International Journal of Computer Vision 50(3): 295-313; Dec 2002

Diffusion Snakes: Introducing Statistical Shape Knowledge into the Mumford-Shah Functional

Daniel Cremers, Florian Tischhäuser, Joachim Weickert* and Christoph Schnörr
Computer Vision, Graphics, and Pattern Recognition Group
Department of Mathematics and Computer Science
University of Mannheim, D-68131 Mannheim, Germany
http://www.cvgpr.uni-mannheim.de
{*cremers, floriant, jwei, schnoerr*}*@uni-mannheim.de*

Abstract. We present a modification of the Mumford-Shah functional and its cartoon limit which facilitates the incorporation of a statistical prior on the shape of the segmenting contour. By minimizing a single energy functional, we obtain a segmentation process which maximizes both the grey value homogeneity in the separated regions and the similarity of the contour with respect to a set of training shapes. We propose a closed-form, parameter-free solution for incorporating invariance with respect to similarity transformations in the variational framework. We show segmentation results on artificial and real-world images with and without prior shape information. In the cases of noise, occlusion or 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 snake, geodesic active contours

1. Introduction

Over the last years many variational approaches to image segmentation have been proposed. They make use of image information such as edges [17, 4, 20], homogeneity requirements on the statistics of the regions being segmented [25, 38, 5, 35] or a combination of both [26]. However, given large amounts of noise, clutter and occlusion, the information contained in the image 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 [16, 37, 31, 19].

Recently, statistical shape knowledge has been introduced in a level set implementation of geodesic active contours [21]: 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

* Present address: Department of Mathematics and Computer Science, Saarland University, D-66041 Saarbrücken, Germany

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 [21] in several aspects: We incorporate statistical shape knowledge in the evolution process of a Mumford-Shah based segmentation [25], 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 a modification of the Mumford-Shah functional which allows an explicit parameterization of the contour as a closed spline curve. In particular, we modify the usual length constraint on 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 modification turns out to be vital, since the contour normal vector becomes ill-defined as soon as control points overlap. As a result, we obtain the *diffusion snake*, a hybrid model which combines the external energy of the Mumford-Shah functional with the internal energy of the snakes [17].

We then build a statistical shape model on a set of training shapes by approximating the corresponding 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 [7, 18, 21, 31]. Instead, we derive a shape energy with support on the full space of possible contour deformations. This allows to treat the shape energy term and the Mumford-Shah energy in one single energy functional. Rather than iterating a two-step process of contour evolution and optimization of the shape dissimilarity, a knowledge-driven segmentation process is obtained by gradient descent on a single functional. The shape term in the resulting evolution equation is similar to that introduced in [21] 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.

Moreover, we incorporate invariance of the shape prior with respect to similarity transformations of the contour in the variational approach. We derive a closed-form, parameter-free solution by aligning the evolving contour with the training shapes before applying the statistical shape energy. Consequently, the evolving contour is restricted to a submanifold of familiar shapes, while being free to translate, rotate and scale.

We present numerical studies of our methods in real-world images: Segmentation results with and without prior knowledge demonstrate the capacity of the shape prior to compensate for missing or misleading information due to noise, clutter and occlusion. We also present a similar implementation of the *cartoon limit* of the Mumford-Shah functional [24, 25]. Numerical results exhibit interesting differences in the convergence properties of the two models. We compare segmentation results of our spline-based Mumford-Shah implementation to those obtained by a level set implementation of geodesic active contours [4, 20]. Finally, we present two different numerical approximations of the diffusion process underlying the curve evolution. This paper comprises and extends work which has been presented on scientific meetings [11, 10].

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

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

The functional E_i measures how well the contour and the associated segmentation u approximate 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.

In the following, 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. Shape learning and shape statistics are then conveniently defined on the distribution of spline control points.

2.1. Spline-based Mumford-Shah Segmentation

The variational approach to image segmentation proposed by Mumford and Shah [25] consists in minimizing the following energy functional:

$$E_i(u, C) \;=\; \frac{1}{2}\int\limits_{\Omega}(f-u)^2\,dx \;\;+\;\; \lambda^2\,\frac{1}{2}\int\limits_{\Omega-C}|\nabla u|^2\,dx \;\;+\;\; \nu\,||C||. \qquad (2)$$

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 $||C||$.

