A Multiphase Level Set based Segmentation Framework with Pose Invariant Shape Priors

Michael Fussenegger[1], Rachid Deriche[2], and Axel Pinz[1]

[1] Institute of Electrical Measurement and Measurement Signal Processing
Graz University of Technology
Schießstattgasse 14b, 8010 Graz, Austria
{Fussenegger, Axel.Pinz}@tugraz.at
http://www.emt.tugraz.at
[2] INRIA, 2004 route des Lucioles, BP 93
06902 Sophia Antipolis, France
Rachid.Deriche@sophia.inria.fr
http://www-sop.inria.fr/odyssee/

Abstract. *Level set based segmentation has been used with and without shape priors, to approach difficult segmentation problems in several application areas. This paper addresses two limitations of the classical level set based segmentation approaches: They usually deliver just two regions - one foreground and one background region, and if some prior information is available, they are able to take into account just one prior but not more. In these cases, one object of interest is reconstructed but other possible objects of interest and unfamiliar image structures are suppressed.*

The approach we propose in this paper can simultaneously handle an arbitrary number of regions and competing shape priors. Adding to that, it allows the integration of numerous pose invariant shape priors, while segmenting both known and unknown objects. Unfamiliar image structures are considered as separate regions. We use a region splitting to obtain the number of regions and the initialization of the required level set functions. In a second step, the energy of these level set functions is robustly minimized and similar regions are merged in a last step. All these steps are considering given shape priors. Experimental results demonstrate the method for arbitrary numbers of regions and competing shape priors.

1 Introduction

Segmenting an image into its semantically significant components is one of the fundamental problems in computer vision. Standard segmentation approaches are driven by low-level cues such as intensity, color or texture. But very often this segmentation of given objects is an ill-posed problem, therefore these methods have to fail. To overcome this limitation, prior knowledge can be used to constrain the segmentation process. Modelling this interaction between the data-driven and the model-based process has become an important topic in the research on image segmentation in the field of computer vision.

The integration of prior knowledge (in our case shape priors) into PDE based segmentation methods has delivered promising results (see [1–7]). Normally, the knowledge of one single shape prior is introduced into the contour evolution in a way that corrupted versions of a familiar object are reconstructed and all unfamiliar image structures are suppressed and often the localization of the shape must be known. Leventon et al. [3] use a Gaussian model to describe their shape priors. They assume a uniform distribution over pose parameters, that include translation and rotation. Rousson and Paragios [4] propose a similarity transformation (scale, rotation and translation) for the shape prior that allows to segment familiar objects with an unknown position in the image scene. But like the approach of Leventon et al. they can handle only one shape prior and unfamiliar image structures are ignored. Cremers et al. ([8], [6]) presented an approach with dynamic labeling, that allows to use more than one shape prior and does not suppress unfamiliar image structures. The problems of this approach are on one side the segmentation in only two regions and on the other side the incorrect segmentation of foreground objects, when one or more objects are very similar to the background. Lately, Raviv et al. [7] present a novel approach that allows a projective transformation of the shape prior, but their approach is also limited to one region. In all these approaches, it is nearly impossible to obtain the number or shapes of the unfamiliar objects in the scene. One possible way to solve that problem is to expand the level set based segmentation to an approach that allows to segment more than two regions.

For more than two regions, the level set idea loses part of its attractiveness. Therefore, there is only little related work on this problem. Paragios and Deriche [9] avoid this assumption by calculating the means of a Gaussian mixture estimation of the image histogram. The number of mixture coefficients determines the number of regions for the segmentation. Chan et al. [10] use a multiphase level set approach to segment many objects (N level-sets are used to intrinsically segment up to 2^N regions). This is a complementary approach to the one advocated in [11] to segment many objects with one level-set assigned to each object with a constraint to prevent the development of overlapping regions and/or vacuums. Brox and Weickert [12] propose a three step split and merge approach. In a first step, they use normal level set based segmentation to split the regions of an image in a recursive way. These regions are used as initialization for a level set based minimization scheme for the variational segmentation model of Zhu and Yuille [13]. In the last step, similar regions are merged to minimize the energy. All these approaches can segment different numbers of significant objects in an image, but do not use prior knowledge. In this paper, we combine the idea of level set based segmentation for multiple regions [12] with the prior knowledge of shapes to a framework which can handle these problems.

