IEEE Workshop on Variational and Levelset Methods,
Vancouver, CA, Jul. 13 2001, © IEEE, to appear.

Diffusion–Snakes: Combining Statistical Shape Knowledge and Image Information in a Variational Framework

Daniel Cremers   Christoph Schnörr   Joachim Weickert

Computer Vision, Graphics, and Pattern Recognition Group
Department of Mathematics and Computer Science
University of Mannheim, 68131 Mannheim, Germany
http://www.cvgpr.uni–mannheim.de
{cremers, schnoerr, jwei}@uni–mannheim.de

Abstract

We present a modification of the Mumford–Shah functional and its cartoon limit which allows the incorporation of statistical shape knowledge in a single energy functional. We show segmentation results on artificial and real–world images with and without prior shape information. In the case of occlusion and strongly cluttered background the shape prior significantly improves segmentation. Finally we compare our results to those obtained by a level–set implementation of geodesic active contours.

Keywords: *image segmentation, shape recognition, statistical learning, variational methods, diffusion, geodesic active contours, diffusion–snake*

1. Introduction

Over the last years many variational approaches to image segmentation have been proposed. They make use of image information such as edges [6, 1, 9], homogeneity requirements on the statistics of the regions being segmented [14, 19] or a combination of both [15]. However, given large amounts of noise, clutter and occlusion, the image information by itself may not be sufficient to define a desired segmentation. Various efforts have been made to include prior information about the shape of the objects of interest in segmentation approaches [18, 16, 8].

Recently statistical shape knowledge has been introduced in a level set implementation of geodesic active contours [10]: a shape model is built over the distribution of surfaces which are obtained by embedding the training curves as zero level sets by the signed distance function. The surfaces are sampled at regular intervals and a principal component analysis is performed in this high–dimensional vector space of sampled surfaces. During segmentation the surfaces are evolved by the level set equation. An additional term is included which causes a relaxation towards the most probable shape. The latter is determined by a separate maximum a posteriori optimization of shape and pose parameters. Numerical results show very promising segmentation properties for 2D and 3D medical image data.

Our approach differs from [10] in several aspects: we incorporate statistical shape knowledge in the evolution process of a Mumford–Shah segmentation [14], which is not implemented as a level set evolution. We are aware that the level set approach has several favorable properties, the main one being the possibility to easily handle topological changes of the contour. On the other hand, level set approaches lead, by virtue of the embedding, to quite high-dimensional shape representations which is not preferable from the statistical learning point of view. Furthermore, there are numerous real–world applications such as silhouettes of many known objects where topological shape changes can be excluded *a priori*. For these reasons, and since we will focus on the statistical learning part in our future work, we prefer to work with low-dimensional explicit shape representations for the time being.

We propose an appropriate modification of the Mumford–Shah functional which allows an explicit parameterization of the contour as a closed spline curve. We then build a shape model on a set of training shapes by approximating the spline control point vectors by a Gaussian probability distribution. At this point, we do not perform a projection into the subspace of learnt contour deformations as is commonly done [2, 7, 10, 16]. Instead, we derive a shape energy which is defined on the full space of possible deformations. This allows to treat the shape energy term and the Mumford–Shah energy in one single energy functional. Instead of performing a two–step process of contour evolution and shape optimization, we can then perform a gradient descent on the full energy functional. The shape term in the resulting evolution equation is similar to that introduced in [10] for the level set evolution of geodesic active contours. However, in our case the relaxation towards the most probable shape is weighted by the inverse of a modified covariance matrix: deformation modes which are less familiar from the learning process will decay faster.

We introduced the general model of combining a spline–based implementation of the Mumford–Shah functional with statistical prior in [3]. The present paper extends this work in several ways:

We propose to modify the length norm of the contour in a way which simplifies the evolution equation and which strongly improves the spline–based formulation since it enforces an equidistant spacing of the spline control points. This turns out to be vital, in particular in the absence of prior knowledge, since the contour normal vector becomes ill–defined as soon as control points overlap.

Moreover, we present numerical studies of our methods in real–world images: segmentation results with and without prior knowledge in images with various degrees of noise and clutter.

We also present a similar implementation of the so–called *cartoon limit* of the Mumford–Shah functional [13, 14]. Numerical results exhibit interesting differences in the convergence properties of the two models.

Finally, we compare segmentation results of our spline–based Mumford–Shah implementation to those obtained by a level set implementation of geodesic active contours [1, 9].

2. Variational integration of shape statistics and segmentation

We propose to combine image information and previously acquired shape information in one variational framework. For a given contour C we define an energy

$$E(C) = E_i(C) + \alpha \cdot E_c(C)\,. \tag{1}$$

The functional E_i measures how well the contour segments the input grey value information. E_c favors contours which are familiar from a learning process. The parameter α allows to define the relative weight of the prior.

