AUTOMATIC EXTRACTION OF THE PELVICALYCEAL SYSTEM FOR PREOPERATIVE PLANNING OF MINIMALLY INVASIVE PROCEDURES

Katarzyna Heryan[1], Andrzej Skalski[1], Jacek Jakubowski[2], Tomasz Drewniak[3], Janusz Gajda[1]

1) AGH University of Science and Technology, Department of Measurement and Electronics, Al. Mickiewicza 30, 30-059 Cracaw, Poland
(✉ heryan@agh.edu.pl, +48 12 617 3596, skalski@agh.edu.pl, jgajda@agh.edu.pl)
2) Rydygier Memorial Hospital, Department of Urology, Os. Złotej Jesieni 1, 31-826 Cracaw, Poland (jacekjakubowski83@gmail.com)
3) Specialized Municipal Hospital G. Narutowicz, Department of Urology, Prądnicka 35-37, 31-202 Cracaw, Poland (tomdrew@vp.pl)

Abstract

Minimally invasive procedures for the kidney tumour removal require a 3D visualization of topological relations between kidney, cancer, the pelvicalyceal system and the renal vascular tree. In this paper, a novel methodology of the pelvicalyceal system segmentation is presented. It consists of four following steps: ROI designation, automatic threshold calculation for binarization (approximation of the histogram image data with three exponential functions), automatic extraction of the pelvicalyceal system parts and segmentation by the Locally Adaptive Region Growing algorithm. The proposed method was applied successfully on the Computed Tomography database consisting of 48 kidneys both healthy and cancer affected. The quantitative evaluation (comparison to manual segmentation) and visual assessment proved its effectiveness. The Dice Coefficient of Similarity is equal to 0.871 ± 0.060 and the average Hausdorff distance 0.46 ± 0.36 mm. Additionally, to provide a reliable assessment of the proposed method, it was compared with three other methods. The proposed method is robust regardless of the image acquisition mode, spatial resolution and range of image values. The same framework may be applied to further medical applications beyond preoperative planning for partial nephrectomy enabling to visually assess and to measure the pelvicalyceal system by medical doctors.

Keywords: the pelvicalyceal system segmentation, kidney segmentation, kidney compartments, Computed Tomography (CT), kidney cancer.

1. Introduction

In 2013 in Poland, there were 5178 new cases of kidney cancer reported. The incidence rates for males and females reached respectively 16.8 and 10.3 per 100.000 [1]. Widely used imaging techniques, especially ultrasonography, cause that nowadays a majority of new renal tumours are detected at an early stage. This provides an opportunity for removing only the cancer lesion and preserving as much of the healthy kidney parenchyma as possible (*nephron sparing surgery* – NSS, *partial nephrectomy* – PN). In recent years, this approach has become a gold standard in treatment of small renal masses. The NSS offers better functional results comparing with the *radical nephrectomy* (RN), with an acceptable profile of peri- and postoperative complications and a similar long-term oncological outcome.

Although the decision regarding treatment feasibility must include all clinical aspects (the patient's age, general conditions, comorbidities, *etc*.), the preoperative planning based on thorough image analysis is essential. It is particularly important to correctly assess the tumour extent and its boundaries, to distinguish all surrounding structures within and beyond the kidney and to examine their mutual topological relations in order to identify possible conflicts in the operated area. This assessment may be based on conventional 2D CT scans, but it can be also supported by 3D reconstructions together with other post-processing techniques, such as

various types of rendering and segmentation. It was suggested that, in preoperative planning or intraoperative navigation, automated segmentation can outperform rendering, as it can produce semi-transparent dynamic models grouping information obtained from different phases, in comparison with limited to a single phase opaque pictures offered by rendering alone [2]. The approach based on image analysis meets the latest trends in surgical oncology. One of its key stages is a 3D segmentation of anatomical structures (kidney, the *pelvicalyceal system* (PCS), vascular tree). In this paper we focus on the PCS segmentation.

According to the European Association of Urology Guidelines on *Renal Cell Carcinoma* (RCC) [3] both CT and MRI are used to characterise the renal mass. The diagnostic decision on RCC is made on the basis of a change of 15 HU or more on the depicted renal mass on CT scans before and after administration of a contrast, preferably in the nephron-graphic phase to maximise differential diagnosis and detection. If the results of CT scan are ambiguous, then MRI may be considered. Additionally, the triple-phase abdominal contrast-enhanced CT provides useful and detailed information on the renal vascular tree and PCS. Although the MRI tissue resolution is considered as higher than CT's, the possibility of acquiring all required information for surgery planning during one examination led to the fact that a triple-phase abdominal CT scan became a standard diagnostic procedure preceding the renal cancer treatment. Therefore, our research material consisted of triple-phase abdominal contrast-enhanced CT scans.

The major challenge of the PCS segmentation on a CT scan is the lack of synchronization between the contrast agent flow and the image acquisition time, which results in visualization of undesired structures and poor recognition of the target ones. Also, the intensity values within the same structure may vary (PCS inhomogeneity). These inconsistencies impact data processing and require a robust segmentation method that can automatically handle different acquisition protocols.

Since the presented issue is relatively new, there are only a few papers concerning the segmentation of kidney compartments. Most of the previously proposed solutions refer to MRI. The algorithms focus on the kidney segmentation without giving particular attention to the PCS. A comprehensive review of the algorithms proposed for kidney volume assessment on MRI provided in [4] presents different acquisition techniques and tailored segmentation methods. Will *et al.* [5] proposed a renal cortex, medulla and renal pelvis segmentation from MRI by simple thresholding. In the case of cortex, the classification is based on a threshold calculated as a row's mean inside the kidney mask minus a standard deviation. With regard to the image value variation within the same structure (explained in Section 2) this kind of solution cannot be applied to the PCS segmentation. Yang *et al.* [6] after kidney segmentation from DCE-MRI data used Principal Component Analysis to describe the data. Based on this, the k-means classification was applied to label data into particular structures. In the last stage, noise and misclassification problems were reduced. They achieved − according to *Dice Coefficient of Similarity* (DICE) − the result of renal pelvis (a part of PCS, Fig. 1) segmentation equal to 0.69 for disordered kidneys and 0.95 for healthy kidneys. Li *et al.* [7] proposed renal cortex segmentation on the CT data using an optimal surface search method. The quality of multiple surfaces' searching method was improved by a graph construction scheme. The segmentation effectiveness was equal to 0.741 ± 0.032 in the true positive volume fraction (precision) and 0.0008 ± 0.0013 in the false positive volume fraction (1-specificity).

