

NONRIGID OBJECT SEGMENTATION AND OCCLUSION DETECTION IN IMAGE SEQUENCES

Ketut Fundana, Niels Chr. Overgaard, Anders Heyden

Applied Mathematics Group, School of Technology and Society, Malmö University, SE-205 06 Malmö, Sweden

ketut.fundana@ts.mah.se, nco@ts.mah.se, heyden@ts.mah.se

David Gustavsson, Mads Nielsen

DIKU, Copenhagen University, DK-2100 Copenhagen, Denmark

davidg@diku.dk, madsn@diku.dk

Keywords: Segmentation, occlusion, image sequences, variational active contour, variational contour matching

Abstract: We address the problem of nonrigid object segmentation in image sequences in the presence of occlusions. The proposed variational segmentation method is based on a region-based active contour of the Chan-Vese model augmented with a frame-to-frame interaction term as a shape prior. The interaction term is constructed to be pose-invariant by minimizing over a group of transformations and to allow moderate deformation in the shape of the contour. The segmentation method is then coupled with a novel variational contour matching formulation between two consecutive contours which gives a mapping of the intensities from the interior of the previous contour to the next. With this information occlusions can be detected and located using deviations from predicted intensities and the missing intensities in the occluded regions can be reconstructed. After reconstructing the occluded regions in the novel image, the segmentation can then be improved. Experimental results on synthetic and real image sequences are shown.

1 Introduction

Segmentation is an important and difficult process in computer vision which refers to the process 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 and even more when occlusions appear. 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.

Numerous methods have been proposed and applied to this problem. Active contours are powerful methods for image segmentation; either boundary-based such as geodesic active contours (Caselles et al., 1997), or region-based such as Chan-Vese models (Chan and Vese, 2001), 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 pro-

posed to be incorporated into the segmentation process, such as in (Chan and Zhu, 2005; Cremers et al., 2003; Cremers and Soatto, 2003; Cremers and Funka-Lea, 2005; Rousson and Paragios, 2002; Leventon et al., ; Bresson et al., 2006; Tsai et al., 2003; Chen et al., 2002). However, major occlusions is still a big problem. In order to improve the robustness of the segmentation methods in the presence of occlusions, it is necessary to detect and locate the occlusions (Strecha et al., 2004; Gentile et al., 2004; Konrad and Ristivojevic, 2003). Then using this information, the segmentation can be improved. For example, (Thiruvenkadam et al., 2007) proposed that the spatial order information in the image model is used to impose dynamically shape prior constraints only to occluded boundaries.

This paper focuses on the region-based variational approach to segment a non-rigid object in image sequences that may be partially occluded. We 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 variational framework is then augmented with a novel contour flow algorithm, giving a mapping of the intensities inside the contour of one image to the inside of the contour in the next image. Using this mapping, occlusions can be detected and located by simply thresholding the difference between the transformed intensities and the observed ones in the novel image. By using occlusions information, the occluded regions are reconstructed to improve the segmentation results.

2 Segmentation of Image Sequences

In this section, we describe the region-based segmentation model of Chan-Vese (Chan and Vese, 2001) and a variational model for updating segmentation results from one frame to the next in an image sequence.

2.1 Region-Based Segmentation

The idea of the Chan-Vese model (Chan and Vese, 2001) is to find a contour Γ such that the image I is optimally approximated by a 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^* is defined as the solution of the variational problem,

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

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

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

Here $|\Gamma|$ is the arc length of the contour, $\alpha, \beta > 0$ are weight parameters, and

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

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

The gradient descent flow for the problem of minimizing a functional $E_{CV}(\Gamma)$ is the solution to initial

value problem:

$$\frac{d}{dt} \Gamma(t) = -\nabla E_{CV}(\Gamma(t)), \quad \Gamma(0) = \Gamma_0, \quad (5)$$

where Γ_0 is an initial contour. Here $\nabla E_{CV}(\Gamma)$ is the L^2 -gradient of the energy functional $E_{CV}(\Gamma)$, cf. e.g. (Solem and Overgaard, 2005) for definitions of these notions. Then the L^2 -gradient of E_{CV} is

$$\nabla E_{CV}(\Gamma) = \alpha \kappa + \beta \left[\frac{1}{2} (I - \mu_{\text{int}}(\Gamma))^2 - \frac{1}{2} (I - \mu_{\text{ext}}(\Gamma))^2 \right], \quad (6)$$

where κ is the curvature.

