Adaptive Metamorphs Model for 3D Medical Image Segmentation

Junzhou Huang[1], Xiaolei Huang[2], Dimitris Metaxas[1], and Leon Axel[3]

[1] Division of Computer and Information Sciences, Rutgers University, NJ, USA
[2] Department of Computer Science and Engineering, Lehigh University, PA, USA
[3] Department of Radiology, New York University, New York, USA.

Abstract. In this paper, we introduce an adaptive model-based segmentation framework, in which edge and region information are integrated and used adaptively while a solid model deforms toward object boundary. Our 3D segmentation method stems from Metamorphs deformable models [1]. The main novelty of our work is in that, instead of performing segmentation in an entire 3D volume, we propose model-based segmentation in an adaptively changing subvolume of interest. The subvolume is determined based on appearance statistics of the evolving object model, and within the subvolume, more accurate and object-specific edge and region information can be obtained. This local and adaptive scheme for computing edges and object region information makes our segmentation solution more efficient and more robust to image noise, artifacts and intensity inhomogeneity. External forces for model deformation are derived in a variational framework that consists of both edge-based and region-based energy terms, taking into account the adaptively changing environment. We demonstrate the performance of our method through extensive experiments using cardiac MR and liver CT images.

1 Introduction

Automated object boundary extraction is a fundamental problem in medical image analysis. It remains challenging to solve the problem robustly however, due to the common presence of cluttered objects, object texture, image noise, and various other artifacts in medical images. Efficiency is also often a concern, especially in 3D segmentation. Model-based methods have been extensively studied in recent years and achieved considerable success because of their ability to integrate high-level knowledge with low-level image processing [2–5]. The most commonly used models are either deformable models [2, 3, 1, 4, 5], or statistical models [6–8]. While both statistical shape [6] and appearance [7] models have been proposed to capture variations in an object in images, most deformable model frameworks [2, 3, 9] have been using primarily shape information, deriving external image forces from edge or gradient information. These shape-only deformable models often have difficulty in segmenting objects with texture patterns or in noisy images.

Some efforts have been made in the literature to integrate region information into shape-only deformable models. Geodesic Active Regions [4] deal with supervised texture segmentation in a frame partition framework. It assumes that the number of regions in an image is known and the region statistics are learned

off-line using a mixture-of-gaussian approximation. Region Competition [10] performs texture segmentation by combining region growing and active contours after applying a set of texture filters. The method assumes multivariate Gaussian distributions on the filter responses. In [11], an improved active contour deforms on a likelihood map instead of heuristically-constructed edge map; however, the segmented objects are dilated versions of the truth objects, which is caused by artificial neighborhood operations. Metamorphs [1] is proposed as a new class of deformable models that integrate boundary information and nonparametric region statistics for segmentation. In Metamorphs, shape and appearance of a model are defined in a common pixel space and both edge- and region-based energy terms are differentiable with respect to deformation parameters so that boundary and region information are naturally integrated. Another key property of Metamorphs is that the object region statistics are adaptively learned from the evolving model-interior statistics, requiring no other *a priori* information. Good segmentation results are obtained using Metamorphs in 2D images. However, the original Metamorphs framework suffers from efficiency problems when being extended to 3D because it requires computation in the whole image volume. Furthermore, edges in the original framework were computed statically using canny edge detector with a global threshold and did not get updated adaptively as the object model deforms.

In this paper, we propose a new, efficient, 3D model-based segmentation method based on Metamorphs. Our main novelty is to propose a local and adaptive scheme that focuses computation on a subvolume and uses the model interior's gradient and intensity statistics for adaptive edge and region of interest (ROI) computation. During model evolution, on one hand, the online-learning aspect of the 3D Metamorphs model keeps up-to-date appearance statistics of the object and uses this information to constrain model deformation, and on the other hand, the adaptively changing statistics enable the model to better handle image noise and gradual changes in intensity. In addition, we develop a new ROI based balloon energy term, from which one can derive anisotropic external forces that efficiently deform the model toward ROI boundary.

In the remainder of the paper, we briefly review Metamorphs [1] in section 2. We introduce the new 3D adaptive Metamorphs models in section 3. In section 4, experimental results are presented and we conclude in section 5.

2   Metamorphs Deformable Models

Considering traditional deformable models as "active contours" or "evolving curve fronts", the deformable shape and appearance models, termed "Metamorphs" [1], are "deforming disks or volumes" that have not only boundary shape but interior appearance.