The outline of the paper is as follows: Section 2 shows a level set formulation that can easily be extended with a single shape prior. In section 3, we enhance this prior by explicit pose parameters and demonstrate the effect of a simultaneous pose optimization. In section 4, we introduce a multi region segmentation method similar to [12], that is extended with shape prior knowledge. It can handle an arbitrary number of known and unknown objects, which is also the central contribution of this work. We demonstrate that our approach is capable of reconstructing corrupted versions of multiple known objects in a scene containing other unknown objects.

2 Two Region Segmentation with a Shape Prior

There are different level set formulations, which could be possible choices [14–17]. In this work, we use the level set formulation proposed by Paragios and Deriche [17, 18] to minimize the energy for an object region:

$$E_D(\Phi, p_1, p_2) = -\int_{\Omega} (H(\Phi) \log p_1 + (1 - H(\Phi)) \log p_2) d\mathbf{x} + \nu \int_{\Omega} |\nabla H(\Phi)| d\mathbf{x}, \quad (1)$$

with the level set function $\Phi : \Omega \rightarrow \mathbb{R}$ with $\Phi(\mathbf{x}) > 0$ if $\mathbf{x} \in \Omega_1$ and $\Phi(\mathbf{x}) < 0$ if $\mathbf{x} \in \Omega_2$ and the Heaviside function $H(\Phi)$ with $\lim_{\Phi \rightarrow -\infty} H(\Phi) = 0$, $\lim_{\Phi \rightarrow \infty} H(\Phi) = 1$ and $H(0) = 0.5$. p_1 and p_2 are the probability densities $p_i = p(\mathbf{x}|\Omega_i)$ of the regions Ω_1 and Ω_2 which cover the whole image domain Ω with no overlap. For color images, we use the following multivariate Gaussian density:

$$p(\mathbf{x}|\Omega_i) = \frac{1}{(2\pi)^{d/2} |\Sigma_i|} e^{-\frac{1}{2}(\mathbf{x}-\mu_i)^T \Sigma_i^{-1} (\mathbf{x}-\mu_i)}, \quad (2)$$

with the mean μ_i and the covariance matrix Σ_i of the multivariate color distribution of the region Ω_i. The last term of equation 1 takes into account the length of the contour weighted by the parameter ν. To add an isotropic Gaussian shape prior to the energy equation 1, we define a straight forward extension

$$E(\Phi, \Phi_0, p_1, p_2) = E_D(\Phi, p_1, p_2) + \lambda E_S(\Phi, \Phi_0), \quad (3)$$

with

$$E_S(\Phi, \Phi_0) = \int_{\Omega} (\Phi - \Phi_0)^2 d\mathbf{x}, \quad (4)$$

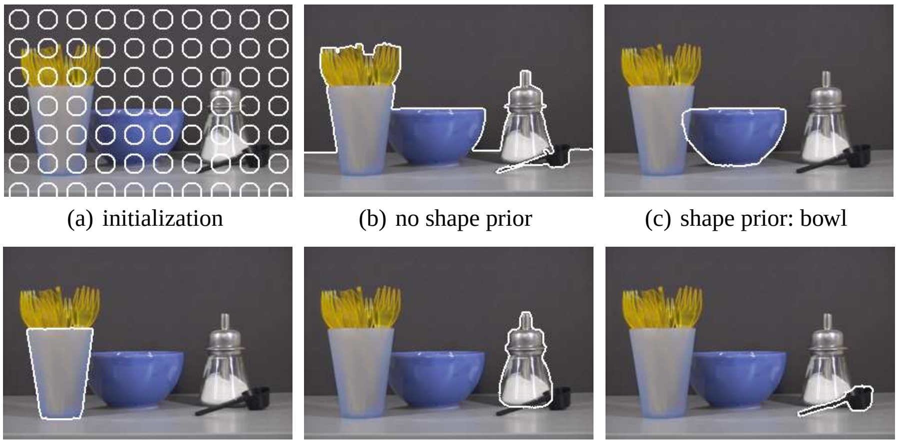
where Φ_0 is the level set of the given training shape or the mean of a set of training shapes. $\lambda \geq 0$ indicates the weight of the prior. Typically, λ is set to a value between 0.5 and 2.0.

