

Variational Segmentation of Image Sequences Using Deformable Shape Priors

Ketut Fundana, Niels Chr. Overgaard, and Anders Heyden

Applied Mathematics Group
School of Technology and Society
Malmö University, Sweden
{ketut.fundana,nc,heyden}@ts.mah.se

Abstract. The segmentation of objects in image sequences is an important and difficult problem in computer vision with applications to e.g. video surveillance. In this paper we propose a new method for variational segmentation of image sequences containing nonrigid, moving objects. The method is based on the classical Chan-Vese model augmented with a novel frame-to-frame interaction term, which allow us to update the segmentation result from one image frame to the next using the previous segmentation result as a shape prior. The interaction term is constructed to be pose-invariant and to allow moderate deformations in shape. It is expected to handle the appearance of occlusions which otherwise can make segmentation fail. The performance of the model is illustrated with experiments on real image sequences.

Keyword: Variational formulation, segmentation, tracking, region-based, level sets, interaction terms, deformable shape priors.

1 Introduction

In this paper we address the problem of segmentation in image sequences using region-based active contours and level set methods. Segmentation is an important and difficult process in computer vision, with the purpose of dividing a given image into one or several meaningful regions or objects. This process is more difficult when the objects to be segmented are moving and nonrigid. The shape of nonrigid, moving objects may vary a lot along image sequences due to, for instance, deformations or occlusions, which puts additional constraints on the segmentation process.

There have been a number of methods proposed and applied to this problem. Active contours are powerful methods for image segmentation; either boundary-based such as geodesic active contours [1], or region-based such as Chan-Vese models [2], which are formulated as variational problems. Those variational formulations perform quite well and have often been applied based on level sets. Active contour based segmentation methods often fail due to noise, clutter and occlusion. In order to make the segmentation process robust against these effects, shape priors have been proposed to be incorporated into the segmentation

process. In recent years, many researchers have successfully introduced shape priors into segmentation methods such as in [3,4,5,6,7,8,9].

We are interested in segmenting nonrigid moving objects in image sequences. When the objects are nonrigid, an appropriate segmentation method that can deal with shape deformations should be used. The application of active contour methods for segmentation in image sequences gives promising results as in [10,11,12]. These methods use variants of the classical Chan-Vese model as the basis for segmentation. In [10], for instance, it is proposed to simply use the result from one image as an initializer in the segmentation of the next.

The main purpose of this paper is to propose and analyze a novel variational segmentation method for image sequences, that can both deal with shape deformations and at the same time is robust to noise, clutter and occlusions. The proposed method is based on minimizing an energy functional containing the standard Chan-Vese functional as one part and a term that penalizes the deviation from the previous shape as a second part. The second part of the functional is based on a transformed distance map to the previous contour, where different transformation groups, such as Euclidean, similarity or affine, can be used depending on the particular application.

This paper is organized as follows: in Sect. 2 we discuss region-based segmentation, the level set method, and gradient descent procedures. In Sect. 3 we describe the segmentation model proposed. Experimental results of the model are presented in Sect. 4. We end the paper with some conclusions and future work plans.

2 Theoretical Background

2.1 Region-Based Segmentation

We begin with a brief review of the classical Chan-Vese segmentation model [2]. In this model a gray scale image is considered to be a real valued function $I : D \rightarrow \mathbf{R}$ defined on the *image domain* $D \subset \mathbf{R}^2$, usually a rectangle. A point $\mathbf{x} \in D$ is often referred to as a pixel, and the function value $I = I(\mathbf{x})$ as the *pixel value*, or the *gray scale value*. The Chan-Vese model is an active contour model. The idea is to find a contour Γ , by which we mean a finite union of disjoint, simple, closed curves in D , such that the image I is optimally approximated by a single gray scale value μ_{int} on $\text{int}(\Gamma)$, the *inside* of Γ , and by another gray scale value μ_{ext} on $\text{ext}(\Gamma)$, the *outside* of Γ . The optimal contour Γ^* and the corresponding pair of optimal gray scale values $\boldsymbol{\mu}^* = (\mu_{\text{int}}^*, \mu_{\text{ext}}^*)$ are defined as the solution of the variational problem,