Model Shape and Appearance Representations. The object model shape in Metamorphs is represented implicitly as a distance function. In this way, model shape is defined by a distance map "image", in the common pixel space as image intensity. The intensity distribution of a Metamorphs model's interior region is represented using a nonparametric kernel-based method [12].

Model Dynamics. The deformations that a Metamorphs model can undergo are defined using a space warping technique, Free Form Deformations (FFD)

[1]. FFD parameterizes warping deformations of the volumetric space in which the model is embedded, hence deforming both model boundary and interior simultaneously. When applied to segmentation, a Metamorphs model is initialized covering a seed region inside the object, and external image forces are derived in a variational framework consisting of both edge- and region-based energy terms. Both types of energy terms are differentiable with respect to the common set of FFD deformation parameters.

3 3D Adaptive Metamorphs Models

Shape Representation. The 3D extension of Metamorphs implicit shape representation is straightforward. Let $\Phi : \Omega \rightarrow R^+$ be a Lipschitz function that refers to the distance transform for the model shape $\mathcal{M}$. By definition Ω is bounded since it refers to the image domain. The shape defines a partition of the domain: the region that is enclosed by $\mathcal{M}$, $[\mathcal{R}_\mathcal{M}]$, the background $[\Omega - \mathcal{R}_\mathcal{M}]$, and on the model, $[\partial\mathcal{R}_\mathcal{M}]$ (In practice, we consider a narrow band around the model $\mathcal{M}$ in the image domain as $\partial\mathcal{R}_\mathcal{M}$). Given these definitions the following implicit shape representation is considered:

$$\Phi_\mathcal{M}(\mathbf{x}) = \begin{cases} 0, & \mathbf{x} \in \partial\mathcal{R}_\mathcal{M} \\ +D(\mathbf{x}, \mathcal{M}) > 0, & \mathbf{x} \in \mathcal{R}_\mathcal{M} \\ -D(\mathbf{x}, \mathcal{M}) < 0, & \mathbf{x} \in [\Omega - \mathcal{R}_\mathcal{M}] \end{cases}$$

where $D(\mathbf{x}, \mathcal{M})$ refers to the min Euclidean distance between the image pixel location $\mathbf{x} = (x, y, z)$ and the model $\mathcal{M}$. One example of shape representation is shown in Fig. 1 (2.e).

Appearance Representation. The nonparametric kernel-based method is used to represent 3D model-interior intensity distribution [1]. Suppose the model is deforming on image I, the image region bounded by current model $\Phi_\mathcal{M}$ is $\mathcal{R}_\mathcal{M}$, the nonparametric distribution over pixel intensity values i can be derived as:

$$\mathbf{P}(i|\Phi_\mathcal{M}) = \frac{1}{V(\mathcal{R}_\mathcal{M})} \iint_{\mathcal{R}_\mathcal{M}} \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(i-I(\mathbf{y}))^2}{2\sigma^2}} d\mathbf{y} \tag{1}$$

where $V(\mathcal{R}_\mathcal{M})$ denotes the volume of $\mathcal{R}_\mathcal{M}$, and σ is a constant specifying the width of the gaussian kernel. Using this nonparametric approximation, the intensity distribution of the model interior gets updated automatically while the model deforms. One example is shown in Fig. 1 (1.b).

Model Deformations. 3D Incremental Free Form Deformations (IFFD) is used to define model deformations. The essence of FFD is to deform the model by manipulating a regular control lattice F overlaid in its volumetric embedding space. Hence the model deformation parameters, $\mathbf{q}$, using FFD are deformations of control points in the lattice F:

$$\mathbf{q} = \{(\delta F^x_{h,w,s}, \delta F^y_{h,w,s}, \delta F^x_{h,w,s})\}; \ (h, w, s) \in [1, H] \times [1, W] \times [1, S]$$

The deformed position of a pixel $\mathbf{x} = (x, y, z)$ given the control lattice deformation from the initial regular configuration, F^0, to a new configuration, $F^0 + \delta F$, can be calculated through interpolation using cubic B-spline basis functions:

$$D(\mathbf{q}; \mathbf{x}) = \sum_{l=0}^{3} \sum_{m=0}^{3} \sum_{n=0}^{3} B_l(u) B_m(v) B_n(r)(F^0_{i+l,j+m,k+n} + \delta F_{i+l,j+m,k+n}) \tag{2}$$