The minimization of the energy term can now be estimated according to the gradient descent equation

$$\frac{\partial \Phi}{\partial t} = \delta(\Phi) \Big[\nu div \Big(\frac{\nabla \Phi}{|\nabla \Phi|} \Big) - log \frac{p_1}{p_2} \Big] - 2\lambda(\Phi - \Phi_0), \quad (5)$$

where $\delta(\Phi)$ is the derivative of $H(\Phi)$ with respect to its argument. The probability densities p_1 and p_2 are estimated with equation 2.

Figure 1 shows the different results of the level set segmentation with and without a shape prior. If a shape prior is used, the pose and position of the object of interest is assumed to be known. We show the original image containing different objects and the background 1(a), and the result of the standard level set segmentation without a shape prior 1(b). Subsequently, we present four results with different shape priors. With a high weight on the shape prior, the region of the specified object is correctly segmented, even if the object is partly occluded (see 1(c), 1(e)). All other objects, which are not in accordance with the given shape are suppressed. This problem is solved in section 4, but first we introduce a pose invariant formulation for the shape prior.



(a) initialization (b) no shape prior (c) shape prior: bowl

(d) shape prior: cup (e) shape prior: sugar sprinkler (f) shape prior: spoon

Fig. 1. (a) Original image with the level set initialization. (b) Level set segmentation result without prior knowledge. (c) With a shape prior of a bowl, (d) a cup, (e) a sugar sprinkler and (f) with a shape prior of a measuring spoon. With the prior knowledge of the shape, the region of its corresponding object is segmented correctly. However, it is not possible to segment the regions of two or more objects in one image, with one single level set function.

3 A Pose-Invariant Formulation

In the results shown in Figure 1, the pose and position of the object of interest is assumed to be known, but that will not be the case in realistic segmentation problems. If the object of interest is no longer presented at the same location, with the same scale and orientation as the shape prior Φ_0, a segmentation with the formalism of section 2 has to fail. Possible solutions are presented in [4], [5] and [7], where a set of pose parameters is associated with the given prior Φ_0. In our approach, we use the work of [4]. Compared to [7] the work of [4] is limited to similarity transformation, but tests have shown that on more complex scenes with more regions it is much more robust and therefore better for our use.

Rousson and Paragios [4] assume a global deformation $\mathcal{A}$ between Φ and Φ_0 that involves the parameters $[\mathcal{A} = (s;\theta;\mathbf{T})]$ with a scale factor s, a rotation angle θ and a translation vector **T**. The corresponding shape energy

$$E_S(\Phi,\Phi_0(\mathcal{A})) = \int_\Omega \delta(\Phi)(s\Phi - \Phi_0(\mathcal{A}))^2 d\mathbf{x} \tag{6}$$

is simultaneously optimized with respect to the segmentation level set function Φ and the pose parameters s, θ and **T**. The function is expanded with $\delta(\Phi)$, so that the shape prior is only estimated within the vicinity of the zero-crossing of the level set representation, which has a better performance than considering the whole image domain.