In the level set framework (Osher and Fedkiw, 2003), a curve evolution, $t \mapsto \Gamma(t)$, can be represented by a time dependent level set function $\phi : \mathbf{R}^2 \times \mathbf{R} \rightarrow \mathbf{R}$ as $\Gamma(t) = \{\mathbf{x} \in \mathbf{R}^2 ; \phi(\mathbf{x}, t) = 0\}$, $\phi(\mathbf{x}) < 0$ and $\phi(\mathbf{x}) > 0$ are the regions inside and the outside of Γ , respectively. The normal velocity of $t \mapsto \Gamma(t)$ is the scalar function $d\Gamma/dt$ defined by

$$\frac{d}{dt} \Gamma(t)(\mathbf{x}) := -\frac{\partial \phi(\mathbf{x}, t)/\partial t}{|\nabla \phi(\mathbf{x}, t)|} \quad (\mathbf{x} \in \Gamma(t)). \quad (7)$$

Recall that the outward unit normal \mathbf{n} and the curvature κ can be expressed in terms of ϕ as $\mathbf{n} = \nabla \phi / |\nabla \phi|$ and $\kappa = \nabla \cdot (\nabla \phi / |\nabla \phi|)$.

Combined with the definition of gradient descent evolutions (5) and the formula for the normal velocity (7) this gives the gradient descent procedure in the level set framework:

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

where $\phi(\mathbf{x}, 0) = \phi_0(\mathbf{x})$ represents the initial contour Γ_0 .

2.2 The Interaction Term

The interaction $E_I(\Gamma_0, \Gamma)$ between a fixed contour Γ_0 and an active contour Γ may be regarded as a shape prior and be chosen in several different ways, such as the area of the symmetric difference of the sets $\text{int}(\Gamma)$ and $\text{int}(\Gamma_0)$, cf. (Chan and Zhu, 2005), and the pseudo-distances, cf. (Cremers and Soatto, 2003).

Let $\phi = \phi(\mathbf{x})$ and $\phi_0 = \phi_0(\mathbf{x})$ denote the signed distance functions associated with Γ and Γ_0 , respectively, where \mathbf{x} is a generic point in the image domain R . By assuming that Γ_0 is already optimally aligned with Γ in the appropriate sense, then the interaction term proposed in this paper has the form:

$$E_I(\Gamma, \Gamma_0) = \int_{\text{int}(\Gamma)} \phi_0(\mathbf{x}) d\mathbf{x}. \quad (8)$$

The area of the symmetric difference, which has been used in (Chan and Zhu, 2005) and (Riklin-Raviv et al., 2007) has the form:

$$E_I^{\text{SD}}(\Gamma, \Gamma_0) = \text{area}(\Omega \triangle \Omega_0), \quad (9)$$

where the notation $\Omega \Delta \Omega_0 := (\Omega \cup \Omega_0) \setminus (\Omega \cap \Omega_0)$ to denote the symmetric difference of the two sets $\Omega = \text{int}(\Gamma)$, $\Omega_0 = \text{int}(\Gamma_0)$. The pseudo-distance has the form:

$$E_I^{\text{PD}}(\Gamma, \Gamma_0) = \frac{1}{2} \int_R [\phi(\mathbf{x}) - \phi_0(\mathbf{x})]^2 d\mathbf{x}, \quad (10)$$

which has been studied, with various minor modifications, in (Rousson and Paragios, 2002), (Paragios et al., 2003), and (Cremers and Soatto, 2003).