We propose for E_i a modification of the Mumford–Shah functional and its cartoon limit, which facilitates a parameterization of the contour as a closed spline curve, since shape learning and shape statistics are easily defined on the distribution of spline control points (see e.g. [2]).

2.1 Spline–based Mumford–Shah segmentation

The variational approach to image segmentation proposed by Mumford and Shah [14] consists in minimizing the following energy functional:

$$\begin{aligned} E_i(u, C) \;&=\; \frac{1}{2}\int_{\Omega} (f-u)^2\,dx \;+\; \lambda^2\,\frac{1}{2}\int_{\Omega - C} |\nabla u|^2\,dx \\ &\quad + \nu\,||C||\,. \end{aligned} \tag{2}$$

The input image f is approximated on the image plane Ω by a piecewise smooth function u, which may be discontinuous across a boundary C, its length being denoted by $||C||$.

In [3] we implemented this functional by representing the boundary as a closed spline curve:

$$C : [0,1] \to \Omega\,, \qquad C(s) = \sum_{n=1}^{N} \vec{p}_n\, B_n(s)\,, \tag{3}$$

where B_n are quadratic B–spline basis functions and $\vec{p}_n = (x_n\,,\, y_n)^t$ are the spline control points. Certainly this restricts the possible boundary formations, not allowing open boundaries, contour spitting, etc. However, it enabled us to incorporate learnt shape statistics. As is commonly done, we implemented the contour length in [3] by the $\mathcal{L}_1$–norm of C_s:

$$||C|| = \int_0^1 |C_s(s)|\,ds\,. \tag{4}$$

This produces a term proportional to the curvature in the evolution equation of the contour. In our framework of spline contours this term is not only computationally costly, but it also does not restrict the spline control points from clustering in one place. Once control points overlap, the normal becomes ill–defined. Since the contour is evolved along its normal, the process becomes instable. This problem only arises in cases where the shape prior is absent, since the latter restricts the control points to the learnt distribution, where control points are set equidistantly.

A slight modification of the original functional solves the problem of control point clustering: replacing the original length norm (4) by the squared $\mathcal{L}_2$–norm one obtains the modified functional

$$\begin{aligned} E_i(u, C) \;&=\; \frac{1}{2}\int_{\Omega} (f-u)^2\,dx \;+\; \lambda^2\,\frac{1}{2}\int_{\Omega - C} |\nabla u|^2\,dx \\ &\quad + \nu\,\mathcal{L}_2^2(C)\,, \end{aligned} \tag{5}$$

where

$$\mathcal{L}_2^2(C) = \int_0^1 C_s^2\,ds\,. \tag{6}$$

Further modifications of the Mumford–Shah functional with respect to length and curvature measures have been considered in [11].

Minimizing the squared $\mathcal{L}_2$–norm (6) with respect to C leads to an Euler–Lagrange equation of the simple form

$$C_{ss}(s) = 0\,, \quad \text{for all } s \in [0,1]. \tag{7}$$

For the quadratic B–spline curve this is equivalent to

$$\vec{p}_i = \frac{\vec{p}_{i-1} + \vec{p}_{i+1}}{2}\,, \quad i = 1, \ldots, N. \tag{8}$$

Therefore, by minimizing (5), each control point $\vec{p}_i$ tends to be centered between its two neighbors. This is what makes (5) well suited for the spline–based implementation. Moreover, experimental results show that the common argument in favor of the $\mathcal{L}_1$–norm (4), namely that it allows discontinuities in the boundary, is of a more theoretical nature since in our case a sufficiently fine parameterization allows the formation of arbitrarily sharp corners.

The same modification can be performed for the *cartoon model* [13, 14], which is obtained as $\lambda \to \infty$. Replacing the $\mathcal{L}_1$–norm by the $\mathcal{L}_2$–norm we obtain the modified cartoon limit:

$$E_i(C) = \frac{1}{2}\int_{\Omega_i} (f - u_i)^2 dx + \frac{1}{2}\int_{\Omega_o} (f - u_o)^2 dx + \nu \mathcal{L}_2^2(C). \tag{9}$$

During minimization of (9), the segmented image u is restricted to the constants u_i and u_o which will take on the mean value of f over the inside and the outside areas of the contour, denoted by Ω_i and Ω_o.

2.2 From learnt shape statistics to a shape energy

The explicit parameterization of the contour allows to represent a set of sample shapes in a vector space and to approximate their distribution statistically. To this end, the images of training objects are binarized, a spline contour is fit to the boundary and the set of training contours is aligned with respect to Euclidean similarity transformations [5] and cyclic permutation of the control points.