We implement this functional by representing the boundary as a closed spline curve:

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

where B_n are periodic, quadratic B-spline basis functions [14, 2] and $p_n = (x_n\,,\,y_n)^t$ are the spline control points. Certainly this parametric representation restricts the class of permissible contours, not allowing open boundaries, and contour splitting or merging. However, it facilitates the generation and incorporation of a statistical prior on the shape of the boundary, which will be done in Section 2.2.

As is commonly done, we originally implemented the contour length in [11] by the L_1-norm of C_s:

$$||C|| = \int\limits_0^1 |C_s(s)|\,ds. \qquad (4)$$

This produces a term proportional to the curvature in the evolution equation of the contour. However, in our framework of spline contours this term 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.

We therefore replace the original length norm (4) by the squared L_2-norm, to obtain the functional for the *diffusion snake*

$$\textbf{(DS)} \quad E_i(u, C) = \frac{1}{2}\int\limits_{\Omega}(f-u)^2\,dx \;+\; \frac{1}{2}\lambda^2\int\limits_{\Omega-C}|\nabla u|^2\,dx \;+\; \nu\,\mathrm{L}_2^2(C), \quad (5)$$

where

$$\mathrm{L}_2^2(C) = \int\limits_0^1 C_s^{\,2}\,ds. \qquad (6)$$

The diffusion snake (5) can be considered a hybrid model which combines the external energy of the Mumford-Shah functional with the internal energy of the snakes [17]. Further modifications of the Mumford-Shah functional with respect to length and curvature measures have been considered in [22].

Minimizing the squared 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

$$p_i = \frac{p_{i-1} + p_{i+1}}{2}, \quad i = 1, \dots, N. \tag{8}$$

Therefore, by minimizing (5), each control point p_i tends to be centered between its two neighbors. This is what makes (5) well suited for our spline-based implementation. Moreover, we experimentally verified that a sufficiently fine parameterization enables the formation of arbitrarily sharp corners.

The same modification can be performed for the *cartoon model* [24, 25], which is obtained as $\lambda \to \infty$. Replacing the L_1-norm by the L_2-norm, we obtain the *simplified diffusion snake*:

$$\textbf{(SDS)} \qquad E_i(\{u_k\}, C) = \frac{1}{2} \sum_k \int_{\Omega_k} (f - u_k)^2 dx + \nu \, L_2^2(C). \tag{9}$$

During minimization of (9), the segmented image u is restricted to the constants $\{u_k\}$ which will take on the mean grey value of the input image f on the set of regions $\{\Omega_k\}$ separated by the contour C.

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 similarity transformations [15, 7] and cyclic permutation of the control points.

The distribution of control point vectors

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

is assumed to be Gaussian. The mean control point vector μ 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 quadratic shape energy (13) 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 with support 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:

Let D be the diagonal matrix of the ordered non-zero eigenvalues $\sigma_1 \geq \ldots \geq \sigma_r > 0$ of the covariance matrix Σ, and V the matrix of corresponding eigenvectors. The covariance matrix is then regularized by replacing all zero eigenvalues by an appropriate constant $\sigma_\perp > 0$:

$$\Sigma_\perp = V\, D\, V^t + \sigma_\perp \left(I - V V^t \right). \tag{11}$$

The inverse of this matrix is well-defined, and the corresponding probability distribution

$$\mathcal{P}(z) \propto \exp\left(-\frac{1}{2} (z-\mu)^t\, \Sigma_\perp^{-1}\, (z-\mu) \right) \tag{12}$$

is differentiable on the full space. It associates a finite non-zero value with any shape z.

The Gaussian probability (12) corresponds to the quadratic energy

$$E_c(z) = -\log(\mathcal{P}(z)) + \text{const.} = \frac{1}{2} (z-\mu)^t\, \Sigma_\perp^{-1}\, (z-\mu), \tag{13}$$

which is similar to the Mahalanobis distance. Note that this shape energy is automatically derived from a set of binary training images.