$$E_{CV}(\boldsymbol{\mu}^*, \Gamma^*) = \min_{\boldsymbol{\mu}, \Gamma} E_{CV}(\boldsymbol{\mu}, \Gamma), \quad (1)$$

where E_{CV} is the well-known Chan-Vese functional,

$$E_{CV}(\boldsymbol{\mu}, \Gamma) = \alpha \int_{\Gamma} d\sigma + \beta \left\{ \frac{1}{2} \int_{\text{int}(\Gamma)} (I(\mathbf{x}) - \mu_{\text{int}})^2 d\mathbf{x} + \frac{1}{2} \int_{\text{ext}(\Gamma)} (I(\mathbf{x}) - \mu_{\text{ext}})^2 d\mathbf{x} \right\}. \quad (2)$$

Here $d\sigma$ denotes the arc length element, and $\alpha, \beta > 0$ are weight parameters. The first term in E_{CV} is the total length of the contour: It serves to regularize the optimal contour. The second term is the fidelity term, which penalizes deviations of the piecewise constant image model from the actual image I .

For any fixed contour Γ , not necessarily the optimal one, it turns out that the best choice of the gray scale values $\boldsymbol{\mu} = (\mu_{\text{int}}, \mu_{\text{ext}})$ corresponds to the mean value of the pixel values inside and the outside Γ , respectively:

$$\mu_{\text{int}} = \mu_{\text{int}}(\Gamma) = \frac{1}{|\text{int}(\Gamma)|} \int_{\text{int}(\Gamma)} I(\mathbf{x}) \, d\mathbf{x}, \quad (3)$$

$$\mu_{\text{ext}} = \mu_{\text{ext}}(\Gamma) = \frac{1}{|\text{ext}(\Gamma)|} \int_{\text{ext}(\Gamma)} I(\mathbf{x}) \, d\mathbf{x}. \quad (4)$$

Here the symbol $|A|$ denotes the area of the subset $A \subset \mathbf{R}^2$. Now, if we introduce the so-called “reduced” Chan-Vese functional

$$E_{CV}^R(\Gamma) := E_{CV}(\boldsymbol{\mu}(\Gamma), \Gamma), \quad (5)$$

then the optimal contour Γ^* can be found by solving the simpler minimization problem

$$E_{CV}^R(\Gamma^*) = \min_{\Gamma} E_{CV}^R(\Gamma). \quad (6)$$

Once Γ^* is found we have $\boldsymbol{\mu}^* = \boldsymbol{\mu}(\Gamma^*)$, of course. The minimization problem in (6) is solved using a gradient descent procedure, which will be recalled in the next section, after the material on the level set representation and the kinematics of moving surfaces have been presented.

2.2 The Level Set Method and Gradient Descent Evolutions

A simple closed curve Γ can be represented as the zero level set of a function $\phi : \mathbf{R}^2 \rightarrow \mathbf{R}$ as

$$\Gamma = \{\mathbf{x} \in \mathbf{R}^2 ; \phi(\mathbf{x}) = 0\}. \quad (7)$$

The sets $\text{int}(\Gamma) = \{\mathbf{x} ; \phi(\mathbf{x}) < 0\}$ and $\text{ext}(\Gamma) = \{\mathbf{x} ; \phi(\mathbf{x}) \geq 0\}$ are then the inside and the outside of Γ , respectively. Geometric quantities such as the outward unit normal \mathbf{n} and the curvature κ can be expressed in terms of ϕ as

$$\mathbf{n} = \frac{\nabla\phi}{|\nabla\phi|} \quad \text{and} \quad \kappa = \nabla \cdot \frac{\nabla\phi}{|\nabla\phi|}. \quad (8)$$

The function ϕ is usually called the *level set function* for Γ , cf. e.g. [13].

