

3D Kidney Segmentation from CT Images Using a Level Set Approach Guided by a Novel Stochastic Speed Function

Fahmi Khalifa¹, Ahmed Elnakib¹, Garth M. Beache², Georgy Gimel'farb³,
Mohamed Abo El-Ghar⁴, Rosemary Ouseph⁵, Guela Sokhadze¹,
Samantha Manning¹, Patrick McClure¹, and Ayman El-Baz^{1,*}

¹ BioImaging Laboratory, Bioengineering Department, University of Louisville,
Louisville, KY, USA

² Diagnostic Radiology Department, University of Louisville, Louisville, KY, USA

³ Department of Computer Science, University of Auckland, Auckland, New Zealand

⁴ Urology and Nephrology Department, University of Mansoura, Mansoura, Egypt

⁵ Department of Medicine, University of Louisville, Louisville, KY, USA

Abstract. Kidney segmentation is a key step in developing any non-invasive computer-aided diagnosis (CAD) system for early detection of acute renal rejection. This paper describes a new 3-D segmentation approach for the kidney from computed tomography (CT) images. The kidney borders are segmented from the surrounding abdominal tissues with a geometric deformable model guided by a special stochastic speed relationship. The latter accounts for a shape prior and appearance features in terms of voxel-wise image intensities and their pair-wise spatial interactions integrated into a two-level joint Markov-Gibbs random field (MGRF) model of the kidney and its background. The segmentation approach was evaluated on 21 CT data sets with available manual expert segmentation. The performance evaluation based on the receiver operating characteristic (ROC) and Dice similarity coefficient (DSC) between manually drawn and automatically segmented contours confirm the robustness and accuracy of the proposed segmentation approach.

1 Introduction

Kidney segmentation from abdominal CT images is an essential step for many high-level processing tasks, such as localization of pathology, radiotherapy planning, and computer-integrated surgery. However, due to image noise, acquisition artifacts, gray level inhomogeneities, and similar visual appearances of adjacent structures, accurate segmentation of the kidney still remains a challenge [1]. In recent years, many automated and semi-automated approaches have been developed to address these challenges. In particular, Pohle and Toennies [2] developed an automatic region-growing algorithm for segmenting anatomical structures. Their approach estimated the homogeneity criterion from the characteristics of the images to be segmented. However, due to gray level similarities

* Corresponding author:- Tel.: (502)-852-5092, E-mail: aselba01@louisville.edu

between the kidney and the abdominal tissues, as well as sensitivity to initial seed point locations, region growing-based approaches are not effective for CT kidney segmentation.

Wang et al. [3] proposed a constrained optimization deformable contour method in which the degree of contour interior homogeneity is computed as an extra constraint within the level set energy minimization framework. However, deformable models fail in the case of noise, spurious edges, poor image resolution, diffused boundaries, or occluded shapes if they do not include a *priori* knowledge to constrain the contour deformation within an admissible range. Tsagaan et al. [4] proposed an automatic deformable model-based approach to segment the kidney from CT images. Their method is based on a non-uniform rational B-spline surface representation and statistical shape information of the kidney, which is then incorporated into the objective function as an additional energy term. Their results, evaluated on 33 CT images, seem reasonable in some cases, but differ markedly from the manually segmented results in others.

Recently, Huang et al. [5] proposed a multiphase level set approach with multi-dynamic shape models to segment the kidneys on abdominal CT images. Spiegel et al. [6] proposed a kidney segmentation framework based on the active shape model (ASM) that was combined with a curvature-based non-rigid registration approach to solve the point correspondence problem of the training data. In general, knowledge-based approaches are computationally intensive, and their accuracy depends on the size of the training data. Freiman et al. [7] proposed a model-based kidney segmentation approach from CT images based on maximum a posteriori-Markov random field (MAP-MRF) estimation of the current image. The MAP-MRF estimation is obtained by using the graph min-cut technique. Campadelli et al. [8] proposed an automatic, gray-level based segmentation framework based on a multiplanar fast marching method. Their segmentation performance was evaluated based only on visual assessment.

To overcome the aforementioned limitations, we introduce a 3-D extension of our previous 2-D stochastic guiding force presented in [9] to guide the evolution of a 3-D geometric deformable model to extract the kidney region from CT images. The proposed 3-D stochastic guiding force accounts for a 3-D shape prior, 1st-order intensity model, and a 3-D 2nd-order spatial interaction model between the kidney voxels and its background.