Furthermore, we can distinguish two major groups of segmentation algorithms for kidney or/and its compartments: *region growing* (RG) and *level set* (LS) methods. In the RG group three papers are worth mentioning. Pohle *et al.* [8] developed a 2D *adaptive region growing* (ARG) method consisting of two steps. In the first one, the region inhomogeneity is iteratively estimated starting from one seed pixel and after each region doubling the ad hoc coefficients are used to compensate for constant underestimation of the standard deviation. The estimated

www.czasopisma.pan.pl   PAN   www.journals.pan.pl
POLSKA AKADEMIA NAUK

Metrol. Meas. Syst., Vol. 24 (2017), No. 1, pp. 3–18.

parameters (median and both upper and lower standard deviations) are then used in the second RG step. Important assumptions of this method cover the requirement of region compactness, the seed location not on the structure boundary and a sufficient number of pixels in the region. The second paper, written by Abiria *et al.* [9], concerns an RG algorithm (EdgeWave) constrained by a user-specified morphological erosion radius and edge strength. The seed set can be defined by thresholding the selected region, which is facilitated by the average value of a selected region that is displayed in a control window. In the third paper [10], due to the false assumption that there are only two types of renal perfusion signals in the object area, the RG is used to find an approximate solution to Chan-Vese (ChV)-based energy functional [11] to overcome this. The RG is driven by the correlation coefficient quantifying the similarity between the average temporal sequences of two adjacent regions. Among the LS methods developed for kidney segmentation, two can be distinguished [12, 13]. The first one applies a 3D kidney segmentation to DW-MRI images and is based on an adaptive shape, prior guided by the first- and second-order visual appearance features of DWI-MRI data. The kidney and the background models are provided by integrating these features into a joint Markov-Gibbs random field. Although this method is well designed for the kidney segmentation, it is inadequate for the PCS segmentation due to a variety of PCS shapes requiring a huge database to create a reliable model. In the second paper, a 4D LS framework combining both spatial and temporal information is proposed to segment the cortex, medulla and collecting system from a dynamic MRI. Again, since our material is included in a CT database, the temporal information cannot be provided. Although the LS algorithms have been proposed for the MR data and have taken into account their specificity, the common denominator of these two methods is the usage of the ChV term [11].

In this paper we propose an automatic PCS segmentation method on the delayed CT data phase. The presented methodology is based on our previous work – the *Locally Adaptive Region Growing* technique (LARG) [14] driven by an image value with the automatic initialization stage. In order to provide a reliable assesment of the proposed method, its comparison with ARG [8], EdgeWave [9] and ChV [11] methods was performed on the same CT database.

This paper is constructed as follows: Section 2 presents the CT database description together with associated challenges to be addressed; Section 3 highlights the medical importance of PCS segmentation; the proposed segmentation methodology is described in Section 4; the results and conclusion are discussed in Sections 5 and 6, respectively.

2. Material and associated challenges

The research material consisted of 24 abdominal CT scans that were acquired in standard medical procedures preceding the oncological surgery. The studies were performed in various medical centres and on different devices. Individual studies differed in the acquisition mode and using various protocols. This resulted in acquiring a set of data that vary in their image intensity values and spatial resolution (spacing: 0.609−0.885 mm, slice thickness 1−5 mm). The dataset for each patient consisted of three phases (the arterial, venous and delayed ones). In each phase the contrast passes different compartments of the kidney. In the delayed phase most of the contrast agent should be excreted in urine, depicting selectively the urinary tract (including the PCS). Therefore, this phase was chosen for the PCS segmentation. In detail, the PCS segmentation process involves segmentation of several structures that are indicated in Fig. 1. For the sake of simplicity, we refer to all of them as the PCS.

The major challenge is the lack of synchronization between the contrast agent flow and the image acquisition time. This results in visualization of undesired structures and poor recognition of the target ones. For instance, at the beginning of excretory (delayed) phase there is still a high concentration of contrast agent in collecting ducts making papillae barely

distinguishable from calyces. Also, the intensity values within the same structure may vary (from this point of view the PCS is not a homogenous structure).
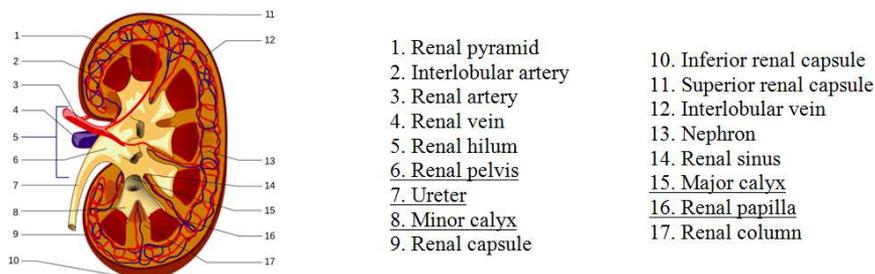

1. Renal pyramid
2. Interlobular artery
3. Renal artery
4. Renal vein
5. Renal hilum
6. Renal pelvis
7. Ureter
8. Minor calyx
9. Renal capsule

10. Inferior renal capsule
11. Superior renal capsule
12. Interlobular vein
13. Nephron
14. Renal sinus
15. Major calyx
16. Renal papilla
17. Renal column

Fig. 1. The kidney compartments. The parts of the PCS are underlined; [F1].

To ensure proper validation of the segmentation results the manual segmentation was performed (48 outlines of PCSs), using the ITK-SNAP software [15], by an experienced urologist. Although the PCS is a continuous structure, in some of the manual outlines there are visible gaps, most commonly in the calyceal neck topography and in the transition between renal pelvis (Fig. 1.6) and ureter (Fig. 1.7). These structures are the narrowest sections of the upper urinary tract. The calyceal neck width normally ranges from 2 to 4mm in 17–48% of patients, depending on the group of calyces, and in the case of 3% of patients the width of the central calyceal is below 2mm [16]. Thus, due to an insufficient transversal resolution and/or thickness of the scanned slices (here, up to 5mm), these structures are incompletely depicted in a CT scan. The discontinuous representation of PCS is another challenge that affects both the accuracy of manual outlines and the effectiveness of segmentation. The situation is also worsened by two following reasons. The PCS is a complex, tree- like structure, which makes the process of manual segmentation of branches into 2D slices more difficult (sometimes a branch is only partially visible on one slice). However, it can be resolved by tracking the outlines simultaneously in other planes (coronal and sagittal), which is extremely time-consuming. The second issue is the slice visualization mode selected for the manual segmentation. Fig. 2 shows how various window setups (*window level* – WL and window width – WW) lead to different manual outlines. The way of performing the manual segmentation has a significant impact on its volume and therefore on the automatic segmentation evaluation. Moreover, the same window setups may give different results depending on an image acquisition protocol (different intensity range and image resolution). Standardization of this process is still a matter of debate.
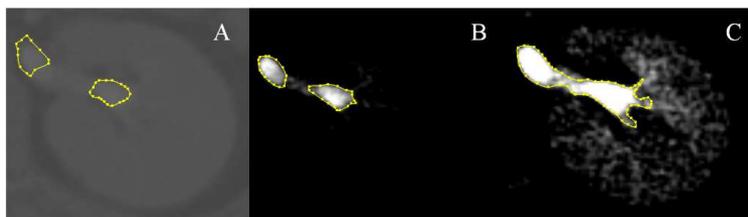

Fig. 2. An example of visualizing CT slices with different windows' parameters; (A) WW = 4137, WL = 1023; (B) WW = 216, WL = 238; (C) WW = 101, WL = 143. The influence of visualization parameters on manual contouring results (marked with a solid line) can be seen.