A curve evolution, that is, a time dependent curve $t \mapsto \Gamma(t)$ can be represented by a time dependent level set function $\phi : \mathbf{R}^2 \times \mathbf{R} \rightarrow \mathbf{R}$ as $\Gamma(t) = \{\mathbf{x} \in \mathbf{R}^2 ; \phi(\mathbf{x}, t) = 0\}$. Let us consider the kinematics of curve evolutions. It does not make sense to “track” points as there is no way of knowing the tangential motion of points on $\Gamma(t)$. The important notion is that of *normal velocity*. The normal velocity of a curve evolution $t \mapsto \Gamma(t)$ is the scalar function defined by

$$v(\Gamma)(\mathbf{x}) = \frac{d}{dt} \Gamma(t)(\mathbf{x}) := - \frac{\partial\phi(\mathbf{x}, t)/\partial t}{|\nabla\phi(\mathbf{x}, t)|} \quad (\mathbf{x} \in \Gamma(t)). \quad (9)$$

The normal velocity is independent of the curve representation, in particular of the choice of level set function ϕ for Γ , and is therefore a geometric property of the evolution, cf. [14]. The set of possible normal velocities $v = v(\Gamma)$ of moving contours $t \mapsto \Gamma(t)$ passing through the contour Γ at time $t = 0$ is an infinite dimensional vector space. This vector space can be endowed with a natural scalar product and a corresponding norm, cf. [14],

$$\langle v, w \rangle_\Gamma = \int_\Gamma v(\mathbf{x})w(\mathbf{x}) d\sigma \quad \text{and} \quad \|v\|_\Gamma^2 = \langle v, v \rangle_\Gamma, \quad (10)$$

where v, w are normal velocities and $d\sigma$ is the arc length element. In the following we therefore denote the vector space of normal velocities at Γ by $L^2(\Gamma)$.

The scalar product (10) is important in the construction of gradient descent flows for energy functionals $E(\Gamma)$ defined on curves. Suppose $v \in L^2(\Gamma)$ is a fixed normal velocity, and let $t \mapsto \Gamma(t)$ be any moving contour which satisfies $\Gamma(0) = \Gamma$, and $(d/dt)\Gamma(0) = v$. Then the Gâteaux variation $dE(\Gamma)v$ of the functional $E = E(\Gamma)$ at the contour Γ is defined as the derivative,

$$dE(\Gamma)v := \left. \frac{d}{dt} E(\Gamma(t)) \right|_{t=0}. \quad (11)$$

Suppose there exists a function $\nabla E(\Gamma) \in L^2(\Gamma)$ such that E 's Gâteaux variation $dE(\Gamma)v$ at Γ can be expressed in terms of the scalar product (10) in the following manner,

$$dE(\Gamma)v = \langle \nabla E(\Gamma), v \rangle_\Gamma \quad \text{for all } v \in L^2(\Gamma). \quad (12)$$

Then the vector $\nabla E(\Gamma)$ it is called the L^2 -gradient of E at Γ . It is unique if it exists. The gradient descent flow for the problem of minimizing $E(\Gamma)$ can now be defined as the initial value problem:

$$\frac{d}{dt}\Gamma(t) = -\nabla E(\Gamma(t)), \quad \Gamma(0) = \Gamma_0, \quad (13)$$

where Γ_0 is an initial contour specified by the user.

As an example, relevant for the application in this paper, notice that the L^2 -gradient of the reduced functional E_{CV}^R defined in (5) is given by:

$$\nabla E_{CV}^R(\Gamma; \mathbf{x}) = \alpha\kappa + \beta \left[\frac{1}{2}(I(\mathbf{x}) - \mu_{\text{int}}(\Gamma))^2 - \frac{1}{2}(I(\mathbf{x}) - \mu_{\text{ext}}(\Gamma))^2 \right], \quad (\mathbf{x} \in \Gamma), \quad (14)$$

where $\kappa = \kappa(\mathbf{x})$ is the curvature at $\mathbf{x} \in \Gamma$. If we combine the definition of gradient descent evolutions in (13) with the formula (9) for the normal velocity, then we get the gradient descent procedure in the level set framework:

$$\frac{\partial}{\partial t}\phi(\mathbf{x}, t) = \left(\alpha\kappa + \beta \left[\frac{1}{2}(I(\mathbf{x}) - \mu_{\text{int}}(\Gamma))^2 - \frac{1}{2}(I(\mathbf{x}) - \mu_{\text{ext}}(\Gamma))^2 \right] \right) |\nabla\phi(\mathbf{x}, t)|, \quad (15)$$

with $\phi(\mathbf{x}, 0) = \phi_0(\mathbf{x})$, where ϕ_0 is the level set function for the initial contour Γ_0 . It is understood that the gray scale values $\mu_{\text{int}}(\Gamma)$ and $\mu_{\text{ext}}(\Gamma)$ are given by (3) and (4), respectively.

3 Segmentation of Image Sequences

3.1 A Variational Updating-Model

In this section we are going to present the basic principles behind our variational model for updating segmentation results from one frame to the next in an image sequence.

Let $I_j : D \rightarrow \mathbf{R}$, $j = 1, \dots, N$, be a succession of frames from a given image sequence. Also, for some integer k , $1 \leq k \leq N$, suppose that all the frames I_1, I_2, \dots, I_{k-1} have already been segmented, such that the corresponding contours $\Gamma_1, \Gamma_2, \dots, \Gamma_{k-1}$ are available. In order to take advantage of the prior knowledge obtained from earlier frames in the segmentation of I_k , we propose the following method: If $k = 1$, i.e. if no previous frames have actually been segmented, then we just use the classical Chan-Vese model, as presented in Sect. 2. If $k > 1$, then the segmentation of I_k is given by the contour Γ_k which minimizes an *augmented* Chan-Vese functional of the form,

$$E_{CV}^A(\Gamma_{k-1}, \Gamma) := E_{CV}^R(\Gamma) + \gamma E_I(\Gamma_{k-1}, \Gamma), \quad (16)$$

where E_{CV}^R is the reduced Chan-Vese functional defined in (5), $E_I = E_I(\Gamma_{k-1}, \Gamma)$ is an *interaction term*, which penalizes deviations of the current active contour Γ from the previous one, Γ_{k-1} , and $\gamma > 0$ is a coupling constant which determines the strength of the interaction. The precise definition of E_I is described in the next section.

3.2 The Interaction Term

The interaction $E_I(\Gamma_0, \Gamma)$ between a fixed contour Γ_0 and an active contour Γ , used in (16), may be chosen in several different ways. Two common choices are the so-called pseudo-distances, cf. [6], and the area of the symmetric difference of the sets $\text{int}(\Gamma)$ and $\text{int}(\Gamma_0)$, cf. [3]. We have found that none of the mentioned contour interactions satisfy our needs, and we have therefore chosen to introduce a completely new pose-invariant interaction term.

To describe this interaction term, let $\phi_0 : D \rightarrow \mathbf{R}$ denote the *signed distance function* associated with the contour Γ_0 , that is, the function:

$$\phi_0(\mathbf{x}) = \begin{cases} \text{dist}(\mathbf{x}, \Gamma_0) & \text{for } \mathbf{x} \in \text{ext}(\Gamma_0), \\ -\text{dist}(\mathbf{x}, \Gamma_0) & \text{for } \mathbf{x} \in \text{int}(\Gamma_0). \end{cases} \quad (17)$$

Then the interaction $E_I = E_I(\Gamma_0; \Gamma)$ is defined by the formula,

$$E_I(\Gamma_0, \Gamma) = \min_T \int_{\text{int}(\Gamma)} \phi_0(T^{-1}\mathbf{x}) \, d\mathbf{x}, \quad (18)$$

where the minimum is taken over the *group of Euclidean transformations* $T : \mathbf{R}^2 \rightarrow \mathbf{R}^2$ which preserves the orientation of the plane, that is, transformations T which are compositions of translations and rotations (but not reflections).

