Segmenting Kidney DCE-MRI Using 1st-Order Shape and 5th-Order Appearance Priors

Ni Liu*, Ahmed Soliman*, Georgy Gimel'farb[1], and Ayman El-Baz[2]**

[1] Department of Computer Science, University of Auckland, Auckland, New Zealand
aselba01@louisville.edu
[2] Bioengineering Department, University of Louisville, Louisville, KY, USA

Abstract. Kidney segmentation from dynamic contrast enhanced magnetic resonance images (DCE-MRI) is vital for computer-aided early assessment of kidney functions. To accurately extract kidneys in the presence of inherently inhomogeneous contrast deviations, we control an evolving geometric deformable boundary using specific prior models of kidney shape and visual appearance. Due to analytical estimates from the training data, these priors make the kidney segmentation fast and accurate, offering the prospect of clinical applications. Experiments with 50 DCE-MRI *in-vivo* data sets confirmed that the proposed approach outperforms three more conventional counterparts.

1 Introduction

Accurate delineation of kidney borders in dynamic perfusion images is essential for their automated analysis. However, it meets with challenges due to the need to maintain adequate spatial resolution while acquiring images very quickly to capture the transient first-pass transit event; varying signal intensities (gray levels) over the time course of agent transit; and motion-induced artefacts related to intrinsic pulsate effects, breathing, or transmitted effects from adjacent structures, such as the bowel. To address these challenges, De Priester et al. [1] obtained a kidney mask by thresholding the difference between averaged pre-contrast and early-enhanced images, removing objects smaller than a certain size, and smoothing the remaining object by morphological closing and manual processing. This approach was further expanded by Giele [2], obtaining the kidney contour as the morphological inner gradient, or difference between the initial and eroded mask. Koh et al. [3] segmented kidneys with a morphological 3D H-maxima transform, using rectangular masks and edge information to avoid prior knowledge or training. Nonetheless, similar intensities in the kidney and surrounding background tissues make segmentation by straightforward signal thresholding mostly inaccurate. To circumvent these drawbacks, the kidney and its internal structures were segmented by Chevaillier et al. [4] by using a semi-automated k-means clustering of pixel-wise temporal intensity curves to

* Shared first authorship (equal contribution)
** Corresponding author.

N. Navab et al. (Eds.): MICCAI 2015, Part I, LNCS 9349, pp. 77–84, 2015.
DOI: 10.1007/978-3-319-24553-9_10

classify the pixels. An automated wavelet-based k-means partitioning was applied also by Li et al. [5] for segmenting the kidneys and tested successfully on a small number of subjects (four volunteers and three patients).

However, the most accurate kidney segmentation is obtained by evolving a deformable boundary with due account of visual appearance and shape of the kidney or its element of interest. Sun et al. [6] guided the evolution in a DCE-MRI towards the kidney borders with a variational level set integrating a temporal smoothness constraint and spatial inter-pixel correlations and employed the Chan-Vese's level set framework [7] for segmenting the cortex and medulla. Later on, the accuracy of kidney segmentation using a variational level set was improved by incorporating statistical image data [8], or prior knowledge about the kidney's visual appearance (texture) and shape [9]. A level set-based guidance by Gloger et al. [10] also used the shape prior and Bayesian statistical inference for generating the shape probability maps. But the existing techniques usually take no account of inter-pixel dependencies governing visual appearance and shape of the kidney and its structural elements. As a result, even most efficient today's deformable boundaries remain too sensitive to fuzzy kidney borders and image noise, involving spatially variant contrast and offset deviations.

To partially overcome these limitations and handle spatial inhomogeneities and contrast variations in the DCE-MRI, a probabilistic shape prior is appended below with a new visual appearance prior based on a 5^{th}-order translation and contrast/offset invariant MGRF. In addition to the latter prior with analytical parameter estimation, we integrate the appearance and shape priors into a novel energy function resulting in a more effective stochastic guiding force for a level-set-based evolution of a deformable boundary.