3. Medical importance of PCS segmentation

Besides preoperative PN planning, the anatomy of PCS is one of the crucial elements assessed, basing on imaging studies, during preoperative planning in many other kidney surgeries. Therefore, in order to facilitate its assessment and measurement, a proper segmentation is necessary.

In the surgical treatment of kidney cancer, a crucial parameter is the distance between tumour and the PCS as whole. It corresponds to the risk of opening the PCS and postoperative complications following unsuccessful reconstruction (urinary fistula). In treatment of kidney stones, the information about both shape and spatial configuration of individual components of the PCS (minor and major calyx and renal pelvis, Fig. 1: 8, 15, 6, respectively) are more important, especially when deposits are located in the lower group of calyces. This specific location hinders the possibility of reaching them during flexible ureterorenoscopy (RIRS) and impairs evacuation of remnants after *extracorporeal lithotripsy* (ESWL).

In the late 80s Sampaio [17] published an analysis of a large series of corrosive endo-casts of cadaveric kidneys' PCSs, determining particular calyceal patterns (anatomical variants) and their distribution. His anatomical classification remains the most widely used system in the preoperative assessment in endo-urology and ESWL. A few years later, together with Aragao [18], he published a paper indicating a correlation between the parameters describing the anatomy of the lower group of calyces (calyceal pattern, pelvicalyceal angle, length and width of the lower calyceal neck) and the effectiveness of ESWL. Since then, a large number of studies were carried out on the role of the PCS anatomy and the efficacy of ESWL/RIRS.

The majority of authors [19–22] use original parameters developed by Sampaio; some of them use their modifications, *e.g.* Elbahnasy [23], whereas others developed new ones, *e.g.* Fong [24]. Among all three aforementioned parameters, the pelvicalyceal angle remains the most important predictor of ESWL efficacy. The European Association of Urology guidelines [25] discourage using ESWL if the lower calyx (containing deposit) is either long (> 3 cm) or narrow (< 5mm), and the pelvicalyceal angle is steep (below 90 degrees). Although these parameters reflect spatial 3D configurations, they are routinely determined, basing on plain urography.

It has been reported that using the CT urography with a 3D reconstruction may increase the accuracy of PCS morphometry assessment, because it enables to measure its parameters in the actual PCS plane, not in a projection on the coronal plane, which is the case in plain urography [26–27]. In *percutaneous nephron-lithotomy* (PCNL), specification of an adequate access to the collecting system represents the most important success factor. It can be facilitated by using a 3D model of a particular PCS.

To meet the requirements of the described procedures and provide accurate measurements, a proper PCS segmentation on 3D CT images is essential.

4. PCS segmentation

The problems and challenges described in Section 2 influence the development of a segmentation method. Due to these problems we have proposed a method aiming at overcoming them (Fig. 3).

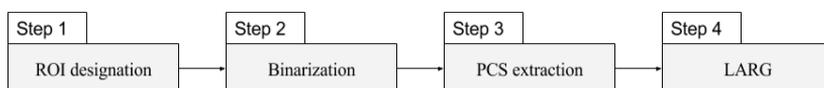| Step 1	Step 2	Step 3	Step 4
ROI designation	Binarization	PCS extraction	LARG

Fig. 3. An outline of the proposed method. A description is provided in the text.

At the beginning, the *region of interest* (ROI) (in this case, the kidney area) is determined (1st step). In the 2nd step binarization is performed in order to acquire the starting seed points for further segmentation. In the 3rd step the PCS parts and ureter are extracted from the binarized ROI image. The final segmentation is performed in the 4th step by the LARG algorithm [14], proposed by the authors.

4.1. Designation of region of interest (ROI), step 1

The 1st step involves manual designation of ROI by indicating the kidney area (maximum span in each anatomical plane) with an adequate margin. It is very important not to crop the kidney as the image is also used for the kidney segmentation, therefore the margins around the kidney reach up to a few pixels (2−3). Additionally, cropping the whole image to ROI enables to significantly reduce the calculation time.

4.2. Binarization of ROI image (BW), step 2

Due to the lack of study protocol standardization (various medical centres, devices, acquisition modes and, therefore, different image spatial resolution and value ranges) it is impossible to explicitly identify universal parameters, in particular the threshold value used for LARG initialization. As a consequence, it becomes necessary to develop an automatic method for accurate threshold determination for each individual case. A characteristic distribution of image values (Fig. 4) enabled to establish the following solution. It is inspired by the method proposed in [28], where a histogram of ROI image values is approximated by the sum of Gaussian functions. The main difference between this and our approaches lies in the way of threshold selection. In [28] the optimal threshold is set at the intersection of particular normal distributions. In the proposed method, the initial segmentation threshold is defined based on the normalized histogram of ROI image values approximated by the sum of three exponential functions:

$$F = \sum_{i=1}^{3} H_i \exp\left(-\frac{(x-a_i)^2}{2b_i^2}\right). \tag{1}$$

The function F (1) is designated by a nonlinear approximation using the Trust-Region-Reflective Least Squares Algorithm [29]. Such an approximation was performed for each of 48 PCS images. In each case, the information extracted from F was used to designate thresholds. The determined image intensity values of a_i were sorted in the ascending order : a_1, a_2, a_3, together with corresponding values of b_i, so that $a_3 = \max(a_i)$. In each case, a threshold was set to $A \cdot a_3 + B \cdot b_3$. In order to calculate parameters A and B, the dataset was divided randomly into two halves constituting the training set and the test set. Then, an optimization procedure, using the Nelder-Mead simplex direct search [30] aimed at minimizing the negation of the sum of DICEs calculated for the training set, was performed. The designated parameter values were as follows: $A = 1.006$, $B = 2.326$. Next, during the remaining steps of the proposed method (3rd and 4th step) only the test set was used. Additionally, to examine the dispersion of A and B parameters' values, a 10-fold partition of the training set was performed and A and B values were calculated independently for each fold. The experiment revealed that the coefficients did not differ significantly for the subsequent folds, as their values were equal to 1.005 ± 0.002 for A and 2.353 ± 0.142 for B. Examples of approximations of the ROI image data by exponential functions are shown in Fig. 4. The calculated threshold value is indicated on the presented plots. Also, the variety of image intensity values depending on the patient CT data can be compared based on these two figures.
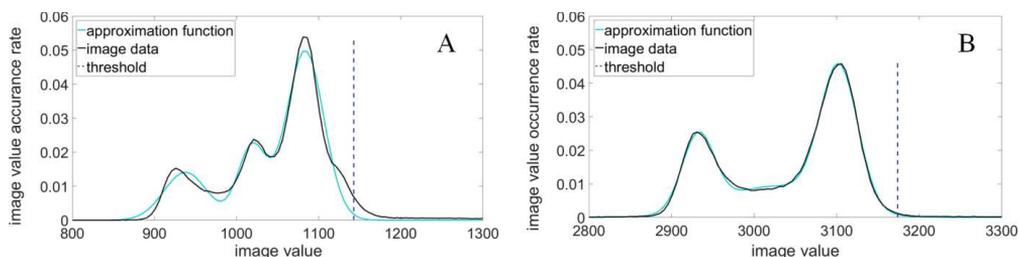
Fig. 4. Histograms of ROI image data normalized with an approximation function, two cases with different image values; $a_3 = 1083.469$, $b_3 = 22.778$ and *threshold* = 1142.95 (A); $a_3 = 3102.805$, $b_3 = 22.792$ and *threshold* = 3174.44 (B).

Since the desired structure – the PCS – is located inside the kidney, the goal of binarization is to separate it from the kidney parenchyma. The threshold should be high enough to achieve this separation even at the cost of discontinuity and loss of some PCS parts. Therefore, this step increases the fragmentation of PCS. For 2 examples from the test set the threshold had to be manually increased to the kidney parenchyma due to over-segmentation.

4.3. Extraction of PCS parts from binarized ROI image BW, step 3

Within an ROI, after the preliminary segmentation, there often occur undesirable elements such as parts of spine or ribs. These elements have image intensity values at a comparable level as the structures with a contrast agent, which limits the possibility of indicating desirable structures based only on their values. In addition, discontinuities of the PCS after the preliminary segmentation preclude designation of the structure of interest by indicating one point located within the PCS. Fig. 5 explains the aforementioned issues (the outcome of initial segmentation – 2nd step).

Fig. 5. Examples of initial segmentation results (2nd step) with indication of occurring problems. Discontinuities of the PCS (light green area) (a); close proximity between the PCS and spine (red area) (b); close proximity between the PCS and rib (red area), with transversal view (c).

Taking into account both the unfavourable mutual spatial location of desired and undesired structures and discontinuities of the PCS, we propose a *Hybrid Level Set* (HLS) method [31] to overcome it. The HLS is a modified classical ChV algorithm. In [32], among others, an additional prior term of ellipsoidal shape is introduced to keep the segmentation result within a desired range and to prevent it from leakage to the surrounding structures. This method was used for kidney segmentation and therefore enabled to identify the structures located inside it. Individual structures were recognized as the PSC when at least their portions were located inside the kidney. Visualizations of kidney segmentation results are shown in Fig. 6, the PCS parts and undesired structures are marked. Examples of the PCS extraction results – E, together with indication of undesirable structures after binarization (BW, 2nd step), are presented in Fig. 7.

Fig. 6. Examples of results after step 3: kidney segmentation by the HLS (light grey color)
with indicated structures located outside the kidney (orange) and extracted PCS (dark blue).
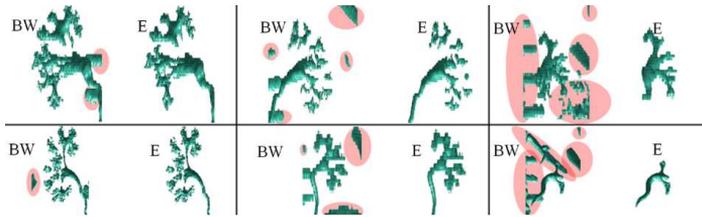

Fig. 7. Examples of the PCS extraction results. In each case the first picture indicate results
after the binarization (BW, 2nd step, undesired structures are highlighted bya red area),
and the second one – the results of extraction process (E, 3rd step).

4.4. Segmentation by Locally Adaptive Region Growing Algorithm, step 4

The LARG algorithm [14] was developed and implemented as a remedy to the problems associated with the lack of homogeneity of voxel intensity values due to an improper contrast propagation. The method described in [14] was adjusted to meet the requirements of PCS segmentation. In comparison with the method proposed in [14], dedicated to vessel segmentation, the regularization was introduced and the region growing conditions were adopted to the PCS segmentation. In the proposed solution the conditions for voxel selection are specified individually for each candidate basing on analysis of its surroundings. The criteria for addition of a next voxel to the previously chosen ones are based on such parameters as the maximum permissible difference between the candidate voxel value and the source of growth *tDiff* (4) and the minimum required intensity value for a voxel candidate *tMin* (5). For data segmentation enhanced by a contrast agent only the lower threshold is meaningful. An outline of LARG algorithm is presented in Fig. 8.

At the beginning the default values are set such as:

− the maximum permissible difference between the candidate voxel value and the source of growth:

$$defaultDiff = 2 \cdot b_3;$$ (2)

− the minimum required value:

$$defaultMin = threshold - 0.5 \cdot b_3.$$ (3)

The parameters *threshold* and b_3 are derived from the 2nd step of the proposed method. This ensures the stability of LARG process providing individual parameter values adequate to a particular image. Since the default parameter values are used only if there are not enough

already chosen voxels, a candidate should be examined basing on the restrictive conditions. Therefore, the *defaultMin* is set slightly below the preliminary segmentation threshold (2nd step) and *defaultDiff* − to $2 \cdot b_3$.