Minimizing over groups of transformations is a standard device to obtain pose-invariant interactions, see [3] and [6].

For any given contour Γ , let $T = T(\Gamma)$ denote the transformation which minimizes the expression on the right hand side of (18). Since this is an optimization problem $T(\Gamma)$ can be found using gradient descent. For simplicity of presentation, suppose we only consider the group of translations $T_{\mathbf{a}} : \mathbf{x} \mapsto \mathbf{x} + \mathbf{a}$, $\mathbf{a} \in \mathbf{R}^2$, and want to determine the optimal translation vector $\mathbf{a} = \mathbf{a}(\Gamma)$. Then we have to solve the optimization problem

$$\min_{\mathbf{a}} \int_{\text{int}(\Gamma)} \phi_0(\mathbf{x} - \mathbf{a}) \, d\mathbf{x} .$$

The optimal translation $\mathbf{a}(\Gamma)$ can then be obtained as the limit, as time t tends to infinity, of the solution to initial value problem

$$\dot{\mathbf{a}}(t) = \int_{\text{int}(\Gamma)} \nabla \phi_0(\mathbf{x} - \mathbf{a}(t)) \, d\mathbf{x} , \quad \mathbf{a}(0) = \mathbf{0} . \quad (19)$$

Similar gradient descent schemes can be devised for rotations and scalings (in the case of similarity transforms), cf. [3], but will not be written out explicitly here.

3.3 The Gradient Descent Equations

The augmented Chan-Vese functional (16) is minimized using standard gradient descent as described in Sect. 2. That is, we solve the initial value problem

$$\frac{d}{dt} \Gamma(t) = -\nabla E_{CV}^A(\Gamma_{k-1}, \Gamma(t)) := -\nabla E_{CV}^R(\Gamma_{k-1}, \Gamma(t)) - \gamma \nabla E_I(\Gamma_{k-1}; \Gamma(t)), \quad (20)$$

with the initial contour $\Gamma(0) = \Gamma_{k-1}$, and pass to the limit $t \rightarrow \infty$. Here ∇E_{CV}^R is the L^2 -gradient of the reduced Chan-Vese functional, see Eq. (14), and ∇E_I is the L^2 -gradient of the interaction E_I , which is given by the formula,

$$\nabla E_I(\Gamma_{k-1}, \Gamma; \mathbf{x}) = \phi_{k-1}(T(\Gamma)\mathbf{x}), \quad (\text{for } \mathbf{x} \in \Gamma), \quad (21)$$

as is easily verified. Here ϕ_{k-1} is the signed distance function for Γ_{k-1} .

4 Numerical Implementation and Experiments

In this section we present the results obtained from experiments using three different image sequences. In the first image of the sequence we use the Chan-Vese model to segment a selected object with approximately uniform intensity. Then the proposed method is applied to segment the image sequences sequentially frame-by-frame, where the segmentation in one frame is used as the initial contour in the next one. The minimization of the functional, giving the optimal

contour, is obtained from the gradient descent procedure (20) which has been implemented in the level set framework outlined in Sect. 2. See also [13].

As illustrated in Fig. 1, the original Chan-Vese model is capable of segmenting a selected object in an image sequence without any problems. Further such results can be found in [10].



Fig. 1. Segmentation of a person in human walking sequence using the classical Chan-Vese model

Another experiment is given in Fig. 3, where a walking person is being segmented. Here the proposed method prevents the segmentation of the wrong objects, as is clearly shown.

However, as pointed out in the above reference, the classical Chan-Vese method will have problems segmenting an object if occlusions appear in the image which cover the whole or parts of the selected object. In Fig 2, we show the segmentation results for a car (the white van) in a traffic sequence, where occlusions occur. The classical Chan-Vese method fails to segment the selected object when it reaches the occlusion (first column). Using the proposed method, including the frame-to-frame interaction term, we obtain much better results (second column).