Minimizing equation 6 leads to the following gradient descent for the level set function Φ:

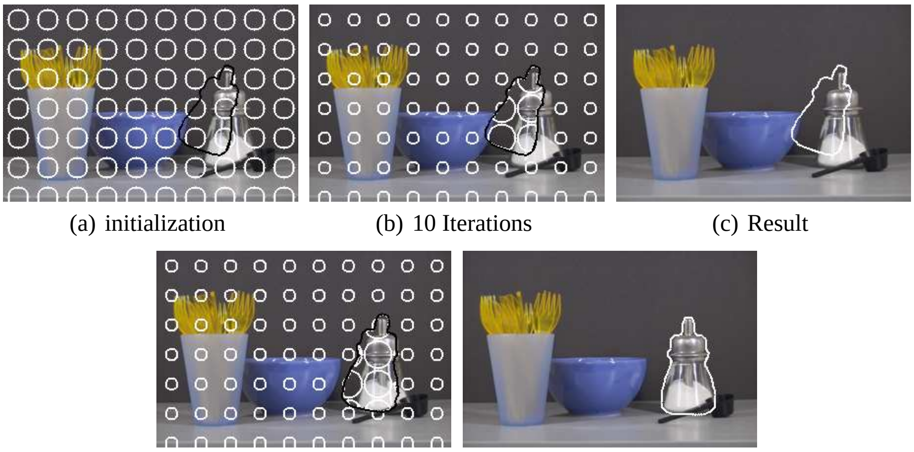
$$\frac{\partial \Phi}{\partial t} = \delta(\Phi)\Big[\nu div\Big(\frac{\nabla \Phi}{|\nabla \Phi|}\Big) - log\frac{p_1}{p_2} - 2\lambda(s\Phi - \Phi_0(\mathcal{A}))\Big]. \tag{7}$$

The transformation $\mathcal{A}$ is also dynamically updated to map Φ and Φ_0 in the best possible way. The calculus of variations for the parameters of $\mathcal{A}$ leads to the system:

$$\begin{aligned} \frac{\partial s}{\partial t} &= 2\int_{\Omega} p * (-\Phi + \nabla \Phi_0(\mathcal{A}) * \frac{\partial}{\partial s}\mathcal{A})d\mathbf{x} \\ \frac{\partial \theta}{\partial t} &= 2\int_{\Omega} p * (\nabla \Phi_0(\mathcal{A}) * \frac{\partial}{\partial \theta}\mathcal{A})d\mathbf{x} \\ \frac{\partial \mathbf{T}}{\partial t} &= 2\int_{\Omega} p * (\nabla \Phi_0(\mathcal{A}) * \frac{\partial}{\partial \mathbf{T}}\mathcal{A})d\mathbf{x}, \end{aligned} \tag{8}$$

with

$$p = \delta(\Phi)(s\Phi - \Phi_0(\mathcal{A})). \tag{9}$$



(a) initialization (b) 10 Iterations (c) Result

(d) 10 Iterations (e) Result

Fig. 2. Evolution of the shape contour (white) considering a shape prior (black). In the first row, the familiar object is forced to appear at a wrong location, without simultaneous pose optimization (figure 2(c)). In the second row, the same initialization is used but the parameters for the pose transformation are optimized. The shape of the familiar object and the shape prior correspond.

Figure 2 shows the resulting segmentation of the sugar bowl with and without the pose invariant formulation. In both cases the location of the shape prior is not identical with the location of the object. Without the pose invariant formulation (top row), the familiar shape is forced to appear in a wrong position 2(c). With the pose invariant formulation (bottom row), the shape of the familiar object and the shape prior correspond 2(e).

4 Multi Region Level Set Segmentation with Shape Priors

Brox and Weickert [12] introduce a split and merge level set based method to segment multiple regions. We expand their three step approach for multi region segmentation with shape priors.