2 The Proposed Level Set-Based Segmentation Approach

In recent years, level set-based deformable models have been applied to medical image segmentation with considerable success because of the flexibility of the evolving boundary and the lack of need for parameterizations. The object-background boundary at each moment t is represented by a zero level $\phi_t(x, y, z) = 0$ of an implicit level set function, namely a distance map $\phi_t(x, y, z)$ of the signed minimum Euclidean distances from every voxel to the boundary (negative for interior and positive for exterior voxels). The distance map is evolved iteratively with the evolution being guided by a speed function $V_n(x, y, z)$ [10]:

$$\phi_{n+1}(x, y, z) = \phi_n(x, y, z) - \tau V_n(x, y, z) |\nabla \phi_n(x, y, z)| \quad (1)$$

where n indicates the time instant $t = n\tau$ (taken with a step $\tau > 0$), and $\nabla \phi_n = [\frac{\partial \phi_n}{\partial x}, \frac{\partial \phi_n}{\partial y}, \frac{\partial \phi_n}{\partial z}]$ is the gradient of $\phi_n(x, y, z)$. Conventional speed functions accounting for image intensities, object edges, gradient vector flow, etc., are unsuccessful on very noisy images with low object-background intensity gradients. The results are improved by involving shape priors (e.g., [5,6,11]). To obtain more accurate results, our stochastic speed function accounts for both the shape prior and appearance features associated with image intensities and their spatial interactions integrated into a 3-D two-level joint MGRF model.

Shape-Appearance Guided Evolution: Let $\mathbf{Q} = \{0, 1, \dots, Q-1\}$ and $\mathbf{L} = \{0, 1\}$ denote a finite set of integer gray values and a binary set of object (“1”) and background (“0”) labels, respectively. Let \mathbf{R} denote a 3-D arithmetic (x, y, z) -lattice that supports a given grayscale CT data $\mathbf{g} : \mathbf{R} \rightarrow \mathbf{Q}$ to be segmented and its goal binary “object - background” region map $\mathbf{m} : \mathbf{R} \rightarrow \mathbf{L}$. The image \mathbf{g} , being co-aligned to a shape prior, and its map \mathbf{m} are described with a joint probability model $P(\mathbf{g}, \mathbf{m}) = P(\mathbf{g}|\mathbf{m})P(\mathbf{m})$ combining a 2^{nd} -order MGRF $P(\mathbf{m})$ of region labels with the shape prior and a conditionally independent random field $P(\mathbf{g}|\mathbf{m})$ of image intensities given the map. The map model $P(\mathbf{m}) = P_s(\mathbf{m})P_h(\mathbf{m})$ has two parts: (i) a shape prior probability $P_s(\mathbf{m})$ and (ii) a 2^{nd} order MGRF model $P_h(\mathbf{m})$ of a spatially homogeneous region map \mathbf{m} for the image \mathbf{g} .

Conditional Intensity Model: To build an initial map \mathbf{m} , we need to estimate the marginal intensities distribution for the object (i.e., kidney) and background. The empirical gray level distribution, $P(\mathbf{g}) = \prod_{(x,y,z) \in \mathbf{R}} P_{\text{mix}}(g_{x,y,z})$, is separated into object and background components, $(p(q|\lambda) : q \in \mathbf{Q}; \lambda \in \mathbf{L})$, by close approximation with a linear combination of discrete Gaussians (LCDG)¹, a modified version of our previous linear combination of continuous Gaussians probabilistic model [12]. This approximation adapts the segmentation to changing appearance, such as non-linear intensity variations caused by patient weight and data acquisition system (scanner type and scanning parameters). The LCDG separates each factor of the empirical gray level distribution more accurately than conventional mixtures of only positive Gaussians, thus yielding a better initial region map formed by voxel-wise classification of the CT images gray values.