The main benefit of our interaction term defined in (8) is that its L^2 -gradient can be computed easily:

$$\nabla_{\Gamma} E_I(\Gamma, \Gamma_0) = \phi_0(\mathbf{x}) = \phi(\Gamma_0; \mathbf{x}) \quad (\mathbf{x} \in \Gamma)$$

and that this gradient is small if Γ is close to the shape prior Γ_0 , and large if the active contour is far from the shape prior. However, $E_I(\Gamma, \Gamma_0)$ is not symmetric in Γ and Γ_0 , which may in general be considered a drawback. However, in our particular application, where we want to use shape information from a previous image frame (Γ_0) to guide the segmentation in the current frame (Γ), the lack of symmetry does not seem to be such a big issue. After all, there is no obvious symmetry between *past* and *present!*. Suppose instead that we wanted to segment an image sequence simultaneously, by considering the stack of frames as a three-dimensional object, then it would be relevant to use a symmetric interaction term between the contours in the individual frames.

The proposed interaction term is constructed to be pose-invariant and to allow moderate deformations in shape. Let $\mathbf{a} \in \mathbf{R}^2$ is a group of translations. We want to determine the optimal translation vector $\mathbf{a} = \mathbf{a}(\Gamma)$, then the interaction $E_I = E_I(\Gamma_0, \Gamma)$ is defined by the formula,

$$E_I(\Gamma_0, \Gamma) = \min_{\mathbf{a}} \int_{\text{int}(\Gamma)} \phi_0(\mathbf{x} - \mathbf{a}) d\mathbf{x}. \quad (11)$$

Minimizing over groups of transformations is the standard device to obtain pose-invariant interactions, see (Chan and Zhu, 2005) and (Cremers and Soatto, 2003).

Since this is an optimization problem $\mathbf{a}(\Gamma)$ can be found using the gradient descent procedure. 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

$$\dot{\mathbf{a}}(t) = \int_{\text{int}(\Gamma)} \nabla \phi_0(\mathbf{x} - \mathbf{a}(t)) d\mathbf{x}, \quad \mathbf{a}(0) = 0. \quad (12)$$

Similar gradient descent schemes can be devised for rotations and scalings (in the case of similarity transforms), cf. (Chan and Zhu, 2005).

2.3 Using the Interaction Term in Segmentation of Image Sequences

Let $I_j : D \rightarrow \mathbf{R}$, $j = 1, \dots, N$, be a succession of N 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 standard Chan-Vese model, as presented in Sect. 2.1. 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,

$$E_{CV}^A(\Gamma_{k-1}, \Gamma_k) := E_{CV}(\Gamma_k) + \gamma E_I(\Gamma_{k-1}, \Gamma_k), \quad (13)$$

where E_{CV} is the Chan-Vese functional, $E_I = E_I(\Gamma_{k-1}, \Gamma_k)$ is an *interaction term*, which penalizes deviations of the current active contour Γ_k from the previous one, Γ_{k-1} , and $\gamma > 0$ is a coupling constant which determines the strength of the interaction. See Algorithm 1.

The augmented Chan-Vese functional (13) is minimized using standard gradient descent (5) described in Sect. 2.1 with ∇E equal to

$$\nabla E_{CV}^A(\Gamma_{k-1}, \Gamma_k) := \nabla E_{CV}(\Gamma_k) + \gamma \nabla E_I(\Gamma_{k-1}; \Gamma_k), \quad (14)$$

and the initial contour $\Gamma(0) = \Gamma_{k-1}$. Here ∇E_{CV} is the L^2 -gradient (6) of the Chan-Vese functional, and ∇E_I the L^2 -gradient of the interaction term, which is given by the formula,

$$\nabla E_I(\Gamma_{k-1}, \Gamma_k; \mathbf{x}) = \phi_{k-1}(\mathbf{x} - \mathbf{a}(\Gamma_k)), \quad (\text{for } \mathbf{x} \in \Gamma_k). \quad (15)$$

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

Algorithm 1 The algorithm for segmentation of N frames image sequence from the second frame $I_2 \dots I_N$.

INPUT: Current frame I_k and the level set function from the previous frame ϕ_{k-1}

OUTPUT: Optimal level set function ϕ_k .

1. **Initialization** Initialize the level set function $\phi_k = \phi_{k-1}$.
 2. **Computation** Compute the optimal translation vector and then the gradient descent of (14).
 3. **Re-initialization** Re-initialize the level set function ϕ_k .
 4. **Convergence** Stop if the level set evolution converges, otherwise go to step 2.
-

3 Occlusion Detection by Contour Matching

In this section we are going to present a variational solution to a contour matching problem. We start with the theory behind the contour matching problem and then describe the algorithm we use to implement it to detect and locate the occlusions. See (Gustavsson et al., 2007) for more detail.