- $\delta F_{i+l,j+m,k+n}$, $(l,m,n) \in [0,3] \times [0,3] \times [0,3]$ are the deformations of pixel **x**'s (sixty four) adjacent control points,
- (i,j,k) is the index (in global reference frame) of the control point located at the origin of a local reference frame defined by **x**'s 64 adjacent control points.
- $B_l(u)$, $B_m(v)$ and $B_n(r)$ are the lth, m^{th} and n^{th} basis function of a Cubic B-spine respectively [1].

3.1 Adaptive Edge and Region Information

(a)   (b)   (c)   (d)   (e)
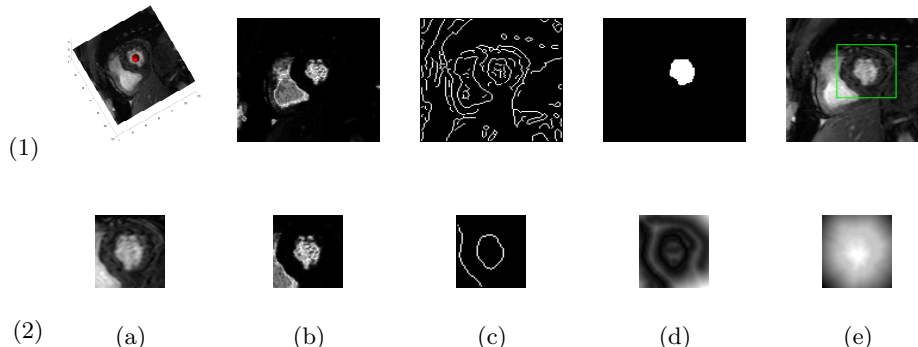
Fig. 1. An example of left ventricle segmentation in a 3D MR image. (1.a) Initial 3D sphere model; (1.b) Intensity probability map (only displaying one slice in a 3D image); (1.c) Edges based on global threshold; (1.d) ROI; (1.e) Localization of a subvolume of interest (IIV); (2.a) one slice from the 3D IIV; (2.b) updated Intensity probability map (2.c) updated edges based on local threshold in IIV; (2.d) distance map based on edges; (2.e) distance map based on ROI boundary.

To efficiently find object boundary, our algorithm starts with a simple-shape object model that is initialized covering a seed region inside the object. Edge and region information are then extracted and integrated adaptively as the model evolves over time. The overall model fitting process has the following steps.

- Step 1: Initialize a simple-shape (e.g. sphere, cylinder) 3D model centered around a seed point [Fig. 1(1.a)].
- Step 2: Compute nonparametric intensity statistics inside the model using Eqn. 1 [Fig. 1(1.b)]. The "Region Of Interest" (ROI) is determined as the largest possible region in the image that overlaps the model and has a consistent intensity distribution as the current model interior [1] (see [Fig. 1(1.d)]). Based on the ROI, we localize a subvolume of interest, which we call the "Interested Image Volume" (IIV) [Fig. 1(1.e)].
- Step 3: In IIV [Fig. 1(2.a)], compute the intensity probability map and update ROI based on a local probability threshold. Edges are also computed in IIV by performing 3D canny edge detection using a local gradient threshold. Comparing Fig. 1(2.b) and Fig. 1(2.c) with Fig. 1(1.b) and Fig. 1(1.c), one can see that the updated intensity probability map and edges are more accurate in our adaptive framework.
- Step 4: With edges and ROI computed in IIV, compute edge- and region-based energy terms(see section 3.2).

– Step 5: Evolve the deformable model, for a fixed number of iterations, toward object boundary based on model dynamics derived from the above energy terms. In [Fig. 2(b)], an intermediate model is shown.
– Step 6: Repeat steps 2-5 until convergence (e.g. when model deformation is sufficiently small). The converged model is shown in [Fig. 2(c)].