The distribution of control point vectors

$$\vec{z} = (x_1, y_1, \ldots, x_N, y_N)^t \tag{10}$$

is assumed to be Gaussian. The mean control point vector $\vec{z}_0$ and sample covariance matrix Σ are determined for the training set.

The advantage of assuming a Gaussian shape probability is that the associated shape energy has the following favorable property: the Mumford–Shah functional has by itself a non–convex dependency upon the contour. Adding the shape energy (12) will "convexify" the total energy. That is, given an arbitrary input image, the total energy (1) will be convex for sufficiently large α.

In general, Σ will not have full rank – especially if the number of sample shapes is smaller than the dimension $2N$ of the underlying vector space. The associated Gaussian probability will vanish for any shape outside the subspace spanned by the training shapes. In order to define a probability density which is continuous in the full $2N$–dimensional space, we artificially extend the Gaussian probability to the space orthogonal to the subspace of learnt deformations in the following way: replace all zero eigenvalues of the covariance matrix Σ by the smallest nonzero eigenvalue encountered. The inverse Σ^* of this matrix is well–defined and the associated probability distribution

$$\mathcal{P}\,(\vec{z}) \;\propto\; \exp\left(-\frac{1}{2}(\vec{z} - \vec{z}_0)^t\, \Sigma^*\, (\vec{z} - \vec{z}_0)\right) \tag{11}$$

is continuous on the full space. It gives a finite non–zero probability value for any shape $\vec{z}$. A similar regularization was proposed in [12]. A more intuitive explanation of this approach is this: given an unfamiliar shape variation, we associate a probability with it which is not zero but smaller or equal to the probability of any shape variation encountered in the learning process. This regularization accounts for the measurement error due to the finite number of training samples. The Gaussian probability (11) corresponds to the quadratic energy

$$E_c\,(\vec{z}) \;=\; \frac{1}{2}(\vec{z} - \vec{z}_0)^t\, \Sigma^*\, (\vec{z} - \vec{z}_0)\,, \tag{12}$$

which is similar to the Mahalanobis distance.

Note that this shape energy is automatically derived from a set of binary training images.

3. Energy minimization by gradient descent

The total energy (1) is iteratively minimized with respect to both the segmenting contour C and the segmented image u. Minimizing the modified Mumford–Shah functional (5) with respect to the contour C (for fixed u) results in the evolution equation

$$\frac{\partial C(s,t)}{\partial t} = (e^+ - e^-) \cdot \vec{n}_s + \nu\, C_{ss}(s,t)\,. \tag{13}$$

The terms e^+ and e^- denote the energy density [14]

$$e^{+/-} = (f - u)^2 + \lambda^2 (\nabla u)^2 \tag{14}$$

outside and inside the contour $C(s)$, respectively, and $\vec{n}$ denotes the outer normal vector on the contour.

For the modified cartoon limit (9), u is piecewise constant and the second term in (14) disappears.

Equation (13) can be converted to an evolution equation for the control points by inserting the definition (3) of the contour as a spline curve. Including the contribution of the shape energy (12) we obtain (for the x–coordinates of the control points):

$$\begin{aligned}\frac{dx_m(t)}{dt} &= \left(\mathbf{B}^{-1}\right)_{mi}\Big[\left(e^+(s_i,t)-e^-(s_i,t)\right)\vec{n}_x(s_i,t)\\ &\qquad\qquad -\nu\left(x_{i-1}-2x_i+x_{i+1}\right)\Big]\\ &\quad -\alpha\left[\Sigma^*\ (\vec{z}-\vec{z}_0)\right]_{2m-1}. \qquad (15)\end{aligned}$$

The equation has been discretized with a set of nodes s_i along the contour to obtain a set of linear differential equations, and the cyclic tridiagonal matrix $\mathbf{B}$ contains the spline basis functions evaluated at these nodes.

The terms in Equation (15) can be interpreted as follows: the first term forces the contour towards the object boundaries, the second term enforces an equidistant spacing of control points and the third term causes a relaxation towards the most probable shape. The index $2m-1$ is simply associated with the x–coordinate of control point m in the notation of (10). Note that the relaxation towards the most probable shape is weighted by the inverse of the modified covariance matrix, such that less familiar shape deformations will decay faster. This interesting property arises automatically due to the proposed variational integration of the prior.