3.1 A Contour Matching Problem

Suppose we have two simple closed curves Γ_1 and Γ_2 contained in the image domain Ω . Find the “most economical” mapping $\Phi = \Phi(x) : \Omega \rightarrow \mathbf{R}^2$ such that Φ maps Γ_1 onto Γ_2 , i.e. $\phi(\Gamma_1) = \Gamma_2$. The latter condition is to be understood in the sense that if $\alpha = \alpha(\gamma) : [0, 1] \rightarrow \Omega$ is a positively oriented parametrization of Γ_1 , then $\beta(\gamma) = \Phi(\alpha(\gamma)) : [0, 1] \rightarrow \Omega$ is a positively oriented parametrization of Γ_2 (allowing some parts of Γ_2 to be covered multiple times).

To present our variational solution of this problem, let \mathcal{M} denote the set of twice differential mappings Φ which maps Γ_1 to Γ_2 in the above sense. Loosely speaking

$$\mathcal{M} = \{\Phi \in C^2(\Omega; \mathbf{R}^2) \mid \Phi(\Gamma_1) = \Gamma_2\}.$$

Moreover, given a mapping $\Phi : \Omega \rightarrow \mathbf{R}^2$, not necessarily a member of \mathcal{M} , then we express Φ in the form $\Phi(x) = x + U(x)$, where the vector valued function $U = U(x) : \Omega \rightarrow \mathbf{R}^2$ is called the *displacement field associated with Φ* , or simply the displacement field. It is sometimes necessary to write out the components of the displacement field; $U(x) = (u_1(x), u_2(x))^T$.

We now define the “most economical” map to be the member Φ^* of \mathcal{M} which minimizes the following energy functional:

$$E[\Phi] = \frac{1}{2} \int_{\Omega} \|DU(x)\|_F^2 d\mathbf{x}, \quad (16)$$

where $\|DU(x)\|_F$ denotes the Frobenius norm of $DU(x) = [\nabla u_1(x), \nabla u_2(x)]^T$, which for an arbitrary matrix $A \in \mathbf{R}^{2 \times 2}$ is defined by $\|A\|_F^2 = \text{tr}(A^T A)$. That is, the optimal matching is given by

$$\Phi^* = \arg \min_{\Phi \in \mathcal{M}} E[\Phi]. \quad (17)$$

The solution Φ^* of the minimization problem (17) must satisfy the following Euler-Lagrange equation:

$$0 = \begin{cases} \Delta U^* - (\Delta U^* \cdot \mathbf{n}_{\Gamma_2}^*) \mathbf{n}_{\Gamma_2}^*, & \text{on } \Gamma_1, \\ \Delta U^*, & \text{otherwise,} \end{cases} \quad (18)$$

where $\mathbf{n}_{\Gamma_2}^*(x) = \mathbf{n}_{\Gamma_2}(x + U^*(x))$, $x \in \Gamma_1$, is the pull-back of the normal field of the target contour Γ_2 to

the initial contour Γ_1 . The standard way of solving (18) is to use the gradient descent method: Let $U = U(t, x)$ be the time-dependent displacement field which solves the evolution PDE

$$\frac{\partial U}{\partial t} = \begin{cases} \Delta U - (\Delta U \cdot \mathbf{n}_{\Gamma_2}^*) \mathbf{n}_{\Gamma_2}^*, & \text{on } \Gamma_1, \\ \Delta U, & \text{otherwise,} \end{cases} \quad (19)$$

where the initial displacement $U(0, x) = U_0(x) \in \mathcal{M}$ specified by the user, and $U = 0$ on $\partial\Omega$, the boundary of Ω (Dirichlet boundary condition). Then $U^*(x) = \lim_{t \rightarrow \infty} U(t, x)$ is a solution of the Euler-Lagrange equation (18). Notice that the PDE (19) coincides with the so-called *geometry-constrained diffusion* introduced in (Andresen and Nielsen, 1999). Thus we have found a variational formulation of the non-rigid registration problem considered there.