2.3. On Regularizing the Covariance Matrix

What is the optimal value for the regularizing constant $\sigma_\perp$ in (11)? A similar regularization of the covariance matrix was proposed in [23, 27, 29]. However, there the underlying purpose is quite different. Namely in [23, 29] the authors intend to obtain a sparse representation by replacing the smallest eigenvalues $\sigma_{r+1}, \ldots, \sigma_{2N}$ by a constant $\sigma_\perp$. By minimizing the Kullback-Leibler distance between the two probability distributions – corresponding to the covariance matrices with and without this replacement – they obtain an optimal value for $\sigma_\perp$, given by the mean of the replaced eigenvalues:

$$\sigma_\perp = \frac{1}{2N-r} \sum_{i=r+1}^{2N} \sigma_i \, . \tag{14}$$

We believe this estimate not to be appropriate in our context. The implicit assumption underlying this derivation is that the covariance matrix estimated from the data is the one corresponding to the true probability distribution. This, however, is not the case since we only have a finite – and in our case even fairly small – number of training samples.

Based on a more intuitive explanation of the regularization, we therefore propose a different estimate of $\sigma_\perp$: Given an unfamiliar shape variation (corresponding to a zero eigenvalue of Σ), we associate a probability with it which is not zero, but which is smaller or equal to the probability of any shape variation encountered in the learning process. This is equivalent to the bound: $0 < \sigma_\perp \leq \sigma_r$, where σ_r is the smallest non-zero eigenvalue of Σ. Lacking a more precise estimate, we fixed

$$\sigma_\perp = \sigma_r/2, \tag{15}$$

which worked well for all our experiments. Moreover, we found in numerical experiments that the curve evolution is not very sensitive to the exact value of $\sigma_\perp$.

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.

3.1. Curve Evolution

Minimizing the diffusion snake (**DS**) functional (5) with respect to the contour C (for fixed u) leads to the Euler-Lagrange equation

$$\frac{\partial E_i}{\partial C} = \left[e^-(s) - e^+(s)\right] \cdot n(s) - \nu \, C_{ss}(s) = 0 \quad \forall s \in [0,1]. \tag{16}$$

The terms e^+ and e^- denote the energy density

$$e^{+/-} = (f-u)^2 + \lambda^2 (\nabla u)^2 \tag{17}$$

in the regions adjoining the contour $C(s)$, and n denotes the outer normal vector on the contour. For the simplified diffusion snake (**SDS**) (9), u is piecewise constant and the second term in (17) disappears.

Solving the minimization problem by gradient descent results in the evolution equation

$$\frac{\partial C}{\partial t} = -\frac{\partial E_i}{\partial C} = \left[e^+(s,t) - e^-(s,t)\right] \cdot n(s,t) + \nu\, C_{ss}(s,t) \quad \forall s, \quad (18)$$

where an artificial time parameter t has been introduced.

Equation (18) 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 (13), we obtain for the coordinates of control point m:

$$\begin{aligned} \frac{dx_m(t)}{dt} &= \sum_i \left(\mathbf{B}^{-1}\right)_{mi} \Big[\left(e^+(s_i,t) - e^-(s_i,t)\right) n_x(s_i,t) \\ &\qquad\qquad\qquad + \nu\left(x_{i-1} - 2x_i + x_{i+1}\right)\Big] \\ &\quad - \alpha \left[\Sigma_\perp^{-1}\,(z-\mu)\right]_{2m-1}, \\ \frac{dy_m(t)}{dt} &= \sum_i \left(\mathbf{B}^{-1}\right)_{mi} \Big[\left(e^+(s_i,t) - e^-(s_i,t)\right) n_y(s_i,t) \\ &\qquad\qquad\qquad + \nu\left(y_{i-1} - 2y_i + y_{i+1}\right)\Big] \\ &\quad - \alpha \left[\Sigma_\perp^{-1}\,(z-\mu)\right]_{2m}. \end{aligned} \quad (19)$$

The equation has been discretized with a set of nodes s_i along the contour to obtain a set of linear differential equations. The cyclic tridiagonal matrix $\mathbf{B}$ contains the spline basis functions evaluated at these nodes: $B_{ij} = B_i(s_j)$, where s_i corresponds to the maximum of B_i.

The three terms in the respective equations in (19) 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 indices $2m-1$ and $2m$ are simply associated with the x- and y-coordinates 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.