(a)   (b)   (c)
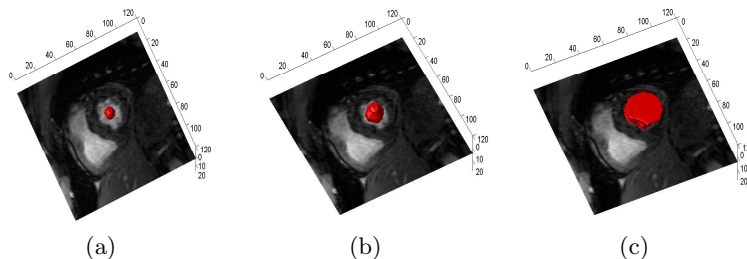
Fig. 2. Model evolution. (a) Initial model plus one slice in a 3D image; (b) Intermediate model; (c) Converged model.

3.2 3D Model Dynamics

Two energy terms are defined to deform the model toward object boundary.

The ROI Based Balloon Term. The ROI-based balloon term is designed to efficiently deform the model toward ROI boundary. After computing the intensity probability map P_I within IIV (see Step 3 in section 3.1, [Fig. 1(2.b)]) based on model-interior intensity statistics, a threshold (typically the mean probability in IIV) is applied on P_I to produce a binary IIV P_B. Assuming the object to be segmented is a solid without holes, we take the outer-most border of the connected component on this binary image P_B that is overlapping the model as the current ROI boundary. We encode this ROI boundary by computing its "shape image", which is its signed distance transform [Fig. 1(2.e)]. Denote this "shape image" as Φ_r, the ROI-based balloon term is defined as:

$$E_B = \frac{1}{S(\partial \mathcal{R}_{\mathcal{M}})} \iint_{\partial \mathcal{R}_{\mathcal{M}}} \Phi_r(\mathbf{x}) \big(\Phi_{\mathcal{M}}(D(\mathbf{q}; \mathbf{x})) \big) d\mathbf{x} \tag{3}$$

By the first component of this energy term $\Phi_{\mathcal{M}}(D(\mathbf{q}; \mathbf{x}))$ alone, the model boundary affinity pixels $\mathbf{x}$ will be mapped outward to locations $D(\mathbf{q}; \mathbf{x})$, where the model shape representation values $\Phi_{\mathcal{M}}(D(\mathbf{q}; \mathbf{x}))$ are smaller. Hence the model would expand and grow like a balloon so as to minimize the value of the energy term. The second component in the energy term, Φ_r, is the ROI "shape image" and encodes the distance value of each pixel from the ROI region boundary. It serves as a weighting factor for the first component so that the speed of model evolution is proportional to the distance of the model from the ROI boundary. Then, the model moves fast when it is far away from the boundary and the underlying $\Phi_r(\mathbf{x})$ values are large; it slows down as it approaches the boundary, and stops at the boundary. This ROI-based balloon term is very effective in countering the effect of un-regularized or inhomogeneous region intensities such as that caused by speckle noise and spurious edges inside the object of interest. Moreover, the ROI term generates adaptively changing balloon forces that expedite model convergence and improve convergence accuracy, especially when the shape of the object is elongated, or has salient protrusions or concavities.

The Adaptive Shape Term. We encode the adaptive edge information using a "shape image" Φ, which is derived from the distance transform of the edge map computed within IIV. In the previous Metamorphs model [1], the edge map is computed with a global threshold applied to the whole image and stay unchanged during model evolution [Fig. 1(1.c)]. In our model, the edge map is computed in a ROI-related, adaptively-changing subvolume instead of in the whole image [Fig. 1(2.c)]. As the model evolves, the edge map is adaptively re-computed based on gradient statistics inside the new model. To evolve a 3D model toward edges, we define an edge-based data term E_E on pixels in a narrow band around the model boundary $\partial\mathcal{R}_{\mathcal{M}}$.

$$E_E = \frac{1}{V(\partial\mathcal{R}_{\mathcal{M}})} \iint_{\partial\mathcal{R}_{\mathcal{M}}} \big(\Phi(D(\mathbf{q};\mathbf{x}))\big)^2 d\mathbf{x} \tag{4}$$

Intuitively, this term will encourage the deformation that maps the model boundary to edge locations where the underlying "shape image" distance values are as small (or as close to zero) as possible.