Implementation. Following (Andresen and Nielsen, 1999), a time and space discrete algorithm for solving the geometry-constrained diffusion problem can be found by iteratively convolving the displacement field with a Gaussian kernel and then project the deformed contour Γ_1 back onto contour Γ_2 such that the constraints are satisfied (see Algorithm 2). The algorithm needs a initial registration provided by the user. In our implementation we have translated Γ_1 and projected it onto Γ_2 and used this as the initial registration. This gives good results in our case where the deformation and translation is quite small. Dirichlet boundary condition - zero padding in the discrete implementation - have been used. By pre-registration and embedding the image into a larger image, the boundary conditions seems to be a minor practical issue. The displacement field is diffused using convolution in each of x and y coordinates independently with a fix time parameter.

3.2 Occlusion Detection

The mapping $\Phi = \Phi(x) : \Omega \rightarrow \mathbf{R}^2$ such that Φ maps Γ_1 onto Γ_2 is an estimation of the displacement (motion and deformation) of the boundary of an object between two frames. By finding the displacement of the contour, a consistent displacement of the intensities inside the closed curve Γ_1 can also be found. Φ maps Γ_1 onto Γ_2 and pixels inside Γ_1 are mapped inside Γ_2 . This displacement field which only depends on displacement - or registration - of the contour (and not on the image intensities) can then be used to map the intensities inside Γ_1 onto Γ_2 . After the mapping, the intensities inside Γ_1 and Γ_2 can be compared and then be classified as the same or different value. Since we can still find the contour in the occluded area, therefore we can also compute the displacement field even in the occluded area.

Algorithm 2 The algorithm for the contour matching**INPUT :** Contours Γ_1 and Γ_2 .**OUTPUT :** Displacement field D .

1. **Initial displacement field** Initial registration of the contours.
 2. **Diffusion** Convolve the displacement field using a Gaussian kernel.
 3. **Deformation** Deform Γ_1 by applying the displacement field D .
 4. **Projection** Project the deformed Γ_1 onto Γ_2 (i.e. find the closest point on the contour Γ_2).
 5. **Updating the displacement field** Update the displacement field according to matching points on the contour Γ_2 .
 6. **Convergence** Stop if the displacement field is stable, otherwise go to step 2.
-

Implementation. Occlusions are detected by comparing the predicted and the observed intensities inside the segmented object. Unfortunately the displacement field is not exact: it is an estimation of the contour displacement and simultaneously an interpolation of the displacement for pixels inside Γ_1 . The intensities in the deformed frame must be interpolated. The interpolation can either be done in the deformed (Lagrange) coordinate or in the original (Euler) coordinate. The next neighbor interpolation scheme in the Euler coordinate has been used. Both the deformed and the current frames are filtered using a low-pass filter to decrease differences due to the interpolation and to the displacement.

The deformed frame, $F_p^{Deformed}(x)$, and the current frame, $F_c(x)$, are compared pixel by pixel using some similarity measures. The absolute differences $|F_p^{Deformed}(x) - F_c(x)|$ are used in our experiments. Different similarity measures require different degree of low-pass filtering. A simple pixel by pixel similarity measure requires more filtering, while a patch based similarity measure may require less or none low-pass filtering. See Algorithm 3.

4 Experimental Results

Following the Algorithm 1, we implement the proposed model to segment a selected object with approximately uniform intensity frame-by-frame. The minimization of the functional is obtained by the gradient descent procedure (14) implemented in the level set framework outlined in Sect. 2.1. Since the Chan-

Algorithm 3 The algorithm for occlusion detection using the displacement field to predict the contents in the next frame inside a contour.**INPUT:** The previous frame F_p , the current frame F_c , displacement field D **OUTPUT:** Occlusion mask.

1. **Deformation** Deform F_p using displacement field D into $F_p^{Deformed}$.
 2. **Interpolation** Interpolate $F_p^{Deformed}$ to get intensities in each grid point.
 3. **Low-pass filtering** Low-pass filter the images $F_p^{Deformed}$ and F_c .
 4. **Similarity measure** Compare $F_p^{Deformed}$ and F_c inside contour Γ_2 using a similarity measure to get a similarity measure for each pixel.
 5. **Thresholding** Find occlusions by thresholding in the similarity measure image.
-

Vese segmentation model finds an optimal piecewise-constant approximation to an image, this model works best in segmenting object that has nearly uniform intensity.