2 Shape-Appearance Guided Deformable Boundary

Geometric (level-set based) deformable object-background boundaries are popular and powerful tools in segmenting medical images due to their flexible and parameter-independent evolving on the (x, y)-plane [7]. At each moment t the boundary is represented by a zero level $\phi_t(x, y) = 0$ of an implicit level-set function – a distance map $\phi_t(x, y)$ of the signed minimum Euclidean distances from every point (x, y) to the boundary. The distance is negative for the interior and positive for the exterior points, and the distance map evolves iteratively [7]: $\phi_{n+1}(x, y) = \phi_n(x, y) - \tau\nu_n(x, y)|\nabla\phi_n(x, y)|$ where n is an integer instant of time $t = n\tau$ measured with a step $\tau > 0$; $\nu_n(x, y)$ is a speed function guiding the evolution, and $\nabla\phi_n = \left[\frac{\partial\phi_n}{\partial x}, \frac{\partial\phi_n}{\partial y}\right]$ denotes the spatial gradient of the distance map. Conventional speed functions accounting for image intensities, object edges, gradient vector flow, *etc.*, are unsuccessful on noisy images with low object-background intensity gradients, such as kidney DCE-MRI. Since kidneys have well-defined shapes and distinct visual appearances, our stochastic speed function combines both the kidney's shape and appearance priors to increase the segmentation accuracy.

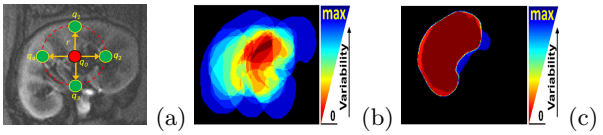
Adaptive 1st-order Kidney Shape Prior. Let $\mathbb{R} = \{(x, y) : x = 0, \ldots, X-1; y = 0, \ldots, Y-1\}$ be a finite 2D lattice supporting greyscale kidney images and their region maps. Our shape prior is modeled with a spatially variant independent random field (IRF) of binary region labels ($\mathbb{L} = \{1(\text{kidney}), 0(\text{background})\}$), such that maps $\mathbf{m} : \mathbb{R} \to \mathbb{L}$ of regions in the DCE-MRI with labels $\mathbf{m} = [m(x, y) : (x, y) \in \mathbb{R}; m(x, y) \in \mathbb{L}]$ are sampled from. To learn the model from a set of training DCE-MRI, $\mathbf{m}_n^\circ$; $n = 1, \ldots, N$, from different subjects, geometric deviations between their kidney shapes are reduced by mutual alignment using non-rigid B-spline-based deformations [11]. After a medical expert delineates kidneys borders in the training images, the shape prior is specified by an empirical joint probability disturbution, $P_{\text{sh}}(\mathbf{m}) = \prod_{(x,y)\in\mathbb{R}} p_{\text{sh}:x,y}(m(x, y))$. Here, each label $m(x, y) = 1$ or 0, and $p_{\text{sh}:x,y}(1) = \frac{1}{n}\sum_{n=1}^{N} m_n^\circ(x, y)$ is the empirical pixel-wise kidney probability for a stack of the co-aligned training images $\mathbf{m}_n^\circ$.

Learnable 5th-order MGRF Appearance Prior. Let $\mathbb{Q} = \{0, \ldots, Q-1\}$ denote a finite set of signals (intensities, or grey levels), in the DCE-MRI, $g : \mathbb{R} \to \mathbb{Q}$, with signals $\mathbf{g} = [g(x, y) : (x, y) \in \mathbb{R}]$. Probabilistic signal dependencies in the images are quantified with an interaction graph, $\Gamma = (\mathbb{R}, \mathbb{A})$, with nodes at the lattice sites (pixels or voxels), $(x, y) \in \mathbb{R}$, and edges, or arcs $((x, y), (x', y')) \in \mathbb{A} \subseteq \mathbb{R}^2$ connecting interdependent, or interacting pairs of the nodes, called neighbours. An MGRF of images is defined by a Gibbs probability distribution (GPD), $\mathbf{P} = \left[P(\mathbf{g}) : \mathbf{g} \in \mathbb{Q}^{|\mathbb{R}|}; \sum_{\mathbf{g}\in\mathbb{Q}^{|\mathbb{R}|}} P(\mathbf{g}) = 1\right]$, factored over a set $\mathbb{C}$ of cliques in Γ supporting non-constant factors, logarithms of which are Gibbs potentials (functions of clique-wise signals) [12].