The subsequent enumeration describes the three steps in detail:

1. Step 1: Splitting
   (a) For each given shape prior a split of Ω according to equation 5 is done, where the foreground region is assigned with the used shape prior and the background region is the new Ω (see figure 4(a)).
   (b) After all shape priors have been used for one split, the last Ω is split recursively using equation 5 without a prior. The final result delivers the expected number of regions in the image and is also the initialization for step 2 (see figure 4(b)).
2. Step 2: Refinement
   (a) The energy of all regions can now be minimized in a global scope with equation 10, considering also the regions assigned to the given shape priors. In the minimizing process, it can happen, that some regions become very small or even vanish. To get rid of these regions, we use the last step (see figure 4(c)).
3. Step 3: Merging
   (a) For all region pairs, where none of the two regions is assigned to a shape prior, the merged and the split energies are calculated using equation 1. If the merged energy ($E_{Merge} = E_D(\Omega_i)$) is smaller than the split energy ($E_{Split} = E_D(\Omega_{i1}) + E_D(\Omega_{i2})$) two regions are merged (see figure 4(d)).

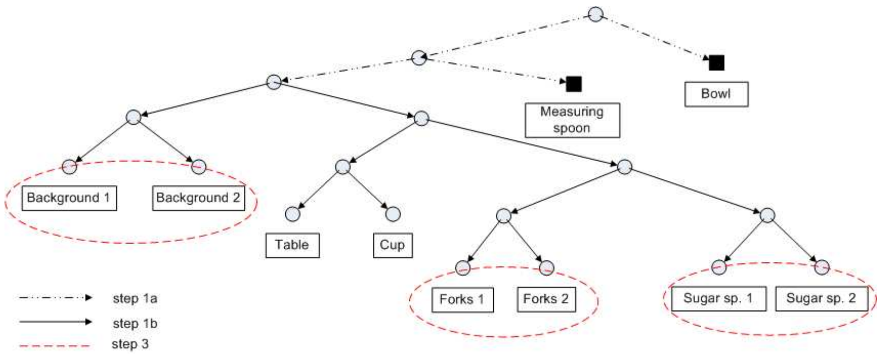
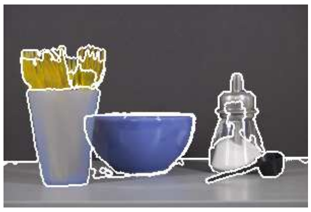


Fig. 3. Tree diagram to show the tried splittings (step 1a and 1b) and the merging (step 3).

Figure 4 shows the results after each of the above steps with two shape priors (bowl and measuring spoon). In 4(d) (final result) all objets are segmented correctly! The splittings steps and the merging step are also shown in figure 3, where every circular node of the tree symbolizes a tried split.

For the refinement we expand the gradient descent of Brox and Weickert [12] as follows:

(a) result after 1(a)



(b) result after 1(b)

(c) result after 2

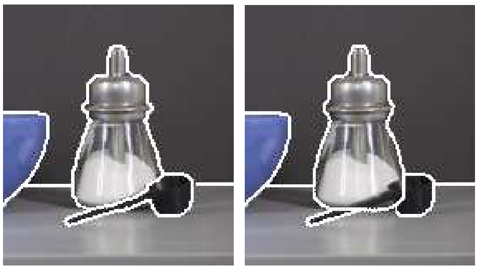
(d) result after 3

Fig. 4. Multi region level set segmentation with two shape priors. (a) Result of region splitting with shape priors (2 Regions), (b) after whole region splitting (10 Regions), (c) after refinement (10 Regions) and (d) final segmentation result after region merging (7 Regions). See also figure 3.

$$\frac{\partial \Phi_i}{\partial t} = \delta(\Phi_i)\Big[log p_i - \max_{j \neq i, H(\Phi_j)>0} log p_j + \frac{\nu}{2} div\Big(\frac{\nabla \Phi_i}{|\nabla \Phi_i|}\Big)$$
$$-2\lambda_i(s\Phi_i - \Phi_{0i}(A_i)) + 2\lambda_j(s\Phi_j - \Phi_{0j}(A_j))\Big], \qquad (10)$$