As shown in [3], minimizing (5) with respect to the segmented image u results in an inhomogeneous linear diffusion of the form

$$\frac{\partial u}{\partial t} = -\frac{dE}{du} = (f-u) + \lambda^2\,\nabla\cdot\left(w_c(x)\,\nabla u\right), \quad (16)$$

where the contour information is encoded by the indicator function

$$w_c:\Omega\to\{0,1\},\quad w_c(x)=\begin{cases}0 & \text{if } x\in C\\ 1 & \text{otherwise}\end{cases}. \quad (17)$$

This underlying diffusion process is what gave rise to the term *diffusion snake*. Equation (16) was solved by a modified explicit scheme and the contour was evolved by iterating Equation (15) in alternation.

For the cartoon limit the diffusion process was replaced by an averaging process. The mean grey values for the two adjacent regions were dynamically updated, and the contour was evolved by Equation (15) in alternation.

4. Experimental results and discussion

In Section 4.1, we give an example of dealing with occlusion in an artificial image. In Section 4.2 we show the segmentation properties of our approach for a real–world image in the absence of a statistical prior, and in Section 4.3 we show how segmentation is improved by including a prior. In Section 4.4 we then present an example of a strongly cluttered background for which the desired segmentation is not obtained unless the statistical prior is included. The different convergence properties of the Mumford–Shah adaptation and its cartoon limit are analyzed in Section 4.5. Finally Section 4.6 offers a comparison to segmentation results obtained for the same images with a level set implementation of geodesic active contours.

4.1 Coping with occlusion

A common difficulty in segmentation problems is the fact that the object of interest may be partially occluded. In this case, the prior shape knowledge can help to recover missing information. Figure 1 shows the input image (black) of an ellipse occluded by a bar, the initial contour (left) and the final segmentation (middle) obtained with the cartoon model. The shape energy was constructed on a set of six ellipses. They were aligned as explained in Section 2.2. Before computing the statistics, one of the shapes was artificially translated in x– and y–direction. In this way, the two translatory degrees of freedom are included in the shape statistics [4].

The effect of this shape energy upon the curve evolution is that the contour deformations are restricted to ellipse–like shapes during the minimization process. Figure 1, right side, shows some intermediate steps in the contour evolution. Unfamiliar object features are ignored and the effective search space for the segmenting contour is drastically reduced.

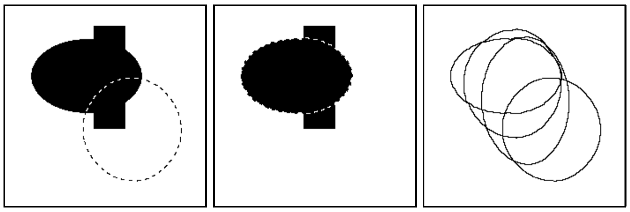

Figure 1. Left: input image and initial contour. Middle: segmentation result for the modified cartoon limit with a prior favoring ellipse–like shapes. Right: a number of intermediate contours indicates the contour evolution.

Similar results were obtained in [3] for the full Mumford–Shah model (2) with $\mathcal{L}_1$–norm on the contour. There we showed that, given a shape prior favoring bars, one is able to segment the bar instead of the ellipse.

4.2 Segmentation results with no prior

In this subsection we will demonstrate the segmentation properties of the two Mumford–Shah adaptations for closed spline curves, in the case when the prior information is com-

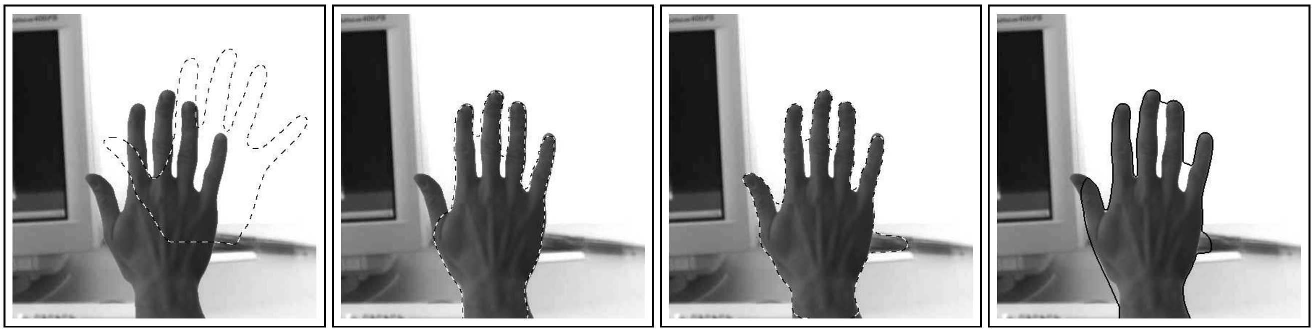

Figure 2. Segmentation with no prior. From left to right: initial contour, final segmentation for the modified Mumford–Shah, its cartoon limit, and a levelset implementation of geodesic active contours.

pletely switched off, i.e. $\alpha = 0$ in Equation (1). The goal is to segment a hand in a given image.