The choice of the coupling constant γ is done manually. It 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. To choose a proper γ is rather problematic in segmentation of image sequences. Using strong prior can give good results when the occlusions occur, but when segmenting the image frame where occlusions do not occur, the results will be close to the prior.

In Fig. 1, we show the segmentation results for a nonrigid object in a synthetic image sequence, where occlusion (the gray bar) occurs. Another experiment on a human walking image sequence shown in Fig. 3 where an occlusion (the superposition of another person) occurs. In both experiments, the standard Chan-Vese method fails to segment the selected object when it reaches the occlusion (Top Row). The result can be improved by adding a frame-to-frame interaction term as proposed in (13) (Bottom Row). In these experiments, we use quite large γ to deal with occlusions. As we can see on the last frame in Fig. 3, the result is close to a similarity transformed of the prior although intensities in between the legs are different from the object.

As described in Sect. 3.1 and Sect. 3.2, occlusion can be detected and located. By using the segmenta-

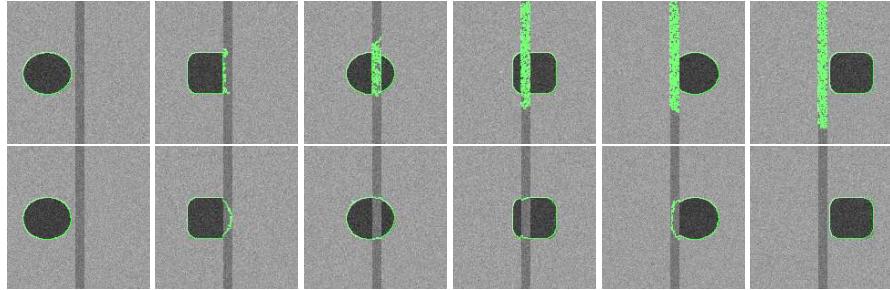


Figure 1: Segmentation of a nonrigid object in a synthetic image sequence with additive Gaussian noise. Top Row: without the interaction term, noise in the occlusion is captured. Bottom Row: with interaction term, we obtain better results.

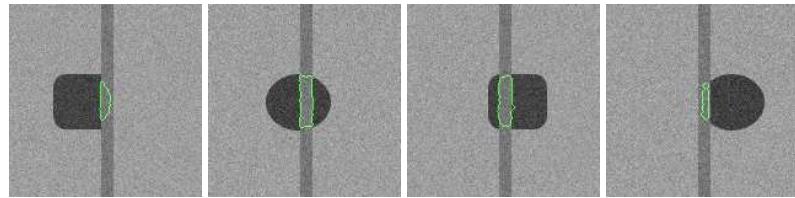


Figure 2: Detected occlusions in the synthetic image sequence.

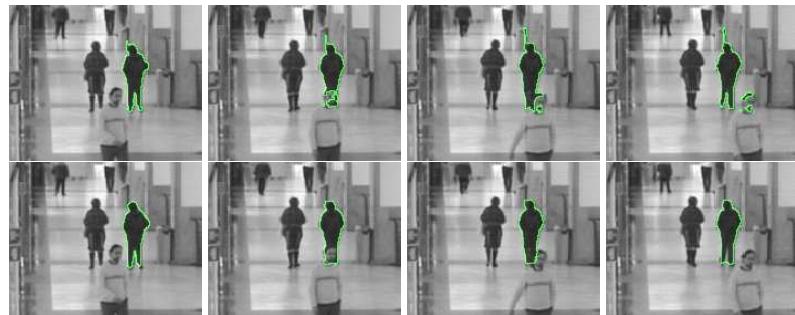


Figure 3: Segmentation of a walking person partly covered by an occlusion in the human walking sequence. Top Row: without interaction term, and Bottom Row: with interaction term

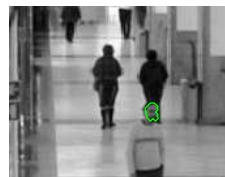


Figure 4: Detected occlusion in the human walking sequence.

tion results of the image sequences, we then implement the Algorithm 2 and 3 to detect and locate the occlusions. In Fig. 2 and Fig. 4, we show the occluded regions in the Frame 2-5 of Fig. 1 and in the Frame 2 of Fig. 3, respectively.