Let a translation-invariant K-order interaction structure on $\mathbb{R}$ be represented by A, $A \geq 1$, families, $\mathbb{C}_a$; $a = 1, \ldots, A$, of K-order cliques, $\mathbf{c}_{a:x,y} \in \mathbb{C}_a$, of the same shape and size. Every clique is associated with a certain pixel, $(x, y) \in \mathbb{R}$, acting as the origin, and supports the same K-variate scalar potential function, $V_a : \mathbb{Q}^K \to (-\infty, \infty)$, depending only on specific ordinal relations between the clique-wise signals. The GPD for this translation- and contrast/offset-invariant MGRF is $P_K(\mathbf{g}) = \frac{1}{Z}\psi(\mathbf{g}) \exp\left(-E_K(\mathbf{g})\right)$ where $E_K(\mathbf{g}) = \sum_{a=1}^{A} E_{K:a}(\mathbf{g})$ and $E_{K:a}(\mathbf{g}) = \sum_{\mathbf{c}_{a:x,y}\in\mathbb{C}_a} V_{K:a}\left(g(x', y') : (x', y') \in \mathbf{c}_{a:x,y}\right)$ denote the Gibbs energy for all the clique families and for each individual family, respectively; $\psi(\mathbf{g})$ is a core distribution (if all the Gibbs potentials are equal to zero), and the partition function Z normalizes the GPD over the parent population of images.

Because kidneys appearance in the DCE-MRI is mostly the same under locally varying contrast, ordinal image descriptors, such as, e.g., local binary (LBP) or ternary patterns (LTP) [13] are more reasonable, than signal co-occurrences. Below, a motivated by the LBP/LTPs new class of high-order MGRFs, introduced in [14,15], is applied to model the kidney appearance priors. Its learning framework generalizes the analytical 2nd-order one in [16].

Given a training image $\mathbf{g}^\circ$, the maximum likelihood estimates (MLE) of the Gibbs potentials for the above generic K-order MGRF model with the simplest core, being an independent random field (IRF) of signals, can be approximated



(a) (b) (c)

Fig. 1. The 5th-order clique (a): signals q_0,q_1,..., q_4 are at the central pixel and its four central-symmetric neighbours at the radial distance r, respectively, and the color-coded shape prior before (b) and after (c) the nonrigid registration.

by generalizing the analytical approximation in [16] of the MLEs of potentials for a generic 2nd-order MGRF:

$$V_{K:a}(\beta) = \frac{F_{K:a:\mathrm{core}}(\beta) - F_{K:a}(\beta|\mathbf{g}^\circ)}{F_{K:a:\mathrm{core}}(\beta)(1-F_{K:a:\mathrm{core}}(\beta))};\ a = 1,\ldots,A;\ \beta \in \mathbb{B}_K$$

where β denotes a numerical code (value) of a particular K-order relation between the K signals on the clique; $\mathbb{B}_K$ is a set of these codes for all K-order signal co-occurrences; $F_{K:a}(\mathbf{g}^\circ)$ is an empirical marginal probability of the relation β; $\beta \in \mathbb{B}_K$, over the K-order clique family $\mathbb{C}_{K:a}$ for the training image $\mathbf{g}^\circ$, and $F_{K:a:\mathrm{core}}(\beta)$ is the like probability for the core distribution.

Algorithm 1. Learning the 5th-order MGRF appearance models.

1. Given a training DCE-MRI $\mathbf{g}^\circ$, find the empirical kidney ($l = 1$) and background ($l = 0$) probability distributions, $\mathbf{F}_{l:5:r}(\mathbf{g}^\circ) = [F_{l:5:r}(\beta|\mathbf{g}^\circ) :\ \beta \in \mathbb{B}]$ of the LBP-based descriptors for different clique sizes $r \in \{1,\ldots,r_{\max}\}$ where the top size $r_{\max} = 10$ in our experiments below.
2. Find the empirical distributions $\mathbf{F}_{5:r:\mathrm{core}} = [F_{5:r:\mathrm{core}}(\beta) :\ \beta \in \mathbb{B}]$ of the same descriptors for the core IRF $\psi(\mathbf{g})$, e.g., for an image, sampled from the core.
3. Find the approximate potentials' MLE $V_{l:5:r}(\beta) = \frac{F_{5:r:\mathrm{core}}(\beta) - F_{l:5:r}(\beta|\mathbf{g}^\circ)}{F_{5:r:\mathrm{core}}(\beta)\cdot(1-F_{5:r:\mathrm{core}}(\beta))}$.
4. Compute partial Gibbs energies of the descriptors for equal and all other clique-wise signals over the training image for the clique sizes $r = 1, 2, \ldots, 10$ to choose the size ρ_l, making both the energies the closest one to another.