The initial seed set consists of voxels from the preliminary segmentation (3rd step, extraction). A set of seed values is sorted in the descending order to reduce the calculation time. Starting from the highest value (primary voxel) a region is growing in specified conditions. For each candidate the parameters *tDiff* and *tMin* are calculated basing only on previously added voxels. Starting from a radius equal to 1 (a cube (3,3,3)) it is checked whether there are enough previously classified voxels as objects to perform calculations. The size of the cube's neighbourhood is gradually increased until either the above condition is met or the maximum permissible size of the surroundings is achieved (max radius 5). The maximum radius is set to 5 to cover the standard diameter of the PCS structure (proper for the CT database transversal plane resolution). For each candidate the parameters are calculated as follows:

$$tDiff = \min\left(3 \cdot \mathrm{std}(\mathbf{tmpCT}) / radius, defaultDiff\right), \tag{4}$$

$$tMin = \max\left(\left(\mathrm{mean}(\mathbf{tmpCT}) - \mathrm{std}(\mathbf{tmpCT}) \cdot radius\right), defaultMin\right), \tag{5}$$

where **tmpCT** is a sub-image of a size defined by the radius and centre of the candidate's location.
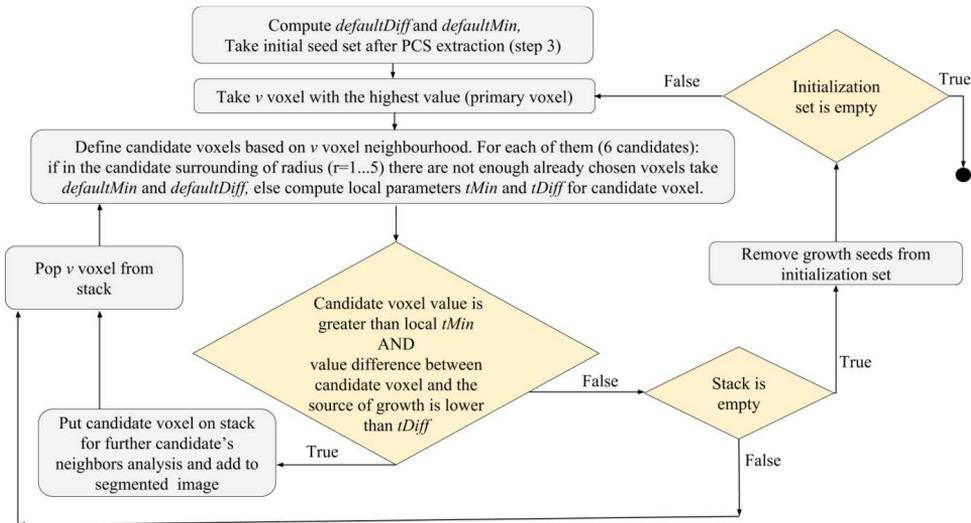

Fig. 8. The LARG – 4th step, an outline of algorithm.

If the number of previously added voxels is not required, the default values: *defaultDiff* (2) *defaultMin* (3) are used. If the parameters were calculated on the basis of a neighbourhood with a relatively large size, we put less confidence to their values (to a smaller extent they reflect the actual threshold for a candidate). Therefore, the regularization, that provides more stringent requirements for voxels located farther away from the candidate, is introduced. Also, the distant voxels may produce a greater variation of values and thus a relatively high *tDiff* (4). It has been observed during experiments that when disregarding the influence of neighbourhood size, the effectiveness of segmentation decreases. In some cases, this leads to overflow of the neighbouring structures. A candidate is added if its value is greater than *tMin* and the difference between the candidate and the source of grow is smaller than *tDiff* (4).

Starting from the primary voxel, voxels are put on the stack and subsequently removed after their growth ends. After emptying the whole stack, chosen voxels are removed from the initial seed set. The following step involves choosing a next primary voxel again. The operation is executed until the initial set of seeds is empty.

There are several advantages of defining region growing conditions locally, instead of using global parameters. First of all, it enables to add voxels whose intensity values descend towards the boundaries and along the vessel branch. Secondly, the voxels with lower intensity values, due to improper contrast propagation, may be also included. Furthermore, it prevents from leakage to unwanted structures. Finally, the continuity of the obtained structure is improved.

5. Results and discussion

Quantitative evaluation of the results from the test set (24 PCSs) was carried out on the basis of the following volumetric metrics derived from a confusion matrix (explanation in Table 1): accuracy, precision, sensitivity, specificity, over- and under-segmentation rate and DICE [33]. Apart from the volumetric metrics, also the distance measures: the maximum (*mHD*) [34] and the averaged Hausdorff distance (*aHD*) [35] were calculated. In general, the Hausdorff distance determines how far two subsets of a metric space are from each other. In particular, *mHD* denotes the maximum distance and *aHD* the average distance between points belonging to two subsets. The *aHD* distance is more suitable for image processing purposes as it decreases the impact of outliers [35]. Assuming that $\mathbf{P}_{mc}$ is a manual segmentation binary mask and p_{mc} is a set of its points' coordinates, $\mathbf{P}_{seg}$ – the binary mask obtained by the algorithm and p_{seg} – a set of its points' coordinates, the aforementioned measures (*mHD* and *aHD*) can be formulated as:

$$mHD = \max\left(h_1\left(\mathbf{P}_{mc}, \mathbf{P}_{seg}\right), h_1\left(\mathbf{P}_{seg}, \mathbf{P}_{mc}\right) \right), \tag{6}$$

$$aHD = \frac{h_2\left(\mathbf{P}_{mc}, \mathbf{P}_{seg}\right) + h_2\left(\mathbf{P}_{seg}, \mathbf{P}_{mc}\right)}{2}, \tag{7}$$

where: $h_1\left(\mathbf{P}_{mc}, \mathbf{P}_{seg}\right) = \max\limits_{p_{mc} \in \mathbf{P}_{mc}} \min\limits_{p_{seg} \in \mathbf{P}_{seg}} \left\| p_{mc} - p_{seg} \right\|$, $h_2\left(\mathbf{P}_{mc}, \mathbf{P}_{seg}\right) = \max\limits_{p_{mc} \in \mathbf{P}_{mc}} \min\limits_{p_{seg} \in \mathbf{P}_{seg}} \left\| p_{mc} - p_{seg} \right\|$ and $\left\| \ \right\|$ denotes the Euclidian distance.

In simple words, the Hausdorff distance specifies how far (meaning a spatial distance) the acquired result boundary is from the manual outline. In order to denote the Hausdorff distance, first corresponding points in two subsets (the resulting boundary points and the expected ground true) are identified. Next, the distances between them are calculated. *mHD* is the longest distance among them and *aHD* is the average one. In this paper, the scores (6), (7) are provided in mm, presenting the distance between the segmentation result and the manual outline with regard to the image resolution.