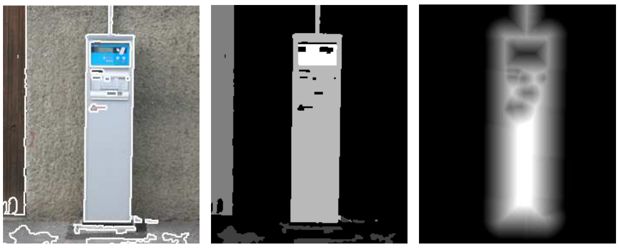
where the maximum criterion ensures that a pixel is only assigned to the region with the highest probability. $\lambda_i > 0$ when Φ_i is assigned to a shape prior and $\lambda_j > 0$ when Φ_j is assigned to a shape prior, they are zero when no shape prior is assigned to the corresponding level set function. When more than one shape prior is used, it can happen that one familiar object is partially occluded by an other familiar object. If all λ have the same value the front object is segmented completely. For increasing value of λ the occluded object is fully reconstructed. That means with a variation of the different λ, we can give each known object different importance. Figure 5(a) demonstrates the results for an equal λ for all shape priors. In figure 5(b) the shape prior of the sugar sprinkler has a higher λ than the shape prior of the measuring spoon.

Figures 6 and 7 show an other example on real images. First in figure 6, we use the multi region level set segmentation without shape priors (figure 6(a), 6(b)) and use the segmented region of the parking ticket machine as a shape prior for the segmentations in figure 7. In all three segmentation results of figure 7(a), 7(b) and 7(c) the partly occluded parking ticket machine is segmented correctly. In all images we use small circles as the initialization for the level set function Φ and a centered level set function Φ_0 (given in black) for the shape prior. The results also illustrate the robustness of the approach.



(a) $\lambda 1 = \lambda 2$ (b) $\lambda 1 > \lambda 2$

Fig. 5. Two segmentation results with varied λ for two shape priors ($\lambda 1$ sugar sprinkler, $\lambda 2$ measuring spoon).



(a) (b) (c)

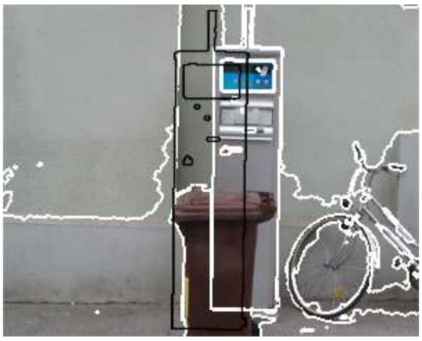
Fig. 6. Multi region level set segmentation without shape prios to get a shape prior of the parking ticket machine. (a) and (b) show the segmented image. (c) shows the resulting level set function Φ_0 of the shape, that is subsequently used as shape prior for the segmentation in fig. 7.

5 Conclusion

We have introduced the framework of level set based segmentation of multiple regions, that allows to integrate an arbitrary number of competing shape priors. Each shape prior is given by a fixed template (a given training shape or the mean of a set of training shapes) and respective pose parameters. An extension to statistical shape priors, with additional deformation modes is straight forward.

First, we have shown the benefit and limitation of using a shape prior with a standard level set based segmentation. The prior knowledge permits the reconstruction of corrupted versions of a familiar object, but suppresses independent unknown objects. Furthermore, we added a pose invariant formulation.

To the end our extension to more level set functions allows us to simply use multiple competing shape priors. And additional, independent unknown objects are not suppressed. Furthermore, the different regions can be much easier distinguished and assigned to the different objects in a scene, compared to the classical approach with only one level set function. The results we have presented in this work demonstrate the power and capacity of our approach.



(a)

(b)

(c)

Fig. 7. Three example segmentations (white) with the initialized shape prior (black) from figure 6. The results also demonstrate the robustness of the approach. In (a) the transformation parameters $\mathcal{A}$ are $s = 0.93$, $\theta = 0.3^\circ$ and $\mathbf{T} = [27, -2]$, in (b) $s = 0.61$, $\theta = -1.6^\circ$ and $\mathbf{T} = [-25, -56]$ and in (c) $s = 0.82, 6$, $\theta = -1.5^\circ$ and $\mathbf{T} = [-42, -31]$.

With its possibility to combine data-driven and recognition-driven information in the segmentation process, it can for example be used to improve an object recognition or detection framework.

