cut/max-flow algorithms. IEEE Trans Patt Anal Machine Intel 2004; 26:1124–1137.
18. Kolmogorov V, Zabih R. What energy functions can be minimized via graph cuts? IEEE Trans Patt Anal Machine Intel 2004; 26:147–159.
19. Cootes T, Edwards G, Taylor C. Active appearance models. IEEE Trans Patt Anal Machine Intel 2001; 23:681–685.
20. Mitchell SC, Lelieveldt BPF, van der Geest RJ, et al. Multistage hybrid active appearance model matching: segmentation of left and right ventricles in cardiac MR images. IEEE Trans Med Imaging 2001; 20: 415–423.
21. Chen X, Yao J, Zhuge Y, et al. 3D automatic image segmentation based on Graph Cut-Oriented Active Appearance Models. ICIP; 2010. p. 3653–3656.
22. Falcao AX, Udupa JK, Samarasekera S, et al. User-steered image segmentation paradigms: live wire and live lane. Graph Models Image Process 1998; 60:233–260.
23. Chen X, Udupa JK, Alavi A, et al. Automatic anatomy recognition via multi-object oriented active shape models. Med Phys 2010; 37:6390–6401.
24. Ciesielski KC, Chen X, Udupa JK, et al. Linear time algorithms for exact distance transform. J Math Imaging Vision 2011; 39:193–209.
25. Udupa JK, Leblanc VR, Zhuge Y, et al. A framework for evaluating image segmentation algorithms. Comput Medl Imaging Graphics 2006; 30: 75–87.
26. Cox DR, Hinkley DV. Theoretical Statistics. New York: Chapman & Hall, 1974.
27. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. Lancet 1986; 1: 307–310.
28. Bland JM, Altman DG. Measuring agreement in method comparison studies. Stat Methods Med Res 1999; 8:135–160.