3.2. Inhomogeneous Linear Diffusion

In order to minimize the **DS** (5) with respect to the segmented image u, we rewrite the functional in the following way:

$$E_i(u, C) = \frac{1}{2} \int_\Omega (f-u)^2 dx \; + \; \lambda^2 \, \frac{1}{2} \int_\Omega w_c(x) |\nabla u|^2 dx \; + \; \nu \, \mathrm{L}_2^2(C). \quad (20)$$

The contour dependence is now implicitly represented by an indicator function

$$w_c : \Omega \to \{0, 1\}, \qquad w_c(x) = \begin{cases} 0 & \text{if } x \in C \\ 1 & \text{otherwise} \end{cases}. \quad (21)$$

The Euler-Lagrange equation corresponding to this minimization problem is given by:

$$\frac{1}{\lambda^2} \frac{dE}{du} \; = \; \frac{1}{\lambda^2}(u-f) \; - \; \nabla \cdot (w_c(x) \, \nabla u) = 0. \quad (22)$$

This equation corresponds to the steady state of the following diffusion process:

$$\frac{\partial u}{\partial t} \; = \; \nabla \cdot (w_c \, \nabla u) \; + \; \frac{1}{\lambda^2} \, (f-u), \quad (23)$$

in which the contour enters as an inhomogeneous diffusivity defining a boundary to the diffusion process. This underlying diffusion process is what gave rise to the term *diffusion snake*.

Two different schemes have been used to approximate the diffusion process: A simple explicit approximation to the diffusion equation (23), and a more sophisticated multigrid scheme solving the corresponding steady state equation (22). Both schemes are not straightforward because the strongly inhomogeneous coefficient function w_c has to be taken into account. Standard implementations may easily lead to diffusion across the discontinuity curve C and to a slowdown of convergence, in particular in the case of multigrid iteration. The technical details concerning these two schemes are given in Appendix A.

In the case of the cartoon limit (**SDS**), the diffusion process is replaced by an averaging process, such that the image u takes on the mean grey values over a set of regions:

$$u_k \; = \; \frac{1}{|\Omega_k|} \int_{\Omega_k} f \, dx. \quad (24)$$

These are dynamically updated during the evolution of the contour.

4. Invariance in the Variational Framework

By construction, the shape energy $E_c(z)$ in (13) is not invariant with respect to similarity transformations of the shape z. Such an invariance could be introduced by replacing E_c with the energy

$$E_{shape}(z) = \min_{s, R_\theta, t} E_c(s\, R_\theta\, (z - t\,)), \tag{25}$$

where minimization is done with respect to scale, rotation and translation. It appears infeasible to obtain an analytic solution (as a function of z) for the above minimization problem. Moreover, even a numerical optimization (cf. [6]) is problematic as discussed in 5.7.

Therefore, we propose to eliminate the rotation angle and the scale in the following way: Since the training shapes were aligned to their mean shape μ with respect to translation, rotation and scaling and then normalized to unit size, we shall do the same to the argument z of the shape energy before calculating the energy E_c. We eliminate the translation by centering the control point vector:

$$z_c = \left(I_{2N} - \frac{1}{N} T \right) z, \tag{26}$$

where I_{2N} denotes the unit matrix, N is the number of control points, and the $2N \times 2N$-matrix T is given by:

$$T = \begin{pmatrix} 1 & 0 & 1 & 0 & \cdots \\ 0 & 1 & 0 & 1 & \cdots \\ 1 & 0 & 1 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{27}$$

Next, we eliminate rotation and scaling by aligning with respect to the mean of the training data. The final shape energy is then given by:

$$E_{shape}(z) = E_c(\hat{z}), \quad \text{with } \hat{z} = \frac{R_\theta\, z_c}{|R_\theta\, z_c|}, \tag{28}$$

where θ denotes the angle corresponding to the optimal rotation of the control point polygon z_c with respect to the mean shape μ. We shall not go into detail about the derivation of R_θ. Derivations can be found in [33, 13]. The final result is given by the formula:

$$\hat{z} = \frac{M\, z_c}{|M\, z_c|}, \quad \text{with } M = I_N \otimes \begin{pmatrix} \mu^t\, z_c & -\mu \times z_c \\ \mu \times z_c & \mu^t\, z_c \end{pmatrix}, \tag{29}$$

where $\otimes$ denotes the Kronecker product and $\mu \times z_c := \mu^t R_{\pi/2}\, z_c$.