Spatial Voxel Interaction Model: To smooth the evolution of the level set and get more accurate segmentation, spatially homogeneous interactions between the region labels are modeled with a generic MGRF of a region map that accounts only for 3-D voxelwise interactions between each region label and its characteristic neighborhood. For simplicity, we restrict the interaction structure to the nearest voxel 26-neighbors (Fig. 1). By symmetry considerations, we assume that the potentials are independent of the relative orientation of each voxel pair and depend only on whether the labels are equal or not. Under these

¹ For complete details, please see: <https://louisville.edu/speed/bioengineering/faculty/bioengineering-full/dr-ayman-el-baz/supplemental-materials>

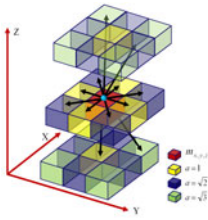


Fig. 1. Pairwise voxel interaction for the 26 neighbors in a 3-D MGRF image model

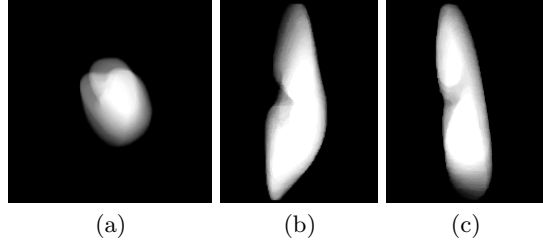


Fig. 2. 3D kidney shape model projected onto 2D axial (a), coronal (b), and sagittal (c) planes for visualization

restrictions, it is a 3-D extension of the conventional auto-binomial, or Potts model, differing only in that the potentials are estimated analytically.

The 26-neighborhood has three types of symmetric pairwise interactions specified by the absolute distance a between two voxels in the same and adjacent CT slices ($a = 1, \sqrt{2}$, and $\sqrt{3}$, respectively): (i) the closest pairs with the inter-voxel coordinate offsets $\mathbf{N}_1 = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$; (ii) the farther diagonal pairs with the offsets $\mathbf{N}_{\sqrt{2}} = \{(0, 1, \pm 1), (1, 0, \pm 1), (1, \pm 1, 0)\}$; and (iii) the farthest diagonal pairs with the offsets $\mathbf{N}_{\sqrt{3}} = \{(1, \pm 1, \pm 1)\}$. The potentials of each type are bi-valued because only the coincidence of the labels is taken into account: $\mathbf{V}_a = \{v_{a,\text{eq}}; v_{a,\text{ne}}\}$ where $v_{a,\text{eq}} = \gamma$ if $l = l'$ and $v_{a,\text{ne}} = -\gamma$ if $l \neq l'$; $a \in \mathbf{A} = \{1, \sqrt{2}, \sqrt{3}\}$. Let $\mathbf{N} = \{\mathbf{N}_a, a \in \mathbf{A}\}$ and $f_{a,\text{eq}}(\mathbf{m})$ denote the relative frequency of the equal label pairs in the equivalent voxel pairs $\{((x, y, z), (x+\xi, y+\eta, z+\zeta)): (x, y, z) \in \mathbf{R}; (x+\xi, y+\eta, z+\zeta) \in \mathbf{R}; (\xi, \eta, \zeta) \in \mathbf{N}\}$, then the MGRF model is:

$$P_h(\mathbf{m}) \propto \exp \sum_{(x,y,z) \in \mathbf{R}} \sum_{(\xi,\eta,\zeta) \in \mathbf{N}} \mathbf{V}_a(m_{x,y,z}, m_{x+\xi,y+\eta,z+\zeta}) \quad (2)$$

The initial region map results in the approximate analytical maximum likelihood estimates of the potentials²: $v_{a,\text{eq}} = -v_{a,\text{ne}} \approx 2f_{\text{eq}}(\mathbf{m}) - 1$; that allow for computing the voxel-wise probabilities $p_{h:x,y,z}(m_{x,y,z} = \lambda)$ of the labels; $\lambda \in \mathbf{L}$, at each step of the boundary evolution.

Probabilistic Shape Prior: To enhance the segmentation accuracy, the expected shape of the goal object is constrained with a probabilistic shape prior (Fig. 2). A training database collected from different subjects are co-aligned by rigid, namely affine, 3-D transformations maximizing their mutual information(MI) [13]. The shape prior is a spatially-variant independent random field of region labels $P_s(\mathbf{m}) = \prod_{(x,y,z) \in \mathbf{R}} p_{s:x,y,z}(m_{x,y,z})$ for the co-aligned, manually segmented training CT images. The factors are the empirical voxel-wise object

² For complete proof, please see: <https://louisville.edu/speed/bioengineering/faculty/bioengineering-full/dr-ayman-el-baz/supplemental-materials>

$p_{s:x,y,z}(1)$, and background $p_{s:x,y,z}(0) = 1 - p_{s:x,y,z}(1)$ probabilities. Each CT data to be segmented is initially co-aligned to the training database.