Both the full Mumford–Shah functional and its cartoon limit differ from many segmentation functionals such as the classical snake approach [6] in that they do not fit a contour to the local image gradient. It is the diffusion process or the averaging in case of the cartoon limit, which collects grey value information from the entire image plane. The effect is that the contour converges to the final segmentation over larger distances.

Figure 2, left side, shows a grey level image of a hand in front of a background and the initial contour (dashed line). The second image shows the final segmentation obtained with the modified Mumford–Shah model. Due to a large weight of the term minimizing the length of the contour, the thumb is cut off and the fingers are not fully segmented. If the parameter ν in (5) and (9) is decreased, the final contour is allowed to increase in length. The resulting segmentation is shown in the third image of Figure 2 for the cartoon limit. The hand is approximated better. However, some of the clutter in the background is included in the segmentation while the fingers are still not fully segmented. The right image shows a segmentation obtained by a level set implementation of geodesic active contours (see Section 4.6).

4.3 Knowledge–driven segmentation

In the above example – though it does not contain a lot of clutter – the final segmentation is not quite the desired one. The segmentations in Figure 2 were obtained without any statistical prior. In order to include prior information we constructed a shape energy upon a set of six binarized hand images as explained in Section 2.2, the hand in Figure 2 *not* being part of the training set. Again, the translatory degrees of freedom were enabled by shifting one of the training shapes in two orthogonal directions and including the obtained control point sets in the calculation of the covariance matrix.

We then performed a gradient descent upon the full energy (1) for the modified Mumford–Shah model in Equation (5) with the same initial contour as in Figure 2, left side. Figure 3 shows three steps in the contour evolution from the initialization (left) to the final segmentation (right).

The statistical prior effectively restricts the contour deformations to the subspace of learnt deformations. However, due to the embedding of the shape probability into the full space of possible deformations, as explained in Section 2.2, some deformation outside this subspace is still feasible – as can be seen in the intermediate steps in Figure 3. This *flexibility* turns out to strongly improve the ability of the system to evade incorrect local segmentations. The final segmentation is cut at the wrist, since the training shapes were all cropped there for simplicity.

The question of which value to assign to the length–governing parameter ν in Equations (5) and (9), discussed in Section 4.2, becomes obsolete: the effective restriction of shape deformations imposed by the prior allows to drop the additional length minimization term. However, for the purpose of analyzing the effect of the prior we kept the value of ν constant throughout all our experiments.

Segmentation results of equal quality as in Figure 3 were obtained by including statistical shape prior in the cartoon model.

4.4 Coping with clutter

The scene in Figure 2 contained little clutter. Therefore segmentation results were rather good even in the case when no prior knowledge was included. Once the amount of clutter is increased, this changes. Figure 4 shows an example of a hand in front of a strongly cluttered background. Note that the grey value of the background is approximately the same as that of the object of interest. Without the statistical prior, none of the segmentation approaches is able to extract the object of interest. Note that due to the underlying diffusion process the modified Mumford–Shah converges more

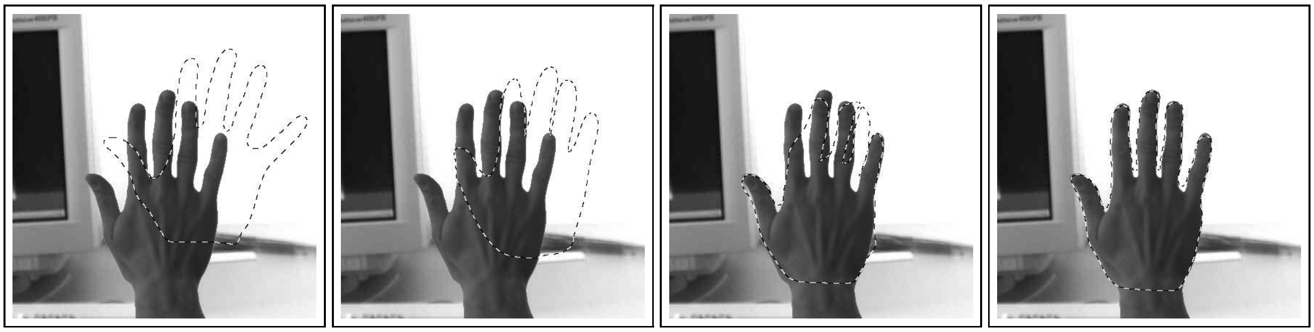

Figure 3. Segmentation with statistical prior: contour evolution from initial to final step.