In order to determine the gradient of the similarity invariant shape energy in (28), we will denote the differentiation of vector-valued functions $f : I\!R^n \to I\!R^m$ by:

$$\frac{df}{dx} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}. \tag{30}$$

To incorporate the similarity invariant shape prior into the segmentation process, we simply replace the last term in the evolution equation (19) by the gradient on the similarity invariant shape energy (28). With results (29) and (26), this is obtained by applying the chain rule:

$$\frac{dE_{shape}(z)}{dz} = \frac{dE_c(\hat{z})}{d\hat{z}} \cdot \frac{d\hat{z}}{dz} = \frac{dE_c(\hat{z})}{d\hat{z}} \cdot \frac{d\hat{z}}{dz_c} \cdot \frac{d\hat{z}_c}{dz}. \tag{31}$$

The three terms in this product can be interpreted as follows:

- The first term is the gradient of the original energy evaluated for the aligned shape $\hat{z}$. It contains the shape information extracted from the training set. For the quadratic energy (13), it is given by:

$$\frac{dE(\hat{z})}{d\hat{z}} = \left(\Sigma_{\perp}^{-1} (\hat{z} - \mu)\right)^t. \tag{32}$$

  It causes a relaxation of the aligned shape $\hat{z}$ towards the mean shape μ, weighted by the inverse of the regularized covariance matrix. This weighting causes unfamiliar deformations from the mean to decay faster.

- The second term in the product of (31) takes into account the influence of changes in the centered shape z_c onto the aligned shape $\hat{z}$. In matrix notation it is given by:

$$\frac{d\hat{z}}{dz_c} = \frac{M' z_c + M}{||M z_c||} - \frac{(M z_c)(M z_c)^t (M' z_c + M)}{||M z_c||^3}, \tag{33}$$

  where M is the matrix defined in (29) and M' denotes the tensor of rank 3 given by:

$$M' = \frac{dM}{dz_c}. \tag{34}$$

The entries of this constant sparse tensor are given by:

$$M'_{ijk} = \frac{dM_{ik}}{d(z_c)_j} = \begin{cases} \mu_j, & i = k \\ \mu_{j+1}, & i = k+1,\ i \text{ even},\ j \text{ odd} \\ -\mu_{j-1}, & i = k+1,\ i \text{ even},\ j \text{ even} \\ -\mu_{j+1}, & i = k-1,\ i \text{ odd},\ j \text{ odd} \\ \mu_{j-1}, & i = k-1,\ i \text{ odd},\ j \text{ even} \\ 0, & \text{otherwise} \end{cases} . \tag{35}$$

- The third term in the product of (31) accounts for the change of the centered shape z_c with the input shape z. According to definition (26), it is given by:

  $$\frac{dz_c}{dz} = \left(I_{2N} - \frac{1}{N} T \right) . \tag{36}$$

  This term centers the energy gradient, as a direct consequence of the translation invariance of the shape energy. Accordingly, the force which minimizes the shape energy has no influence on the translation of the contour. Similarly, rotation and scaling of the shape are not affected by the shape optimization, due to the term (33) in the evolution equation (31).

Note that, due to the analytical minimization with respect to similarity transformations based on the control point representation of the spline curve, no additional parameters enter the above evolution equation to account for scale, rotation and translation. The advantages of this approach over modeling explicit parameters for translation, scale and rotation – cf. [6] – will be discussed in Section 5.7.

In a given segmentation task, invariance with respect to the group of similarity transformations may be too general. Obviously, one can – depending on the specific application – incorporate only a subgroup of the suggested invariances, such as translation invariance. Moreover, since there exists a similar closed form solution for the optimal alignment of two polygons with respect to the more general affine group [33], the above approach could be extended to define a shape prior which is invariant with respect to affine transformations. However, we will not elaborate this possibility.

5. Experimental Results and Discussion

In Section 5.1, we give an example of dealing with occlusion in an artificial image. In Section 5.2, we show the segmentation properties of our approach for a real-world image in the absence of a statistical prior, and in Section 5.3, we show how segmentation is improved by including a prior. In Section 5.4, 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 diffusion snake and its cartoon limit are analyzed in Section 5.5. Section 5.6 offers a comparison to segmentation results obtained for the same images with a level set implementation of geodesic active contours. In Section 5.7, we present results on rotation and scale invariance and discusses the advantages of the parameter-free implementation over other approaches. In Section 5.8, we investigate the potential and limitations of the shape prior in dealing with occlusion in real-world images. In Section 5.9, we demonstrate the capacity of the shape prior to cope for missing information due to noise.