Stochastic Speed Function: All of the above models contribute to the voxel-wise guidance of the evolving level set. Let κ be the mean contour curvature. Let $\vartheta(x, y, z)$ specify the evolution magnitude and direction:

$$\vartheta(x, y, z) = \begin{cases} -P_{1:x,y,z} & \text{if } P_{1:x,y,z} > P_{0:x,y,z} \\ P_{0:x,y,z} & \text{otherwise} \end{cases} \quad (3)$$

where $P_{1:x,y,z} = \frac{\Omega_{1:x,y,z}}{\Omega_{1:x,y,z} + \Omega_{0:x,y,z}}$ and $P_{0:x,y,z} = 1 - P_{1:x,y,z}$. Here, $\Omega_{1:x,y,z} = p(q|1)p_{h:x,y,z}(1)p_{s:x,y,z}(1)$; $\Omega_{0:x,y,z} = p(q|0)(1 - p_{h:x,y,z}(1))(1 - p_{s:x,y,z}(1))$, and $p_{h:x,y,z}(1)$ is the probability of the object label in the Potts model $P_h(\mathbf{m})$. The stochastic speed function in Eq. (1) is defined as in [9]: $V(x, y, z) = \kappa\vartheta(x, y, z)$. The steps of our segmentation approach are summarized in Algorithm 1.

Algorithm 1. Key Steps for Level Set Segmentation

1. Construct the probabilistic shape prior from the training data sets.
 2. Approximate the empirical gray level distribution by using the LCDG with two dominant Gaussian modes.
 3. Form an initial region map \mathbf{m} using the estimated LCDG models.
 4. Find the Gibbs potentials for the MGRF model from the initial map.
 5. Find the stochastic speed function defined in Eq. (3).
 6. Evolve the level set using the determined speed function.
-

3 Experimental Results

The proposed approach has been tested on *in-vivo* 3-D CT kidney data sets (7 for training and 14 for testing and performance evaluation). The images were acquired by a GE light speed plus scanner (General Electric, Milwaukee, USA) using the following parameters: 120 KV, 250 mA, slice thickness of 0.9 mm, and FOV of 360 mm. To minimize the effect of inter-observer variations, each CT data is segmented by three independent experts and the ‘‘ground truth’’ is considered as the common segmented part of their segmentations. The CT images have marginal intensity distributions that are mixtures of two dominant modes: one mode for the kidney object and another mode for the background. Basic density estimation steps using the LCDG models are illustrated in Fig. 3 and Fig. 5 demonstrates the results of kidney segmentation.

The performance of our segmentation approach is evaluated based on the voxel-based overlap measured by using the Dice similarity coefficient (DSC) [14] and the average perpendicular distance (APD) between the automatic segmentation (\mathbf{C}) and the ground truth (\mathbf{G}). To measure the distances, one has to accurately co-locate the point-to-point correspondences between the borders. In

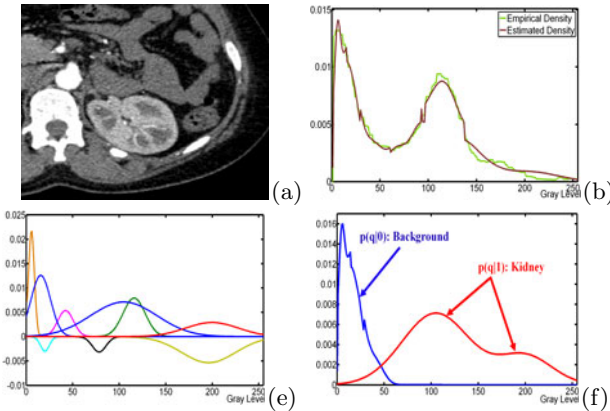


Fig. 3. LCDG-based modeling of the marginal probability distribution: (a) typical CT image, (b) its gray level density estimation (the final estimated density (brown) for the empirical density (green)), (c) sign-alternate LCDG components, and (d) final estimated LCDG for each class

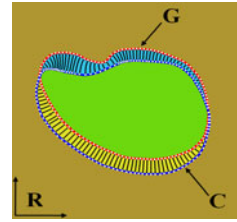


Fig. 4. 2-D illustration of the correspondences found by the solution of the Laplace equation between the ground truth **G** (manual expert segmentation) and automatic segmentation **C** obtained by the proposed segmentation approach

this paper, the correspondences, or matches between the borders' points (see e.g., Fig. 4), are found by solving the Laplace equation:

$$\nabla^2\psi = \frac{\partial^2\psi}{\partial^2x^2} + \frac{\partial^2\psi}{\partial^2y^2} + \frac{\partial^2\psi}{\partial^2z^2} = 0 \tag{4}$$

for a scalar potential field ψ . Table 1 represents the DSC and APD statistics obtained for all test data sets.