The proposed method was applied successfully in each of 24 PCSs belonging to the test set. Both visual and quantitative analyses were accomplished to evaluate its performance. A visual comparison between manual outlines and the segmentation results is shown in Fig. 9. The provided examples indicate a good quality of the segmentation together with a substantial improvement achieved when applying LARG after PCS extraction. The visual assessment is supported by DICE provided for each case. The results are presented in Table 2. After the 3rd step of the proposed method the mean DICE is above 84%. Further improvements are achieved when applying the 4th step – LARG (87%). The calculated average Hausdorff distance (0.46 ± 0.36 mm) is satisfactory from a medical point of view.

www.czasopisma.pan.pl   PAN   www.journals.pan.pl
POLSKA AKADEMIA NAUK

Metrol. Meas. Syst., Vol. 24 (2017), No. 1, pp. 3–18.

Table 1. A confusion matrix together with definitions of evaluation metrics: accuracy,
precision, sensitivity, specificity, DICE, over-segmentation and under-segmentation rates.

<table>
<tr><td rowspan="2"></td><td rowspan="2"></td><td colspan="2">Actual value</td><td colspan="2"></td><td colspan="2"></td></tr>
<tr><td>object</td><td>background</td></tr>
<tr><td rowspan="2">Predicted value</td><td>object</td><td>TP
(true positive)</td><td>FP
(false positive)</td><td colspan="2">Precision (Prec)
=TP/(TP+FP)</td><td colspan="2">Over-segmentaton (OS)
= FP/TP</td></tr>
<tr><td>background</td><td>FN
(false negative)</td><td>TN
(true negative)</td><td colspan="2">Accuracy (Acc)
=(TP+TN)/ (TP+FP+FN+TN)</td><td colspan="2">Under-segmentation (US)
= FN/TP</td></tr>
<tr><td></td><td></td><td>Sensitivity
(Sens)
=TP/(TP+FN)</td><td>Specificity
(Spec)
=TN/(FP+TN)</td><td colspan="2">DICE
=2TP/ (2TP+FN+FP)</td><td></td><td></td></tr>
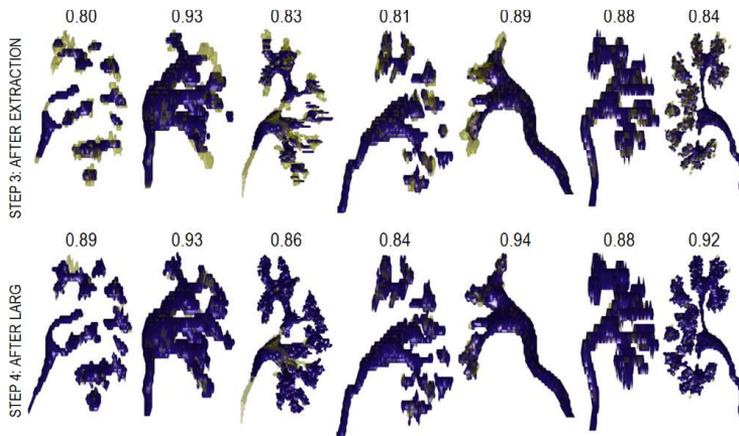</table>

Fig. 9. An example of visualization of the obtained results in a 3D view. The rows indicate two subsequent steps
of the proposed method, the columns represent 7 different cases.The DICE is presented for each case and each
step. The dark blue color represents the output of the algorithm while the green one – the manual outline.

Table 2. A summary of the results indicating improvements after applying
subsequent steps of the proposed method.

	Acc	Prec	Sens	Spec	OS	US	DICE	*mHD* [mm]	*aHD* [mm]
Step 3: **After Extraction**	0.997 ±0.004	0.945 ±0.063	0.778 ±0.114	0.999 ±0.001	0.070 ±0.079	0.315 ±0.219	0.845 ±0.068	9.96 ± 7.60	0.74 ± 0.49
Step 4: **After LARG**	0.997 ±0.004	0.872 ±0.107	0.883 ±0.074	0.998 ±0.003	0.166 ±0.165	0.141 ±0.101	0.871 ±0.060	7.73 ± 6.67	0.46 ± 0.36

The comparative assesment was also performed. The results of the proposed method were compared with those obtained with EdgeWave [9], ARG [8] and ChV [11] methods. For this purpose both ARG and ChV methods were implemented, while the EdgeWave tool was shared with its authors. The ARG was adapted for the 3D case. To assure the assumptions of this method (region compactness) each separatated part of the PCS was estimated and grown separately. Moreover, for the second RG step, only the lower threshold was taken into account, as for data with a contrast agent it is the only meaningful one. While using both thresholds the RG frequently stuck preventing the growth among the whole region. On the other hand, during estimation, it was essential to use both upper and lower thresholds to keep the standard

deviation estimates within reasonable limits. For ARG and ChV the initial mask for the segmentation contained one voxel per each separate region (up to 22 regions) making sure that it is not located on the border of the structure. Parameters for the ChV evolution were adopted using a 10-fold partition and were applied for all data. The same optimization procedure was used to designate ChV parameters, as for *A* and *B* in the 2nd step. In some cases, both ARG (3 cases) and ChV (5 cases) lead to leakage to the surrounding structures. These examples were excluded from evaluation of ChV and ARG methods. Regarding the above, the EdgeWave is a stable tool, however it involves manual indication of the segmentation threshold and therefore is prone to subjective setting. Regarding DICE, the Hausdorff distances, both *mHD* and *aHD*, accuracy, sensitivity, under- and over-segmentation rates, the proposed method provided the best results. However, in respect of precision and sensitivity, the EdgeWave algorithm outperformed other evaluated methods. A relatively low performance of ARG and ChV can be attributed to homogeneity assumption of the segmented region which is not true due to improper contrast propagation. The detailed results of the comparison are presented in Table 3 and Fig. 10.

Additionally, the issue of the efficiency of different methods has been addressed. Table 4 presents the overall computation time required by each method (the proposed method, EdgeWave [9], ChV [11] and ARG [8]) regarding the mean calculation time per case. The calculations were performed on a PC equipped with Intel® Core™ i3-3240 CPU, 3.4GHz, 2 cores with 32GB RAM. The input images for each method were derived from the 1st step (ROI designation) and therefore the computation time of this step is excluded from comparison. Such a comparison has several limitations that should be mentioned before making a direct reference to the values presented in Table 4. The ARG was implemented and modified by us and perhaps is different from the original one proposed by Authors regarding its computational efficiency. The mean calculation time required by the EdgeWave was obtained by dividing the total processing time by the number of cases. In addition to the algorithm worktime (approximately 2s per case), also the user-interaction time is taken into account. Moreover, the proposed method, ChV, ARG are Matlab prototypes, whereas the EdgeWave is an executable application. Taking into account the aforementioned issues, interpreting these values should be made with caution. Nevertheless, the EdgeWave was proven to be slightly faster (243.7 s) than the proposed method (250.7 s), whereas the ChV and ARG were much slower (644.3 s and 1047.6 s, respectively).