5.1. 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 basic principle: Depicted are the input image (black) of an ellipse occluded by a bar, the initial contour, and the final segmentation obtained with the cartoon model (**SDS**), both without and with a statistical prior. The statistical shape energy was constructed from a set of six ellipses, as explained in Section 2.2. In order to demonstrate more clearly the effect of the prior on the curve deformation, only translation invariance was included into the shape energy – see Section 4 for details.

Without a statistical shape prior, the contour segments the entire black region corresponding to the object and the occlusion. Upon introduction of the shape prior, however, the evolving contour is restricted to ellipse-like shapes during the minimization process. Figure 1, right side, shows some intermediate steps in the contour evolution: The effective search space for the segmenting contour is drastically reduced, and unfamiliar object features are ignored.

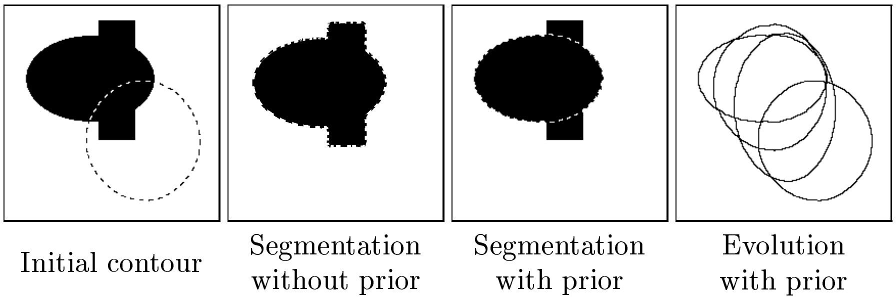


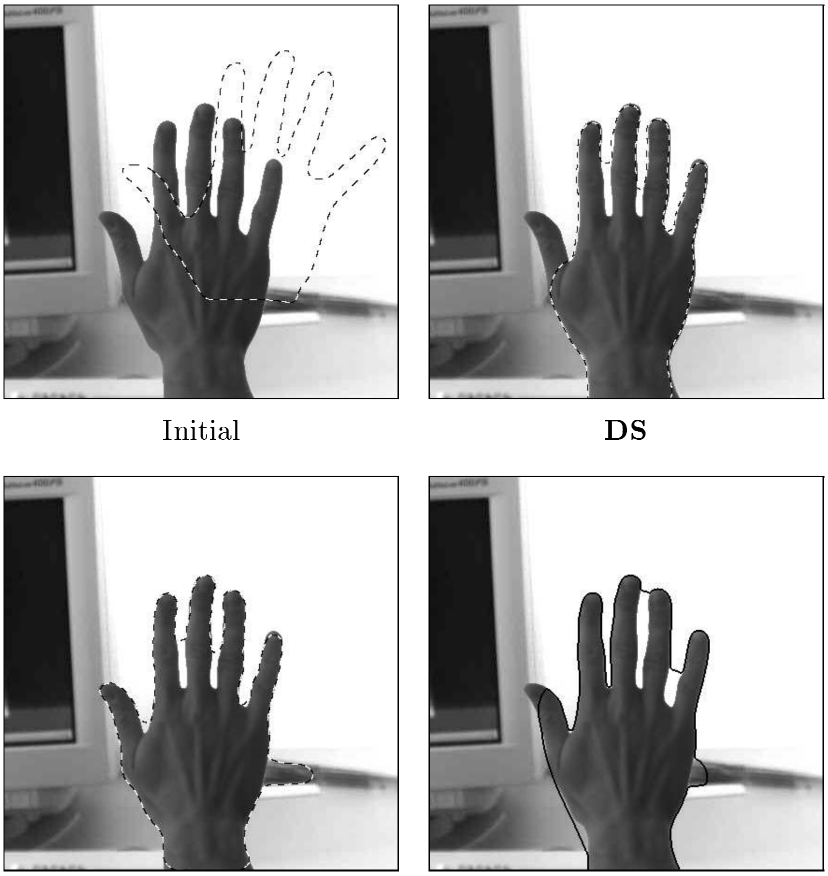
Figure 1. Segmentation results for the **SDS** with a prior favoring ellipse-like shapes. A number of intermediate contours (right) indicates how the contour evolution is restricted to the submanifold of familiar shapes. In the final segmentation with prior, the occluding bar is removed.