To highlight the advantages of our approach, we compared it to the shape-based approach proposed in [11] based on the average volumetric error $E_r = \frac{1}{G} (FP + FN)$, where FP and FN are the false positive and false negative segmentation errors, respectively. Table 2 compares the segmentation results over all the test data sets for our approach and the shape-based approach proposed in [11] with respect to the radiologist's segmentation. Differences between the average volumetric errors for our approach and the shape-based approach [11] are statistically significant by the unpaired *t*-test (the two-tailed *P*-value is $\leq 10^{-4}$). The final 3-D kidney segmentations for two of the test data sets and their associated FP and FN errors are shown in Fig. 6.

Another major metric to test the performance of our segmentation approach is to compute the receiver operating characteristic (ROC). Each point on the graph is generated by using a different cut point (i.e., classification threshold). Figure 7 shows the ROC curves of our three feature approach, intensity-based only (*I*), the intensity and spatial interactions-based (*I + S*), and the intensity and shape-based (*I + P*) level set segmentation. The figure clearly demonstrates that the area under the ROC curve is greatest for our approach ($A_z = 0.9423$).

Table 1. Segmentation accuracies over all test data sets. Note that APD values are in mm.

	DSC	APD
Min.	0.950	0.00
Max.	0.993	3.12
Mean	0.970	1.25
Std.	0.019	0.68

Table 2. Comparative segmentation accuracy over all test data sets.

$E_r, \%$	Algorithm	
	Our	[11]
Min.	1.26	6.76
Max.	7.72	20.08
Mean	3.71	13.95
Std.	2.18	5.96

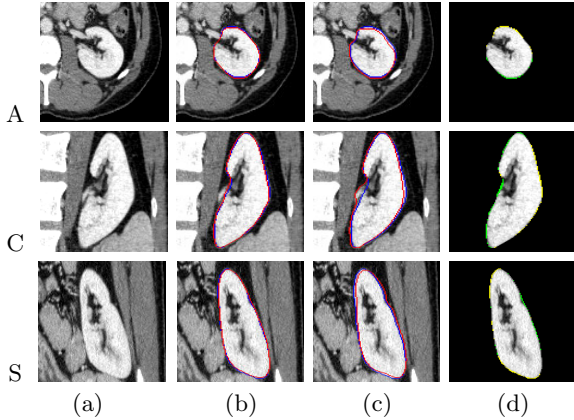


Fig. 5. 3-D kidney segmentation projected onto 2-D axial (A), coronal (C), and sagittal (S) planes for visualization: (a) 2-D CT images, (b) our segmentation (red) compared with the ground truth (blue), (c) the segmentation with the algorithm in [11], and (d) the associated FP (green) and FN (yellow) errors of our segmentation w.r.t. the ground truth

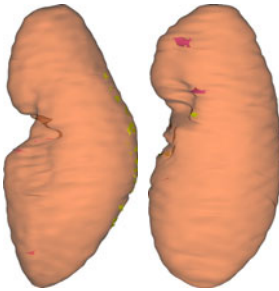


Fig. 6. 3-D visualization for the segmented kidneys using the proposed segmentation approach for two of the test data sets and their associated FP (pink) and FN (yellow) errors

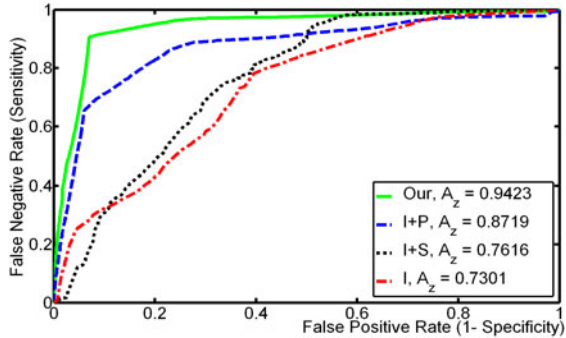


Fig. 7. The ROC curves for different level set-based segmentation guided by: the intensity information only (I ; red), the combined intensity and spatial interactions features ($I + S$; black), the combined intensity and shape features ($I + P$; blue), and the integrated three features (our; green). Note that A_z stands for the area under the curve.