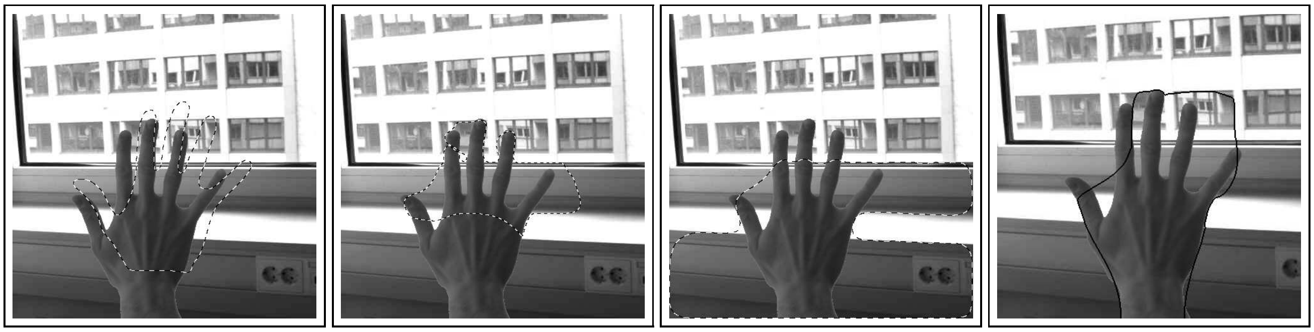

Figure 4. Segmentation with no prior in strongly cluttered background. From left to right: initial contour, segmentation results obtained by the modified Mumford–Shah, its cartoon limit and a level set scheme of geodesic active contours.

locally than its cartoon limit. This will be discussed in more detail in Section 4.5.

As in the previous example, we now include the shape prior and perform a gradient descent on the total energy (1) to obtain the segmentation shown in Figure 5, left side, for the case of the modified full Mumford–Shah functional. Again, the shape in the image was not part of the training set.

The final segmentation produced with the statistical prior is the desired one. Small discrepancies between the object boundary and the final contour in the area between the fingers are probably due to the fact that the shape prior does not fully suppress some shape variability in that area. This could be improved with a more elaborate alignment of the training shapes during shape learning. However, we prefer to explicitly avoid any shape learning that involves the calculation of landmarks or any manual interaction such as the labeling of correspondences.

The segmentation obtained with statistical prior in the case of the cartoon model was not successful, as can be seen in Figure 5, middle. The reason for this failure to capture the object of interest will be discussed next.

4.5 Comparing the modified Mumford–Shah model and its cartoon limit

The full Mumford–Shah functional and its cartoon limit differ in their contour evolution equation in that the former collects grey value information from the area surrounding the contour by means of a diffusion process, whereas the latter does this by separately averaging over the area inside and the area outside the present contour. Therefore, in the full Mumford–Shah model the contour motion is affected mostly by the image information in the neighborhood – at least for not too large values of λ in (5). The cartoon model, however, is equally affected by information in any part of the image. This explains the very different segmentation results obtained for the image in Figures 4 and 5 – both with and without prior.

The segmentation obtained with the cartoon model will be affected by grey value differences on a global scale. To analyze which effect this property has upon the energy landscape, we calculated the value of the full Mumford–Shah functional and its cartoon limit for a fixed contour which we simply translated in x– and y–direction. We used the same input image as in Figure 4. The contour was optimally placed upon the hand boundaries and then shifted in