3.3   3D Model Evolution

A unified gradient-descent based parameter updating scheme can be derived using energy terms introduced above. The following evolution equation is derived for each element $\mathbf{q}_i$ in the model deformation parameters $\mathbf{q}$:

$$\frac{\partial E}{\partial \mathbf{q}_i} = \big(\frac{\partial E_E}{\partial \mathbf{q}_i} + a\frac{\partial E_B}{\partial \mathbf{q}_i}\big) \tag{5}$$

– The motion due to the adaptive shape term is:

$$\frac{\partial E_E}{\partial \mathbf{q}_i} = \frac{1}{S(\partial\mathcal{R}_{\mathcal{M}})} \iint_{\partial\mathcal{R}_{\mathcal{M}}} 2\Phi(D(\mathbf{q};\mathbf{x})) \cdot \big(\nabla\Phi(D(\mathbf{q};\mathbf{x})) \cdot \frac{\partial}{\partial \mathbf{q}_i}D(\mathbf{q};\mathbf{x})\big) d\mathbf{x}$$

– And the motion due to the ROI based Balloon term is:

$$\frac{\partial E_B}{\partial \mathbf{q}_i} = \frac{1}{S(\partial\mathcal{R}_{\mathcal{M}})} \iint_{\partial\mathcal{R}_{\mathcal{M}}} \Phi_r(\mathbf{x}) \big(\nabla\Phi_{\mathcal{M}}(D(\mathbf{q};\mathbf{x})) \cdot \frac{\partial}{\partial \mathbf{q}_i}D(\mathbf{q};\mathbf{x})\big) d\mathbf{x}$$

where a is a constant balancing the contribution of the two parts and the partial derivatives with respect to the deformation (FFD) parameters, $\frac{\partial}{\partial \mathbf{q}_i}D(\mathbf{q};\mathbf{x})$, can be easily derived from the model deformation formula for $D(\mathbf{q};\mathbf{x})$ [Eqn. (2)].

4   Experiments

Heart. We applied our adaptive model to segmenting endocardium of the left ventricle in tagged MRI images. The image data are 4D spatial-temporal short-axis cardiac tagged MR images. A 1.5T GE MR imaging system was used to acquire the images using an EGG-gated tagged gradient echo pulse sequence. Every 30ms, 2 sets of parallel short axis (SA) images were acquired; one with horizontal tags and one with vertical tags. Each set consists of 24 phases, with 16 slices (images) per phase. We collected 768 2D-images for testing. These 768 2D-images formed 48 3D images, each with the size of $192 \times 192 \times 16$. An expert was also asked to draw the left ventricle (LV) endocardium contours in these images for validation purposes.

Some example segmentation results are shown in Figure 3. Quantitative validation is performed by comparing the automated segmentation results with

expert solutions. Denote the expert segmentation in the images as l_{true}, and the results from our method as l_{amm}. The True Positive Fraction (TPF) describes the fraction of the total amount of tissue in the true segmentation that is over-lapped with the segmentation by our method: $FPF = \frac{|l_{amm} \cap l_{true}|}{l_{true}}$. On the entire test image set, the proposed method achieved an average 96.1% TPF. Our algorithm was implemented in MATLAB with embedded C code, and tested on a 2GHz Pentium 4 PC. The running time for LV endocardium segmentation in a 3D image of size $192 \times 192 \times 16$ is around 15 seconds.
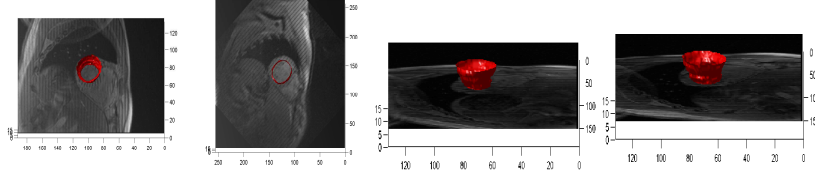
Fig. 3. Left ventricle endocardium segmentation example. Converged 3D model plus one slice in the 3D image from different views.

Liver MRI. Another application to which we apply our model-based method is segmenting tumors in liver CT images. These images were acquired using a LightSpeed Plus GE multi-slice CT scanner. The scans were reconstructed with a 38cm field of view. On average the CT system operated with an image acquisition time of 90 msec. The scanner generated X-rays with a peak voltage of 120 kVp and a maximum current of 260 mA. The axial resolution was 7.5mm and the image resolution of each cross section slice is 512 by 512. We collected 28 3D CT scans of the liver with 392 2D image slices in total. An expert was asked to draw the tumor contours in these test images. Some example segmentation results are shown in Figure 4. On the entire dataset, the proposed method achieved very encouraging segmentation results with an average of 91.7% TPF. The running time for tumor segmentation in these high-resolution CT images varied from 30 seconds to a few minutes depending the tumor size and noise level.