To demonstrate advantages of capturing constrained high-order signal relations, the kidney and background appearances are quantified below, for simplicity, by pixel-wise Gibbs energies for two 5th-order translation- and contrast/offset-invariant MGRFs, each with a single family of fixed-shape central-symmetric cliques $\mathbf{c}_{x,y} = \{(x,y),(x \pm r, y),(x, y \pm r)\}$, shown in Fig. 1. Their potentials and radial distances, r, between the peripheral and central lattice cites are learned from the training image. Each LBP-based clique descriptor accounts for binary ordinal relations between grey values in the central, q_0, and four peripheral pixels, $q_1,\ldots,q_4$, i.e., $b(q_i, q_0) = 0$ if $q_i = q_0$ and 1 otherwise; $i = 1,\ldots,4$, giving 16 codes per clique. To further detail the clique-wise signal relations, our descriptor accounts also for the number τ; $\tau = 1,\ldots,5$, of signals, being equal to or greater than their mean, $\widehat{q} = \frac{1}{5}(q_0 + q_1 + \ldots + q_4)$, making in total up to 80 distinct codes β; $\beta \in \mathbb{B} = \{0,\ldots,79\}$, per clique.

Algorithm 2. Kidney segmentation with a geometric deformable boundary.

1. Equalize a DCE-MRI $\mathbf{g}$ using its cumulative probability distribution of signals.
2. Select among the training images a reference image, $\mathbf{g}_{\text{ref}}$, having the minimum Kullback-Leibler divergence from the equalized DCE-MRI $\mathbf{g}$.
3. Align $\mathbf{g}$ with $\mathbf{g}_{\text{ref}}$ using the non-rigid deformations [11].
4. Evolve a deformable boundary with the speed function depending on the appearance and shape priors: $\nu(x,y) = \kappa\theta(x,y)$ where κ is the mean curvature and $\theta(x,y)$ defines the pixel-wise evolution magnitude and direction: $\theta(x,y) = -P_{1:x,y}$ if $P_{1:x,y} > P_{0:x,y}$ and $\theta(x,y) = P_{0:x,y}$ otherwise where $P_{l:x,y} = \left(\frac{E_{l:5:\rho_l:x,y}(\mathbf{g})}{E_{0:5:\rho_0:x,y}(\mathbf{g})+E_{1:5:\rho_1:x,y}(\mathbf{g})}\right) p_{\text{sh}:l}(x,y)$; $l \in \mathbb{L} = \{1,0\}$.
5. Transfer the final boundary to the initial (non-aligned) DCE-MRI by reversing the non-rigid deformations, which have been estimated for the alignment.

Algorithms 1 and 2 outline learning the kidney and background appearance priors and basic steps of segmenting the kidney DCE-MRI with these and shape priors, respectively. The pixel-wise energies, $E_{l:5:\rho_l:x,y}(\mathbf{g})$; $l \in \mathbb{L}$, summing the learned potentials for the five cliques, containing the pixel (x,y), characterize in Algorithm 2 to what extent that pixel of the image $\mathbf{g}$ can be assigned to the background or kidney in accord with their appearance priors.