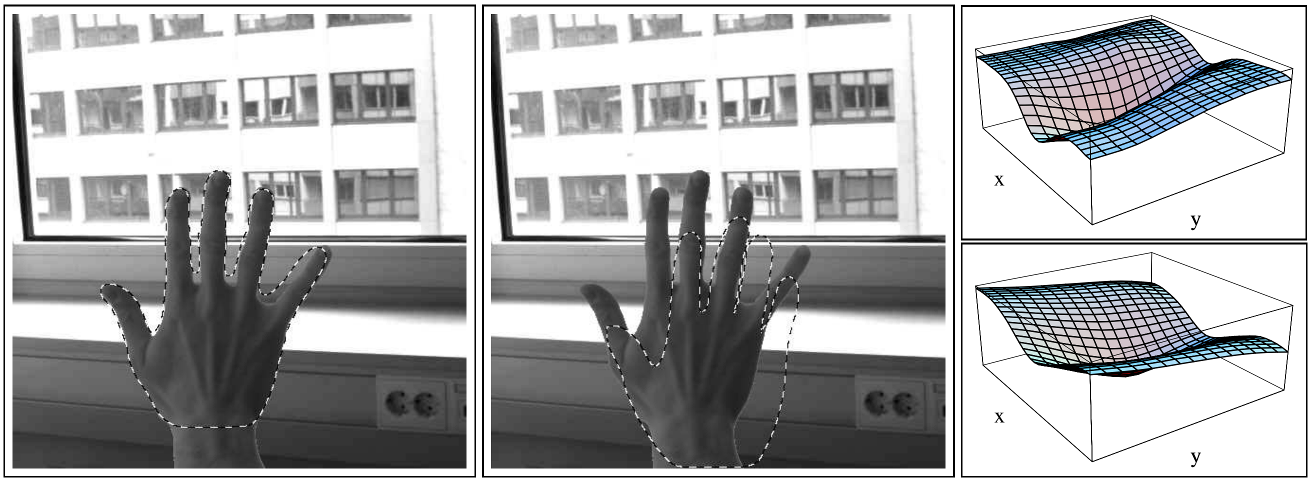

Figure 5. Left and middle: Segmentation with prior in strongly cluttered environment for the modified Mumford–Shah and its cartoon limit. Right: corresponding energy plots (see text).

equidistant steps up to 30 pixels in each direction. The resulting energies are plotted in Figure 5, right side, as a function of the displacement from the optimal position. Note that the bottom of the image corresponds to the top right side of the energy plots. Both energies show a minimum at the optimal position of the contour. However, the energy for the cartoon limit (below) is strongly slanted towards the bottom of the image. This is caused by the global change in brightness from the top of the image towards the bottom. This is what drives the contour to segment the entire bottom part of the image if no prior is given — cf. Figure 4. Even in the case of added shape prior the hand contour is pushed to the bottom of the image in the cartoon model – cf. Figure 5, middle.

4.6 Comparison with geodesic active contours and no shape prior

In order to compare our results to another segmentation approach, we performed a level set implementation of geodesic active contours. We opted for this comparison since the level set formulation of geodesic active contours is one of the most competitive among present segmentation methods. For the same input images f and the same initial contours C, we minimized the energy functional [1, 9]

$$E_f(C) = \int_0^1 g\left(\left|\nabla f_\sigma(C(s))\right|^2\right) |C_s(s)|\, ds\,, \qquad (18)$$

with Gaussian–smoothed input image f_σ, and metric [17]

$$g(s^2) = \begin{cases} 1, & \text{if } s^2 = 0 \\ 1 - \exp\left(-\frac{3.315}{(s/\lambda)^8}\right), & \text{if } s^2 > 0 \end{cases} \,. \qquad (19)$$

Here λ serves as contrast parameter.

We did not include any additional terms such as balloon forces since they assume a prior knowledge about whether the object of interest is inside or outside the initial contour. Moreover, the two segmentation approaches based on the Mumford–Shah functional do not contain any such term either.

Our geodesic active contour implementation used an efficient pyramid additive operator splitting (AOS) scheme that does not require to recalculate a distance transformation in each iteration [17].

The comparison in Figure 2 shows that the segmentation obtained by the Mumford–Shah based models and that obtained by the geodesic active contour model are similar for homogeneous background. However, the comparison with Figure 4 indicates that in a strongly cluttered background the geodesic contours give a more satisfactory approximation of the object of interest – indicating at least its approximate location.

One should however keep in mind, that the model formulations are conceptually very different: whereas the geodesic active contour model is directly governed by the gradient of the smoothed input image, this is different for the Mumford–Shah model – especially for the case of the cartoon limit, which is a region–based rather than an edge–based segmentation approach.

Moreover, in the case of the geodesic active contour model, the final contour is obtained as the zero level set of a higher dimensional surface. In our model formulation the final segmentation curve is obtained in form of a parameterized spline curve. The latter may be of advantage for further post–processing of the contour information.

5 Summary and conclusions

We presented an image segmentation approach which allows the integration of statistical shape knowledge in a single energy functional. To this end we modified the Mumford–Shah functional in a way which facilitates a spline representation of the contour. We proposed to adopt the $\mathcal{L}_2$–norm as a contour measure. This enforces an equidistant spacing of control points which turns out to be vital for a stable contour evolution.

We extracted the contours of a set of binary example shapes in terms of sets of spline control points in a way which involves no manual labeling of correspondences or landmark calculation. Based on a Gaussian approximation of the shape distribution we calculated a shape energy. By a regularization of the generally non–invertible covariance matrix, we embedded the shape energy from the subspace of learnt contour deformations into the full space of possible deformations. This allows to integrate image information and statistical prior in one variational framework. The contour evolution is simply given by a gradient descent on one energy functional rather than by iterating a two–step process of contour evolution and shape optimization.