Having information about the location of the occlusions in the image, the occluded region can be reconstructed in order to improve further the segmen-

tation results. Let Occ be the occlusion masks, e.g. the output after implementing Algorithm 3. Here we reconstruct the occluded regions by assigning the intensity values in the occluded regions with the mean value of the intensities inside the contour but excluding the occluded regions:

$$I(Occ) = \mu_{int},$$

where

$$\mu_{\text{int}} = \mu_{\text{int}}(\Gamma) = \frac{1}{|\text{int}(\Gamma) \setminus \text{Occ}|} \int_{\text{int}(\Gamma) \setminus \text{Occ}} I(\mathbf{x}) d\mathbf{x}.$$

After we reconstruct the occluded regions, we implement the Algorithm 1 again by using smaller coupling constant γ in order to allow more deformation of the contours. As we can see from Fig. 5 and Fig. 6, the results are better if we reconstruct the occluded regions than the ones without reconstruction.

5 Conclusions

We have presented a method for segmentation and occlusion detection of image sequences containing nonrigid, moving objects. The proposed segmentation method is formulated as variational problem in the level set framework, 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. The segmentation results can then be used to detect the occlusions by the proposed method which is formulated as a variational contour matching problem. By using occlusion information, the segmentation can be further improved by reconstructing the occluded regions. Preliminary results are shown and its performance looks promising.

ACKNOWLEDGEMENTS

This research is funded by EU Marie Curie RTN FP6 project VISIONTRAIN (MRTN-CT-2004-005439). The human walking sequence was downloaded from EU funded CAVIAR project (IST 2001 37540) website.

REFERENCES

- Andresen, P. R. and Nielsen, M. (1999). Non-rigid registration by geometry-constrained diffusion. In Taylor, C. and et al, editors, *MICCAI'99*, LNCS 1679, pages 533–543. Springer Verlag.
- Bresson, X., Vanderghenst, P., and Thiran, J.-P. (2006). A variational model for object segmentation using boundary information and shape prior driven by the mumford-shah functional. *International Journal of Computer Vision*, 68(2):145–162.
- Caselles, V., Kimmel, R., and Sapiro, G. (1997). Geodesic active contours. *International Journal of Computer Vision*, 22(1):61–79.
- Chan, T. and Vese, L. (2001). Active contour without edges. *IEEE Transactions on Image Processing*, 10(2):266–277.
- Chan, T. and Zhu, W. (2005). Level set based prior segmentation. In *Proceeding CVPR 2005*, volume 2, pages 1164–1170.
- Chen, Y., Tagare, H. D., Thiruvenkadam, S., Huang, F., Wilson, D., Gopinath, K. S., Briggs, R. W., and Geiser, E. A. (2002). Using prior shapes in geometric active contours in a variational framework. *International Journal of Computer Vision*, 50(3):315–328.
- Cremers, D. and Funka-Lea, G. (2005). Dynamical statistical shape priors for level set based sequence segmentation. In Paragios, N. and et al., editors, *3rd Workshop on Variational and Level Set Methods in Computer Vision*, LNCS 3752, pages 210–221. Springer Verlag.
- Cremers, D. and Soatto, S. (2003). A pseudo-distance for shape priors in level set segmentation. In Faugeras, O. and Paragios, N., editors, *2nd IEEE Workshop on Variational, Geometric and Level Set Methods in Computer Vision*.
- Cremers, D., Sochen, N., and Schnörr, C. (2003). Towards recognition-based variational segmentation using shape priors and dynamic labeling. In Griffin, L. and Lillholm, M., editors, *Scale Space 2003*, LNCS 2695, pages 388–400. Springer Verlag.
- Gentile, C., Camps, O., and Sznaier, M. (2004). Segmentation for robust tracking in the presence of severe occlusion. *IEEE Transactions on Image Processing*, 13(2):166–178.
- Gustavsson, D., Fundana, K., Overgaard, N. C., Heyden, A., and Nielsen, M. (2007). Variational segmentation and contour matching of non-rigid moving object. In *ICCV Workshop on Dynamical Vision 2007*, LNCS. Springer Verlag.
- Konrad, J. and Ristivojevic, M. (2003). Video segmentation and occlusion detection over multiple frames. In Vasudev, B., Hsing, T. R., Tescher, A. G., and Ebrahimi, T., editors, *Image and Video Communications and Processing 2003*, SPIE 5022, pages 377–388. SPIE.