Table 3. The quantitative results of all evaluated methods; the mean values and the standard deviations are given; the results in bold indicate the best method in each category.
3 examples of ARG and 5 of ChV are excluded.

|	Acc	Prec	Sens	Spec	OS	US	DICE	*mHD* [mm]	*aHD* [mm]
Proposed method	**0.997** ±**0.004**	0.872 ±0.107	**0.883** ±**0.074**	0.998 ±0.003	**0.166** ±**0.165**	**0.141** ±**0.101**	**0.871** ±**0.060**	**7.73** ±**6.67**	**0.46** ±**0.36**
EdgeWave	0.994 ±0.007	0.893 ±0.177	0.644 ±0.216	**0.999** ±**0.003**	0.258 ±0.797	2.410 ±8.883	0.710 ±0.201	18.71 ±11.22	2.88 ±2.88
ChV	0.993 ±0.007	**0.901** ±**0.205**	0.647 ±0.277	**0.999** ±**0.003**	0.412 ±1.495	7.237 ±28.41	0.682 ±0.250	17.69 ±12.84	3.08 ±3.50
ARG	0.985 ±0.020	0.703 ±0.323	0.684 ±0.684	0.988 ±0.022	1.128 ±1.809	1.191 ±2.813	0.583 ±0.218	11.82 ±8.16	1.99 ±3.25
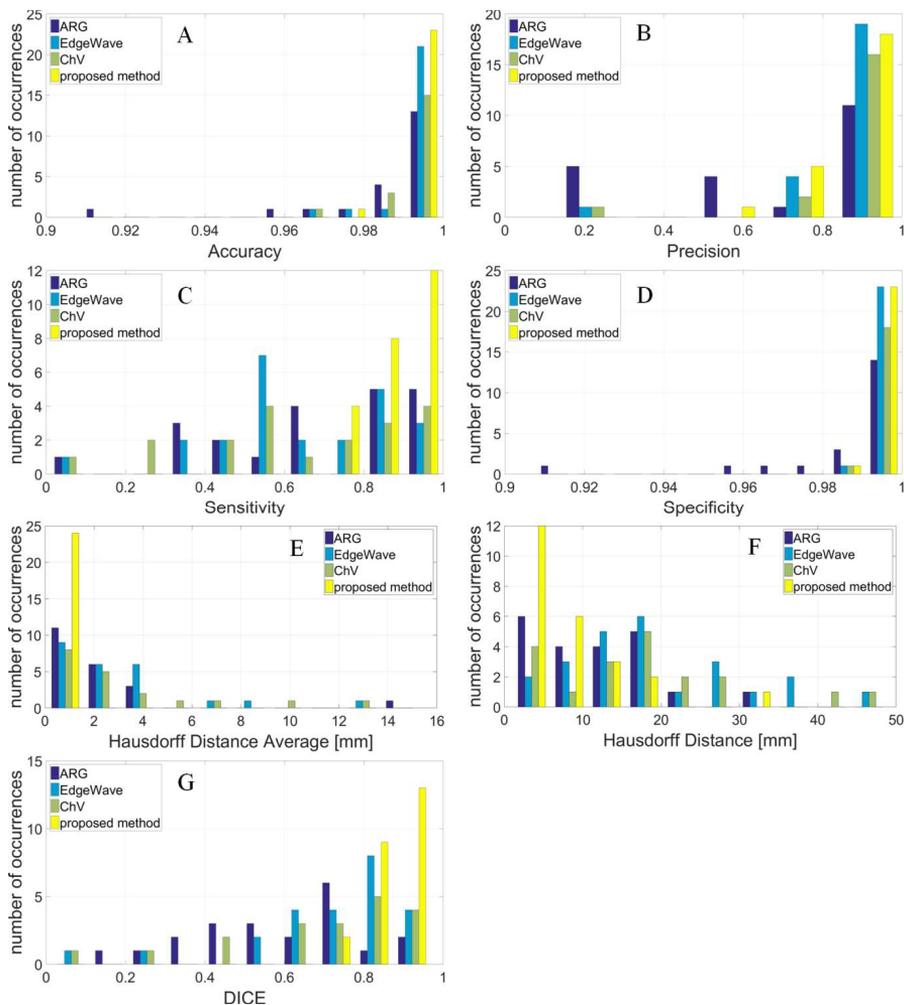
Fig. 10. Compared results of the proposed method, EdgeWave, ChV and ARG in respect of accuracy (A);
precision (B); sensitivity (C); specificity (D); *aHD* (E); *mHD* (F) and DICE (G).
Due to leakage 3 examples of ARG and 5 of ChV are excluded.

Table 4. The methods' efficiency (computation time) – comparison of the proposed method, EdgeWave [9],
ChV [11] and ARG [8] in respect of mean calculation time per case.

	Overall Computation Time [s]			
Proposed method	2nd step: 0.1	3rd step 3 (HLS): 141.6	4th step 4 (LARG): 82.1	**250.7**
EdgeWave				**243.7**
ChV				**644.3**
ARG				**1047.6**

6. Conclusion

This paper proposes a PCS segmentation methodology, consisting of an automatic threshold selection method for binarization (preliminary segmentation), extraction of the PCS parts from

the binarized ROI image by HLS and − finally − the LARG algorithm driven by the image intensity values.

The quantitative evaluation was performed for 24 PCSs belonging to the test set. A significant improvement was observed between the 3rd and 4th steps of the proposed method: the DICE coefficient increased (from 84% to 87%) and the average Hausdorff distance decreased (from 0.74 mm to 0.46 mm). A comparison with the other described methods (ARG, ChV, EdgeWave) revealed that the proposed method outperforms them in respect of DICE coefficient, the maximum and average Hausdorff distances, accuracy, sensitivity, under- and over-segmentation rates. Regarding precision and sensitivity the EdgeWave algorithm provided the best results.

Basing on the proposed PCS segmentations each anatomic detail of the kidney collecting system can be selectively displayed in 3D and used for planning various urological procedures. Additionally, measurements of different objects (the distance between tumour and the PCS, both shape and spatial configuration of individual components of the PCS, the calyceal pattern, the pelvicalyceal angle, a length and width of the lower calyceal neck) are facilitated. The presented PCS segmentation algorithm may constitute a considerable part of a support system for planning minimally invasive procedures. Moreover, the same segmentation framework may be applied to further medical applications. Preoperative imaging and operation planning are the first step in image-guided surgery and now we are planning to study this approach on clinical aspects, especially on a complication rate during PN, PCNL or RIRS. The next step will be development of intraoperative imaging and tracking systems.