In both experiments the coupling constant γ is varied to see the influence of the interaction term on the segmentation results. The contour is only slightly affected by the prior if γ is small. On the other hand, if γ is too large, the contour will be close to a similarity transformed version of the prior.

5 Conclusions and Future Works

We have presented a new method for variational segmentation of image sequences containing nonrigid, moving objects. The proposed method is formulated as variational problem, with one part of the functional corresponding to the Chan-Vese model and another part corresponding to the pose-invariant interaction with a shape prior based on the previous contour. The optimal transformation as well as the shape deformation are determined by minimization of an energy functional using a gradient descent scheme. Preliminary results are shown on several real image sequences and its performance looks promising.

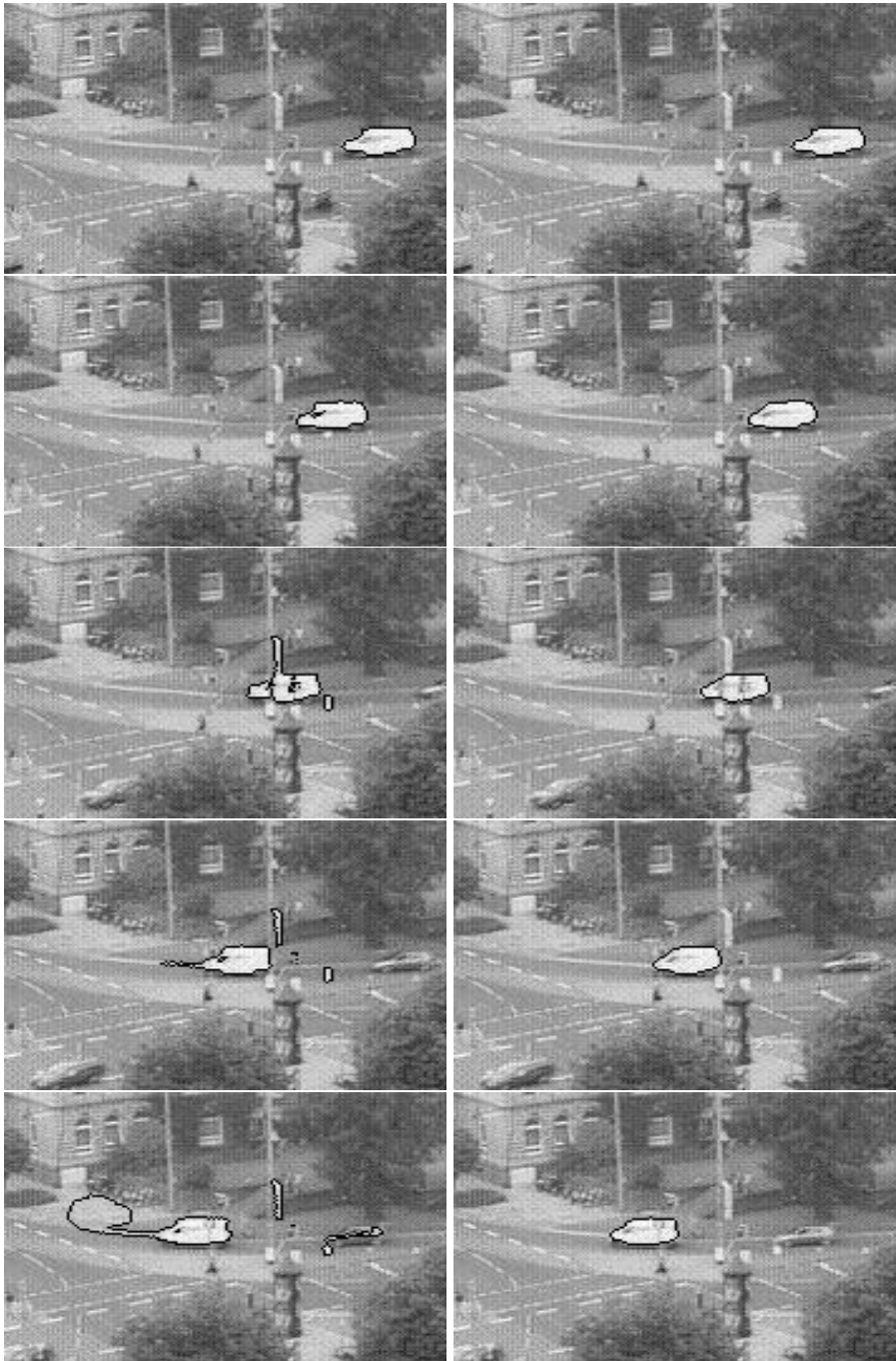


Fig. 2. Segmentation of a car which passes occlusions in the traffic sequence. Left Column: without interaction term, and Right Column: $\gamma = 80$.



Fig. 3. Segmentation of a person covered by an occlusion in the human walking sequence. Left Column: without interaction term, Middle Column: $\gamma = 20$, and Right Column: $\gamma = 70$.

Acknowledgements

This research is funded by EU Marie Curie RTN FP6 project VISIONTRAIN (MRTN-CT-2004-005439). The human walking sequences were downloaded from EU funded CAVIAR project (IST 2001 37540) website and the traffic sequence from KOGS/IAKS Universität Karlsruhe.

References

1. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic active contours. *International Journal of Computer Vision* 22(1), 61–79 (1997)
2. Chan, T., Vese, L.: Active contour without edges. *IEEE Transactions on Image Processing* 10(2), 266–277 (2001)
3. Chan, T., Zhu, W.: Level set based prior segmentation. Technical Report UCLA CAM 03-66, Department of Mathematics, UCLA (2003)