3 Experimental Validation and Conclusions

The proposed approach has been tested on the 3D (2D + time) DCE-MRI data sets collected from 50 subjects (35 men and 15 women from 10 to 56 years old (the mean age of $31_{\pm 11}$ years). The temporal sampling was adequate to characterize the transit of the clinical gadoteric acid contrast agent (Dotarem 0.5 *mmol/mL*; Guerbet, France), injected at the rate of 3-4 *ml/sec*, at the dose of 0.2 *ml/kgBW*. The gradient-echo T1 imaging employed a 1.5 T MRI scanner Signa Horizon LX Echo speed (GE Medical Systems, USA) with a phased-array torso surface coil; slice thickness: 5 *mm*; TR = 30-40 *msec*; TE = 2-3 *msec*; flip angle 70^o; FOV = 38×38 cm^2, and matrix size = 256×160. To obtain representative sampling to characterize perfusion for each patient, a single coronal image section was used at the level of the renal hilum of the transplanted kidney. Approximately 80 repeated temporal frames were obtained at 3 *sec* intervals.

Basic steps of the proposed level set-based segmentation are shown in Fig. 2, which also compares, together with Table 1 and Figs. 3 and 4, our segmentation accuracy with the vector level set (VLS) algorithm by Abdelmunim and Farag [9] and the parametric kernel graph cut (PKGC) with morphological and connectivity post-analysis by Salah et al. [17]. Differences between the mean Dice similarity coefficients (DSC) for our and other algorithms in Table 1 are statistically significant by the paired t-test. In total, embedding the proposed simple 5^{th}-order MGRF appearance model, together with our earlier shape prior, into

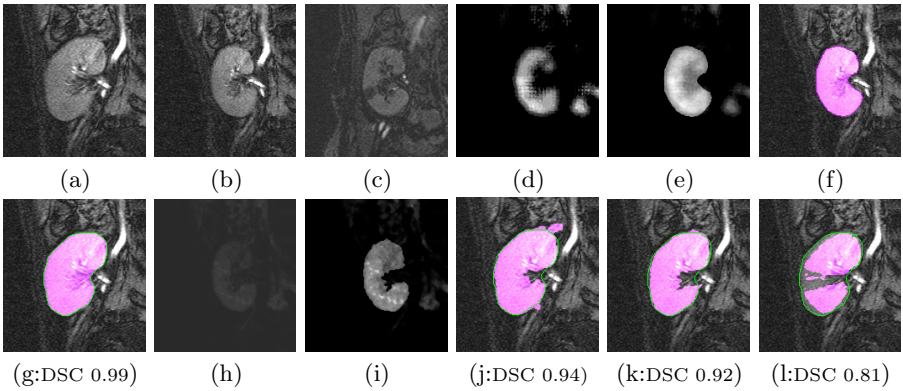


Fig. 2. The image (a) to be segmented; its alignment (b) [11] to a selected training reference image (c); the pixel-wise Gibbs energy (d) of our 5^{th}-order MGRF model; the total energy (e) after fusing that energy with the shape prior; the segmented aligned kidney (f); the final segmentation (g) after reversing to the original image (a); the pixel-wise Gibbs energy (h) for the 2^{nd}-order MGRF appearance model [16]; the total energy (i) after fusing with the shape prior; the PKGC segmentation (j) [17]; the segmentation (k) with the 2^{nd}-order MGRF prior; and the VLS segmentation (l) [9]. The ground truth is in green.

Table 1. Accuracy of our level-set based kidney segmentation with the 5^{th}- or only 2^{nd}-order appearance prior w.r.t. the vector level set (VLS) [9] and parametric kernel graph cut (PKGC) [17] algorithms on the 50 data sets.

	5^{th}-order prior	2^{nd}-order prior	VLS	PKGC
DSC: mean$_{\pm\text{st.dev.}}$	$0.99_{\pm 0.02}$	$0.91_{\pm 0.03}$	$0.90_{\pm 0.08}$	$0.82_{\pm 0.18}$
p-value		$\leq 10^{-4}$	$\leq 10^{-4}$	$\leq 10^{-4}$

the speed function of the level-set-guided boundary evolution results in more accurate segmentation of complex 3D (2D + time) kidney DCE-MRI.

These qualitative and quantitative comparisons use the ground truth obtained manually by an MRI expert. To highlight advantages of the proposed 5^{th}-order MGRF priors, the pixel-wise Gibbs energies were compared in Fig. 2,h–k, and Table 1 with the like energies for the 2^{nd}-order priors [16]. Obviously, the latter priors describe the object appearance less accurately. The considerably more distinct 5^{th}-order pixel-wise Gibbs energies for the kidney and background (see Figs. 2,d,e,h,i) ensure better guidance of the evolving level-sets-based boundary.