Acknowledgements

References

[1]   *Statistical Bulletin of the Ministry of Health*. (2016). Centre for Health Information Systems.

[2]   Ukimura, O., *et al*. (2012). Three-dimensional reconstruction of renovascular-tumor anatomy to facilitate zero-ischemia partial nephrectomy. *European Urology*, 61(1), 211–217.

[3]   Ljungberg, B., *et al*. (2015). Guidelines on Renal Cell Carcinoma. *European Association of Urology*.

[4]   Zöllner, F.G., *et al*. (2012). Assessment of Kidney Volumes From MRI: Acquisition and Segmentation Techniques. *American Journal of Roentgenology,* 199 (5), 1060–1069.

[5]   Will, S., *et al*. (2014). Automated segmentation and volumetric analysis of renal cortex, medulla, and pelvis based on non-contrast-enhanced T1-and T2-weighted MR images. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 27(5), 445–454.

[6]   Yang, X., *et al*. (2015). Automatic Segmentation of Renal Compartments in DCE-MRI Images. *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015*, 3–11.

[7]   Li, X., *et al*. (2011). Renal cortex segmentation using optimal surface search with novel graph construction. In *Medical Image Computing and Computer-Assisted Intervention-MICCAI 2011,* 387–394.

[8]   Pohle, R., Toennies, K.D. (2001). A new approach for model-based adaptive region growing in medical image analysis. *Computer Analysis of Images and Patterns Journal*, 238–246.

[9]   Abiria, B., *et al.* (2014). Performance of an automated renal segmentation algorithm based on morphological erosion and connectivity. *Proc. of SPIE, Medical Imaging 2014: Computer-Aided Diagnosis*, 90352R−90352R−5.

[10] Sun, Y. *et al.* (2002). Kidney segmentation in MRI sequences using temporal dynamics. *Proc. of IEEE International Symposium on Biomedical Imaging*, 98−101.

[11] Chan, T.F., Vese, L.A. (2001). Active contours without edges. *IEEE Trans. on image processing*, 10(2), 266−277.

[12] Shehata, M., *et al.* (2015). A level set-based framework for 3D kidney segmentation from diffusion MR images. *Proc. of International Conference on Image Processing*, 4441−4445.

[13] Song, T., *et al.* (2008). Segmentation of 4D MR Renography Images Using Temporal Dynamics in a Level Set Framework. *Proc. of IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 37−40.

[14] Bugajska, K. *et al.* (2015). The renal vessel segmentation for facilitation of partial nephrectomy. *Proc. of IEEE, SPA: Signal Processing: Algorithms, Architectures, Arrangements and Applications*, 50−55.

[15] Yushkevich, P.A., *et al.* (2006). User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage*, 31(3), 1116−1128.

[16] Pankaj, R., *et al.* (2014.) Radiological Study of Variations in the pelvicalyceal system of kidney. *International Journal of Biological and Medical Research*, 5(3), 4336−4339.

[17] Sampaio, F.J.B., *et al.* (1988). Anatomic classification of the kidney collecting system for endourologic procedures. *Endourology,* 2(3), 247−251.

[18] Sampaio, F.J.B., *et al.* (1992). Inferior pole collecting system anatomy: its probable role in extracorporeal shock wave lithotripsy. *Journal of Urology*, 147, 322−324.

[19] Keeley, F.X. Jr., *et al.* (1999). Clearance of lower-pole stones following shock wave lithotripsy: effect of the infundibulopelvic angle. *European Urology*, 36, 371−375.

[20] Sumino, Y., *et al.* (2002). Predictors of lower pole renal stone clearance after extracorporeal shock wave lithotripsy. *Journal of Urology*, 168, 1344−1347.

[21] Lin, C.C., *et al.* (2008). Predictive factors of lower calyceal stone clearance after extracorporeal shockwave lithotripsy (ESWL): the impact of radiological anatomy. *Journal of the Chinese Medical Association,* 71(10), 496−501.

[22] Resorlu, B., *et al.* (2012). The Impact of Pelvicaliceal Anatomy on the Success of Retrograde Intrarenal Surgery in Patients With Lower Pole Renal Stones. *Urology*, 79(1), 61−66.

[23] Elbahnasy, A.M., *et al.* (1998). Lower caliceal stone clearance after shock wave lithotripsy or ureteroscopy: the impact of lower pole radiographic anatomy. *Journal of Urology*, 159, 676−682.

[24] Fong, Y.K., *et al.* (2004). Lower pole ratio: a new and accurate predictor of lower pole stone clearance after shockwave lithotripsy. *International Journal of Urology*, 11, 700−703.

[25] Türk, C., *et al.* (2015). Guidelines on Urolithiasis. *European Association of Urology*.

[26] Xu, Y., *et al.* (2016). The value of three-dimensional helical computed tomography for the retrograde flexible ureteronephroscopy in the treatment of lower pole calyx stones. *Chronic Diseases and Translational Medicine*, available online Apr. 06, 2016, DOI:10.1016/j.cdtm.2016.02.001.

[27] Sargin, S.Y., *et al.* (2014). The efficacy of radiographic anatomical measurement methods in predicting success after extracorporeal shockwave lithotripsy for lower pole kidney stones. *International Brazilian Journal Of Urology*, 40(3), 337−345.

[28] Sonka, M., *et al.* (2008). *Image processing, analysis, and machine vision.* 3rd ed. Cengage Learning.

[29] Byrd, R.H., *et al.* (1988). Approximate Solution of the Trust Region Problem by Minimization over Two Dimensional Subspaces. *Mathematical Programming*, 40, 247−263.

[30] Lagarias, J.C., *et al.* (1998). Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions. *SIAM Journal of Optimization*, 9(1), 112−147.

[31] Zhang, Y., *et al.* (2008). Medical image segmentation using new hybrid level-set method. *Proc. of BioMedical Visualization MEDIVIS'08 IEEE*, 71–76.

[32] Skalski, A., *et al.* (2017). Kidney segmentation in CT data using hybrid Level-Set Method with ellipsoidal shape constraints. *Metrol. Meas. Syst.*, 24(1), 101–112.

[33] Dice, L. (1945). Measures of the amount of ecologic association between species. *Ecology*, 26(3), 297–302.

[34] Huttenlocher, D.P., *et al.* (1993). Comparing images using the Hausdorff distance. *IEEE Tran. on Pattern Analysis and Machine Intelligence,* 15(9), 850–863.

[35] Dubuisson, M.P., *et al.* (1994). A modified Hausdorff distance for object matching. *Proc. of International Conference on Pattern Recognition*, 566–568.

[F1] *https://en.wikipedia.org/wiki/Kidney#/media/File:KidneyStructures_PioM.svg.*