4. Cremers, D.: Statistical Shape Knowledge in Variational Image Segmentation. Phd thesis, Department of Mathematics and Computer Science, University of Mannheim (July 2002)
5. Cremers, D., Sochen, N., Schnörr, C.: Towards recognition-based variational segmentation using shape priors and dynamic labeling. In: Griffin, L.D, Lillholm, M. (eds.) *Scale Space Methods in Computer Vision*. LNCS, vol. 2695, pp. 388–400. Springer, Heidelberg (2003)
6. Cremers, D., Soatto, S.: A pseudo-distance for shape priors in level set segmentation. In: Faugeras, O., Paragios, N., (eds.): *2nd IEEE Workshop on Variational, Geometric and Level Set Methods in Computer Vision* (2003)
7. Cremers, D., Funke-Lea, G.: Dynamical statistical shape priors for level set based sequence segmentation. In: Paragios, N., Faugeras, O., Chan, T., Schnörr, C. (eds.) *VLSM 2005*. LNCS, vol. 3752, pp. 210–221. Springer, Heidelberg (2005)
8. Rousson, M., Paragios, N.: Shape priors for level set representations. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) *ECCV 2002*. LNCS, vol. 2351, pp. 78–92. Springer, Heidelberg (2002)
9. Leventon, M., Grimson, W., Faugeras, O.: Statistical shape influence in geodesic active contours. In: *CVPR* (2000)
10. Moelich, M., Chan, T.: Tracking objects with the chan-vese algorithm. Technical Report UCLA CAM 03-14, Department of Mathematics, UCLA (March 2003)
11. Paragios, N., Deriche, R.: Geodesic active contours and level set methods for the detection and tracking of moving objects. *IEEE Trans. PAMI* 22(3), 266–280 (2000)
12. Paragios, N., Deriche, R.: Geodesic active regions and level set methods for motion estimation and tracking. *Computer Vision and Image Understanding* 97, 259–282 (2005)
13. Osher, S., Fedkiw, R.: *Level Set Methods and Dynamic Implicit Surfaces*. Springer, Heidelberg (2003)
14. Solem, J.E., Overgaard, N.C.: A geometric formulation of gradient descent for variational problems with moving surfaces. In: Kimmel, R., Sochen, N.A., Weickert, J. (eds.) *Scale-Space 2005*. LNCS, vol. 3459, pp. 419–430. Springer, Heidelberg (2005)