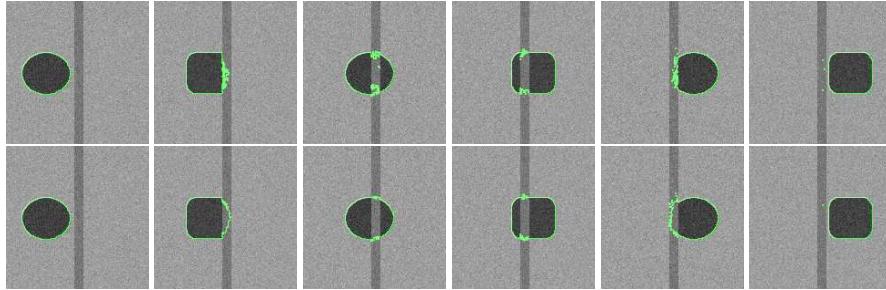


Figure 5: Segmentation of the synthetic image sequence by using smaller coupling constant than the one in Fig. 1. Top row: without reconstruction of the occluded regions. Bottom row: after the occluded regions are reconstructed.



Figure 6: Segmentation of the human walking sequence when by using smaller coupling constant than the one in Fig. 3. Top row: without reconstruction of the occluded regions. Bottom row: after the occluded region is reconstructed.

- Leventon, M., Grimson, W., and Faugeras, O. Statistical shape influence in geodesic active contours. In *Proc. Int'l Conf. Computer Vision and Pattern Recognition*, pages 316–323.
- Osher, S. and Fedkiw, R. (2003). *Level Set Methods and Dynamic Implicit Surfaces*. Springer-Verlag, New York.
- Paragios, N., Rousson, M., and Ramesh, V. (2003). Matching Distance Functions: A Shape-to-Area Variational Approach for Global-to-Local Registration. In Heyden, A. and et al, editors, *ECCV 2002*, LNCS 2351, pages 775–789. Springer-Verlag Berlin Heidelberg.
- Riklin-Raviv, T., Kiryati, N., and Sochen, N. (2007). Prior-based segmentation and shape registration in the presence of perspective distortion. *International Journal of Computer Vision*, 72(3):309–328.
- Rousson, M. and Paragios, N. (2002). Shape priors for level set representations. In Heyden, A. and et al, editors, *ECCV 2002*, LNCS 2351, pages 78–92. Springer Verlag.
- Solem, J. E. and Overgaard, N. C. (2005). A geometric formulation of gradient descent for variational problems with moving surfaces. In Kimmel, R., Sochen, N., and Weickert, J., editors, *Scale-Space 2005*, volume 3459 of *LNCS*, pages 419–430. Springer Verlag.
- Strecha, C., Fransens, R., and Gool, L. V. (2004). A probabilistic approach to large displacement optical flow and occlusion detection. In *Statistical Methods in Video Processing*, LNCS 3247, pages 71–82. Springer Verlag.
- Thiruvenkadam, S. R., Chan, T. F., and Hong, B.-W. (2007). Segmentation under occlusions using selective shape prior. In *Scale Space and Variational Methods in Computer Vision*, volume 4485 of *LNCS*, pages 191–202. Springer Verlag.
- Tsai, A., Yezzy, A., Wells, W., Tempany, C., Tucker, D., Fan, A., Grimson, W. W., and Willsky, A. (2003). A shape-based approach to the segmentation of medical imagery using level sets. *IEEE Transactions on Medical Imaging*, 22(2):137–154.