Our present mixed-code implementation (Matlab and C++) on a T7500 workstation (Intel quad-core processor; 3.33 GHz each with 48 GB of memory) takes about 125 ± 10 *sec* for segmenting 79 DCE-MRI time series images, each of size 256×256 pixels.

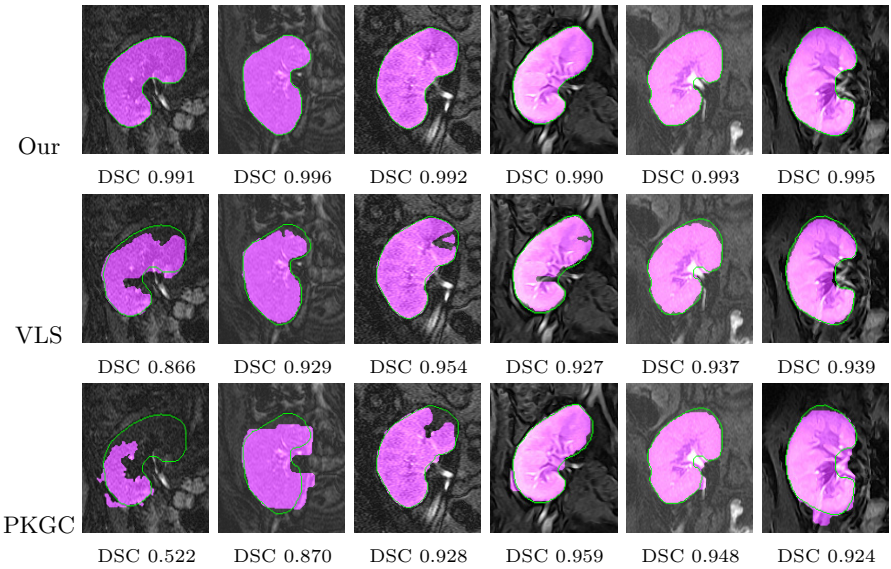


Fig. 3. Comparison to the VLS [9] and PKGC [17] (green – the ground truth).

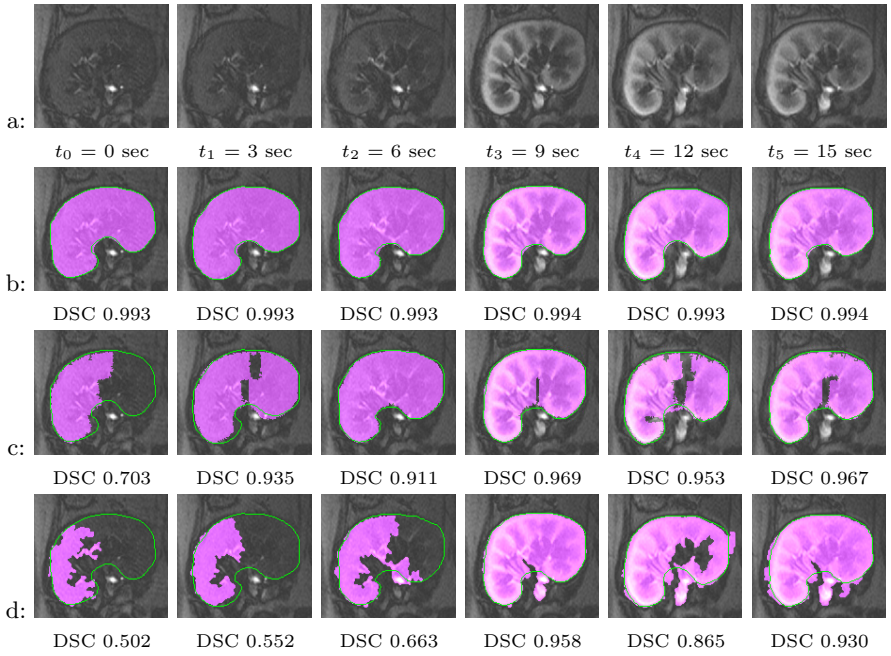


Fig. 4. Original first 6 images (a) of one of the DCE-MRI sequences and our segmentation (b) w.r.t. the VLS [9] (c) and PKGC [17] (d).