4 Conclusions

This paper has presented a novel and automated 3-D approach for the segmentation of the kidney from abdominal CT images. Incorporation of the CT images'

features, namely the 3-D probabilistic shape, the 1st-order intensity, and the 2nd-order spatial interaction into the speed function notably improves the level set evolution and increases the segmentation accuracy and robustness based on both ROC curves and Dice similarity coefficient metrics. Experimental results showed that the proposed segmentation approach outperformed other methods for kidney segmentation. We plan to ultimately include this segmentation method in a kidney-dedicated CAD system designed for the early detection of acute renal rejection and treatment planning.

References

1. Pham, D.L., Xu, C., Prince, J.L.: Current Methods in Medical Image Segmentation. *Annu. Rev. Biomed. Eng.* 2, 315–338 (2000)
2. Pohle, R., Toennies, K.D.: A New Approach for Model-based Adaptive Region Growing in Medical Image Analysis. In: Skarbek, W. (ed.) CAIP 2001. LNCS, vol. 2124, pp. 238–246. Springer, Heidelberg (2001)
3. Wang, X., He, L., Wee, W.: Deformable Contour Method: A Constrained Optimization Approach. *Int. J. Comput. Vision* 59(1), 87–108 (2004)
4. Tsagaan, B., Shimizu, A., Kobatake, H., Miyakawa, K.: An Automated Segmentation Method of Kidney Using Statistical Information. In: Dohi, T., Kikinis, R. (eds.) MICCAI 2002. LNCS, vol. 2488, p. 556. Springer, Heidelberg (2002)
5. Huang Y.-P., Chung, P.-C., Huang, C.-L., Huang, C.-R.: Multiphase Level Set with Multi Dynamic Shape Models on Kidney Segmentation of CT Image. In: IEEE Biomedical Circuits and Systems Conference (BioCas 2009), pp. 141–144 (2009)
6. Spiegel, M., Hahn, D.A., Daumb, V., Waszab, J., Hornegger, J.: Segmentation of Kidneys Using a New Active Shape Model Generation Technique based on Non-rigid Image Registration. *Comput. Med. Imag. Graphics* 33, 19–39 (2009)
7. Freiman, M., Kronman, A., Esses, S., Joskowicz, L., Sosna, J.: Non-parametric Iterative Model Constraint Graph min-cut for Automatic Kidney Segmentation. In: Jiang, T., Navab, N., Pluim, J., Viergever, M. (eds.) MICCAI 2010. LNCS, vol. 6363, pp. 73–80. Springer, Heidelberg (2010)
8. Campadelli, P., Casiraghi, E., Pratisoli, S.: A Segmentation Framework for Abdominal Organs from CT Scans. *Artif. Intell. Med.* 50, 3–11 (2010)
9. Khalifa, F., El-Baz, A., Gimel'farb, G., Ousephand, R., Abu El-Ghar, M.: Shape-Appearance Guided Level Set Deformable Model for Image Segmentation. In: International Conference on Pattern Recognition (ICPR 2010), pp. 4581–4584 (2010)
10. Osher, S., Fedkiw, R.: *Level Set Methods and Dynamic Implicit Surfaces*. Springer, New York (2006)
11. Tsai, A., Yezzi Jr., A., Wells, W., Tempany, C., Tucker, D., Fan, A., Grimson, W.E., Willsky, A.: A Shape-based Approach to Curve Evolution for Segmentation of Medical Imagery. *IEEE Trans. Med. Imag.* 22(2), 137–154 (2003)
12. Farag, A., El-Baz, A., Gimel'farb, G.: Precise Segmentation of Multimodal Images. *IEEE Trans. Image Process.* 15(4), 952–968 (2006)
13. Viola, P., Wells III, W.M.: Alignment by Maximization of Mutual Information. *Int. J. Comput. Vision* 24(2), 137–154 (1997)
14. Dice, L.R.: Measures of the Amount of Ecologic Association Between Species. *Ecological Society of America* 26(3), 297–302 (1945)