We presented a thorough experimental study of the segmentation properties of our approach with and without prior knowledge in the case of real–world images. We showed how the statistical prior can help to deal with noise, occlusion and clutter. Moreover we presented results obtained for the so–called cartoon limit, both with and without clutter. Finally we compared our results to a level set implementation of geodesic active contours.

Our shape prior is invariant with respect to translation. This is not the case for scaling or rotation. Ongoing work is focused on integrating general invariance to similarity transformations into the variational approach presented above. Moreover, we are investigating models of nonlinear shape probability density beyond the Gaussian approximation.

Acknowledgments

We thank N. Ayache, J. Keuchel and C. Kervrann for helpful discussions.

References

[1] V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. In *Proc. IEEE Internat. Conf. on Comp. Vis.*, pages 694–699, Boston, USA, 1995.

[2] T.F. Cootes, C.J. Taylor, D.M. Cooper, and J. Graham. Active shape models – their training and application. *Comp. Vision Image Underst.*, 61(1):38–59, 1995.

[3] D. Cremers, C. Schnörr, J. Weickert, and C. Schellewald. Diffusion–snakes using statistical shape knowledge. In G. Sommer and Y.Y. Zeevi, editors, *Algebraic Frames for the Perception-Action Cycle*, volume 1888 of *Lect. Not. Comp. Sci.*, pages 164–174, Kiel, Germany, Sept. 10–11, 2000. Springer.

[4] D. Cremers, C. Schnörr, J. Weickert, and C. Schellewald. Learning of translation invariant shape knowledge for steering diffusion snakes. In G. Baratoff and H. Neumann, editors, *Dynamische Perzeption*, volume 9 of *Proceedings on Artificial Intelligence*, pages 117–122, Ulm Germany, Nov. 2000. Infix.

[5] C. Goodall. Procrustes methods in the statistical analysis of shape. *J. Roy. Statist. Soc., Ser. B.*, 53(2):285–339, 1991.

[6] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *Int. J. of Comp. Vision*, 1(4):321–331, 1988.

[7] C. Kervrann. *Modèles statistiques pour la segmentation et le suivi de structures déformables bidimensionnelles dans une séquence d'images*. PhD thesis, Université de Rennes I, France, 1995.

[8] C. Kervrann and F. Heitz. Statistical deformable model-based segmentation of image motion. *IEEE Trans. on Image Processing*, 8:583–588, 1999.

[9] S. Kichenassamy, A. Kumar, P.J. Olver, and A. Tannenbaum. Gradient flows and geometric active contour models. In *Proc. IEEE Internat. Conf. on Comp. Vis.*, pages 810–815, Boston, USA, 1995.

[10] M.E. Leventon, W.E.L. Grimson, and O. Faugeras. Statistical shape influence in geodesic active contours. In *Proc. Conf. Computer Vis. and Pattern Recog.*, volume 1, pages 316–323, Hilton Head Island, South Carolina, June 13–15, 2000.

[11] C. Mantegazza. *Su Alcune Definizioni Deboli di Curvatura per Insiemi Non Orientati*. PhD thesis, Dept. of Mathematics, SNS Pisa, Italy, 1993.

[12] B. Moghaddam and A. Pentland. Probabilistic visual learning for object representation. *IEEE Trans. Patt. Anal. Mach. Intell.*, 19(7):696–710, 1997.

[13] J.-M. Morel and S. Solimini. Segmentation of images by variational methods: a constructive approach. *Revista Matematica de la Universidad Complutense de Madrid*, 1(1,2,3):169–182, 1988.

[14] D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. Pure Appl. Math.*, 42:577–685, 1989.

[15] N. Paragios and R. Deriche. Coupled geodesic active regions and level set methods for image segmentation. *IEEE Trans. on Image Processing*. To appear.

[16] Y. Wang and L. H. Staib. Boundary finding with correspondence using statistical shape models. In *Proc. Conf. Computer Vis. and Pattern Recog.*, pages 338–345, Santa Barbara, California, June 1998.

[17] J. Weickert. Applications of nonlinear diffusion filtering in image processing and computer vision. *Acta Mathematica Universitatis Comenianae*. To appear.

[18] S.C. Zhu and D. Mumford. Prior learning and Gibbs reaction–diffusion. *IEEE Trans. Patt. Anal. Mach. Intell.*, 19(11):1236–1250, 1997.

[19] S.C. Zhu and A. Yuille. Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation. *IEEE Trans. Patt. Anal. Mach. Intell.*, 18(9):884–900, 1996.