References

1. de Priester, J., Kessels, A., Giele, E., den Boer, J., Christiaans, M., Hasman, A., van Engelshoven, J.: MR renography by semiautomated image analysis: Performance in renal transplant recipients. J. M.R.I. 14(2), 134–140 (2001)
2. Giele, E., De Priester, J., Blom, J., den Boer, J., Van Engelshoven, J., Hasman, A., Geerlings, M.: Movement correction of the kidney in dynamic MRI scans using FFT phase difference movement detection. J. M.R.I. 14(6), 741–749 (2001)
3. Koh, H., Shen, W., Shuter, B., Kassim, A.A.: Segmentation of kidney cortex in MRI studies using a constrained morphological 3D H-maxima transform. In: Proc. ICARCV, pp. 1–5 (2006)
4. Chevaillier, B., Ponvianne, Y., Collette, J.L., Mandry, D., Claudon, M., Pietquin, O.: Functional semi-automated segmentation of renal DCE-MRI sequences. In: Proc. ICASSP, pp. 525–528 (2008)
5. Li, S., Zöllner, F.G., Merrem, A.D., Peng, Y., Roervik, J., Lundervold, A., Schad, L.R.: Wavelet-based segmentation of renal compartments in DCE-MRI of human kidney: initial results in patients and healthy volunteers. Comput. Med. Imaging Graph. 36(2), 108–118 (2012)
6. Sun, Y., Moura, J., Yang, D., Ye, Q., Ho, C.: Kidney segmentation in MRI sequences using temporal dynamics. In: Proc. ISBI, pp. 98–101 (2002)
7. Chan, T.F., Vese, L.A.: Active contours without edges. IEEE Trans. Image Process. 10(2), 266–277 (2001)
8. Wu, C.H., Sun, Y.N.: Segmentation of kidney from ultrasound B-mode images with texture-based classification. Comput. Meth. Prog. Biomed. 84(2), 114–123 (2006)
9. El Munim, H.E.A., Farag, A.A.: Curve/surface representation and evolution using vector level sets with application to the shape-based segmentation problem. IEEE Trans. Pattern Anal. Mach. Intell. 29(6), 945–958 (2007)
10. Gloger, O., Tonnies, K.D., Liebscher, V., Kugelmann, B., Laqua, R., Volzke, H.: Prior shape level set segmentation on multistep generated probability maps of MR datasets for fully automatic kidney parenchyma volumetry. IEEE Trans. Med. Imaging 31(2), 312–325 (2012)
11. Rueckert, D., Sonoda, L.I., Hayes, C., Hill, D.L., Leach, M.O., Hawkes, D.J.: Non-rigid registration using free-form deformations: application to breast MR images. IEEE Trans. Med. Imaging 18(8), 712–721 (1999)
12. Blake, A., Kohli, P., Rother, C.: Markov random fields for vision and image processing. MIT Press (2011)
13. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. Mach. Intell. 24(7), 971–987 (2002)
14. Liu, N., Gimel'farb, G., Delmas, P., Chan, Y.H.: Contrast/offset-invariant generic low-order MGRF models of uniform textures. In: AIP Conf. Proc., vol. 1559(1), pp. 145–154 (2013)
15. Liu, N., Gimel'farb, G., Delmas, P.: High-order MGRF models for contrast/offset invariant texture retrieval. In: Proc. Int. Conf. Image Vision Computing, pp. 96–101. ACM, New Zealand (2014)
16. Gimel'farb, G., Zhou, D.: Texture analysis by accurate identification of a generic Markov–Gibbs model. In: Gimel'farb, G., Zhou, D. (eds.) Applied Pattern Recognition. SCI, vol. 91, pp. 221–245. Springer, Heidelberg (2008)
17. Salah, M.B., Mitiche, A., Ayed, I.B.: Multiregion image segmentation by parametric kernel graph cuts. IEEE Trans. Image Process. 20(2), 545–557 (2011)