(1)

(2)

(a) (b) (c) (d)
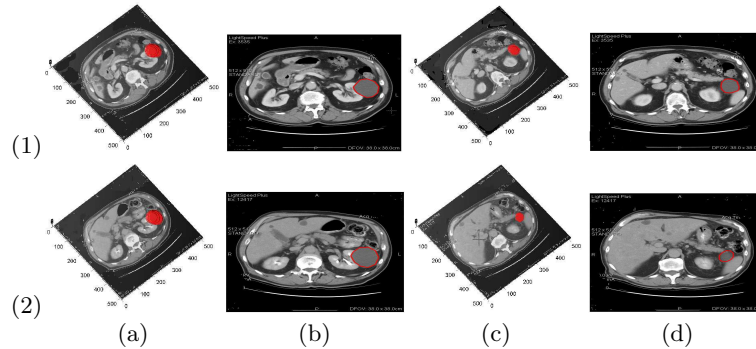
Fig. 4. Tumor segmentation examples. (1.a) The converged model representing the segmented tumor, plus one slice in a 3D CT image; (1.b) 2D view of the segmented contour in the same image slice; (1.c) The model plus another slice of the same 3D image; (1.d) The segmented contour in the same slice as in (1.c); (2) similar to (1) but for a different 3D image.

5 Conclusions

We have presented a new 3D adaptive Metamorphs deformable model. It is a major extension of the 2D Metamorphs model [1]. Instead of computing edge and region information statically and in the whole image, we adaptively calculate a "focus of attention" subvolume, in which edges and region of interest (ROI) are extracted based on object-specific criterion. A new ROI-based balloon term is developed, which is effective in countering the effect of inhomogeneous region intensities such as that caused by speckle noise and spurious edges inside the object. Compared to other works that integrate edge and region information for model-based segmentation, our model is novel in that it is an object model that has both boundary and interior. As the model evolves, its adaptively-changing appearance statistics will re-define its surrounding edges and ROI so that more reliable external image forces can be obtained. Focusing the computation on a subvolume containing the model also makes the framework more efficient.

References

1. Huang, X., Metaxas, D., Chen, T.: Metamorphs: Deformable shape and texture models. In: IEEE Conf. on Computer Vision and Pattern Recognition. (2004) 496–503
2. Kass, M., Witkin, A., Terzopoulos, D.: Snakes: Active contour models. Int'l Journal of Computer Vision **1** (1987) 321–331
3. Metaxas, D.: Physics-Based Deformable Models. Kluwer Academic Publishers (1996)
4. Paragios, N., Deriche, R.: Geodesic active regions and level set methods for supervised texture segmentation. Int'l Journal of Computer Vision **46**(3) (2002) 223–247
5. Vese, L.A., Chan, T.F.: A multiphase level set framework for image segmentation using the Mumford and Shah model. Int'l Journal of Computer Vision **50**(3) (2002) 271–293
6. Cootes, T.F., Taylor, C.J., Cooper, D.H., Graham, J.: Active shape models - their training and application. Computer Vision and Image Understanding **61**(1) (1995) 38–59
7. Cootes, T.F., Edwards, G.J., Taylor, C.J.: Active appearance models. In: Proc. of European Conf. on Computer Vision. Volume 2. (1998) 484–498
8. Leventon, M.E., Grimson, E.L., Faugeras, O.: Statistical shape influence in geodesic active contours. In: IEEE Conf. on Computer Vision and Pattern Recognition. Volume 1. (2000) 1316–1323
9. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic active contours. In: IEEE Int'l Conf. on Computer Vision. (1995) 694–699
10. Zhu, S., Yuille, A.: Region Competition: Unifying snakes, region growing, and Bayes/MDL for multi-band image segmentation. IEEE Trans. on Pattern Analysis and Machine Intelligence **18**(9) (1996) 884–900
11. Pujol, O., Radeva, P.: Texture segmentation by statistical deformable models. Intenational Journal of Image and Graphics **4**(3) (2004) 433–452
12. Comaniciu, D., Meer, P.: Mean shift: A robust approach toward feature space analysis. IEEE Trans. on Pattern Analysis and Machine Intelligence **24**(5) (2002) 603–619