Acknowledgement

This research has been partly funded by the Austrian Science Foundation (FWF, project S9103-N04), the European project IMAVIS HPMT-CT-2000-00040 within the framework of the Marie Curie Fellowships Training Sites Programme and by the European project LAVA (IST-2001-34405). This is gratefully acknowledged.

References

1. Yuille, A., Hallinan, P.: Deformable templates. In: A. Blake and A. Yuille, editors, Active Vison. (1992) 21–38

2. Cootes, T.F., A. Hill, C.J.T., Haslam, J.: Use of active shape models for locating structures in medical images. In: Image and Vison Computing. Volume 12:6. (1994) 355–365
3. Leventon, M.E., Grimson, W.E.L., Faugeras, O.: Statistical shape influence in geodesic active contour. In: Proceedings of Conference Computer Vision and Pattern Recognition. Volume 1. (2000) 316–323
4. Rousson, M., Paragios, N.: Shape priors for level set representations. In: Proceedings of European Conference of Computer Vision. Volume 2351 of LNCS. (2002) 78–92
5. Chen, Y., Tagare, H.D., Thiruvenkadam, S., Huang, F., Wilson, D., Gophinath, K.S., Briggs, R.W., Geiser, E.A.: Using prior shapes in geometric active contours in a variational framework. In: International Journal of Computer Vison. Volume 50(3). (2002) 315–328
6. Cremers, D., Sochen, N., Schnoerr, C.: Multiphase dynamic labeling for variational recognition-driven image segmentation. In: Proceedings of European Conference of Computer Vision. (2004) 74–86
7. Riklin-Raviv, T., Kiryati, N., Sochen, N.A.: Unlevel-sets: Geometry and prior-based segmentation. In: Proccedings of ECCV. (2004) 50–61
8. Cremers, D., Sochen, N., Schnoerr, C.: Towards recognition-based variational segmentation using shape priors and dynamic labeling. In: Proceedings of Scale-Space 2003. (2003) 388–400
9. Paragios, N., Deriche, R.: Coupled geodesic active regions for image segmentation: A level set approach. In: Proceedings of European Conference of Computer Vision. Volume 2. (2000) 224–240
10. Chan, T.F., Shen, J., Vese, L.: Variational PDE models in image processing. In: Notice of American Mathematical Society. Volume 50(1). (2003) 14–26
11. Zhao, H., Chan, T., Merrimann, B., Osher, S.: A variational level set approach to multiphase motion. In: Journal of Computational Physics. Volume 127. (1996) 179–195
12. Brox, T., Weickert, J.: Level set based image segmentation with multiple regions. In: Proceedings of 26th DAGM. (2004) 415–423
13. Zhu, S., Yuille, A.: Region competition: unifying snakes, region growing, and Bayes/MDL for multiband image segmentation. In: IEEE Transaction on Pattern Analysis and Machine Intelligence. Volume 18(9). (1996) 884–900
14. Osher, S.J., Sethian, J.A.: Fronts propagation with curvature depend speed: Algorithms based on Hamilton-Jacobi formulations. In: Journal of Comp. Phys. Volume 79. (1988) 12–49
15. Chan, T., Vese, L.: Active contours without edges. In: IEEE Transaction on Image Processing. Volume 10(2). (2001) 266–277
16. Tsai, A., Yezzi, A.J., Willsky, A.S.: Curve evolution implementation of the mumford-shah functional for image segmentation, denoising, interpolation and magnification. In: IEEE Transaction on Image Processing. Volume 10(8). (2001) 1169–1186
17. Paragios, N., Deriche, R.: Geodesic active regions : a new framework to deal with frame partition problems in computer vision. In: Journal of Visual Communication and Image Representation. Volume 13(1/2). (2002) 249–269
18. Paragios, N., Deriche, R.: Geodesic active regions and level set methods for motion estimation and tracking. In: Computer Vision and Image Understanding. Volume 97(3). (2005) 259–282