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

Both the diffusion snake (**DS**) and its cartoon limit (**SDS**) differ from many segmentation functionals such as the classical snake approach [17] 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. Yet, since the functional is not defined in terms of a presmoothed input image, details of the image information such as the precise location of corners and edges are preserved.

Figure 2, left side, shows a grey level image containing a hand and the initial contour (dashed line). The second image shows the final segmentation obtained with the **DS**. 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 **SDS**. 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 5.6).



Initial   **DS**

SDS   Geodesic Active Contour

Figure 2. Segmentation with no prior. Initial contour (top left), final segmentation for the **DS**, the **SDS**, and a level set implementation of geodesic active contours. Depending on the size of the length constraint, part of the background clutter will be incorporated into the segmentation, although details such as part of the fingers are not yet fully segmented.

5.3. 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, in order to gradually increase the model complexity, we only included translation invariance.

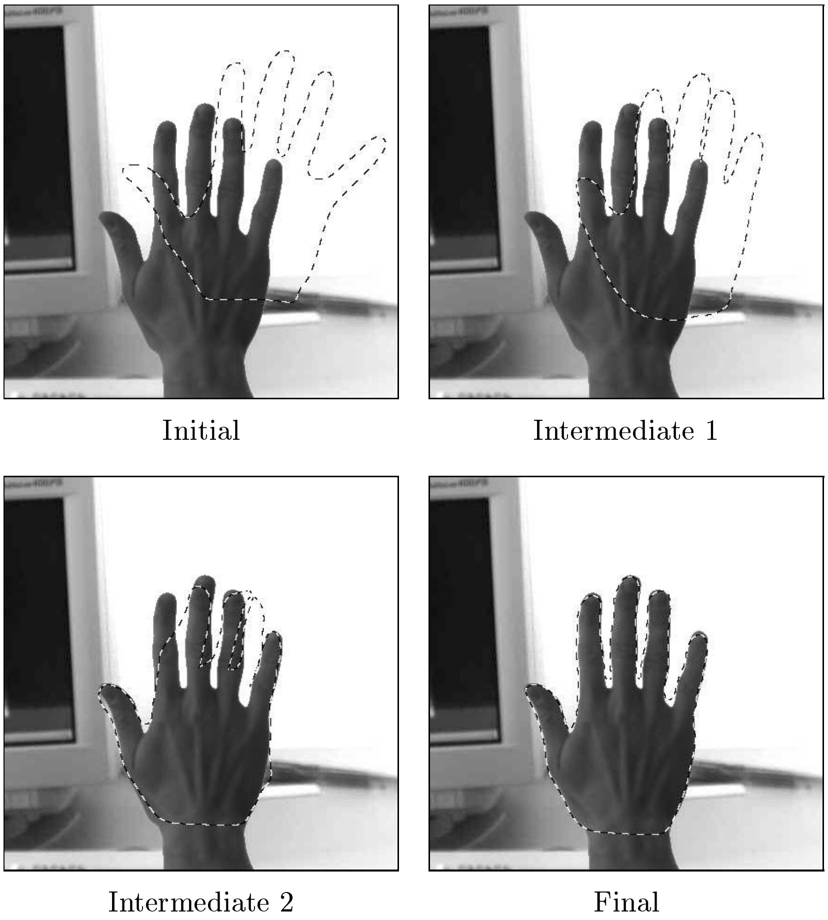


Figure 3. Contour evolution for the diffusion snake with a statistical shape prior. Due to the prior, the evolving contour is effectively restricted to a submanifold of familiar shapes. In contrast to the result without prior in Figure 2, the background clutter is not included in the segmentation, and the fingers are fully segmented.

The training shapes all had the same rotation and scale as the object in the image. Results including scale and rotation invariance will be shown separately later on.

We then performed a gradient descent upon the full energy (1) for the **DS** in equation (5) with the same initial contour as in Figure 2, left side. Figure 3 indicates the contour evolution from the initialization (top left) to the final segmentation (bottom 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 5.2, 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 these experiments.

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

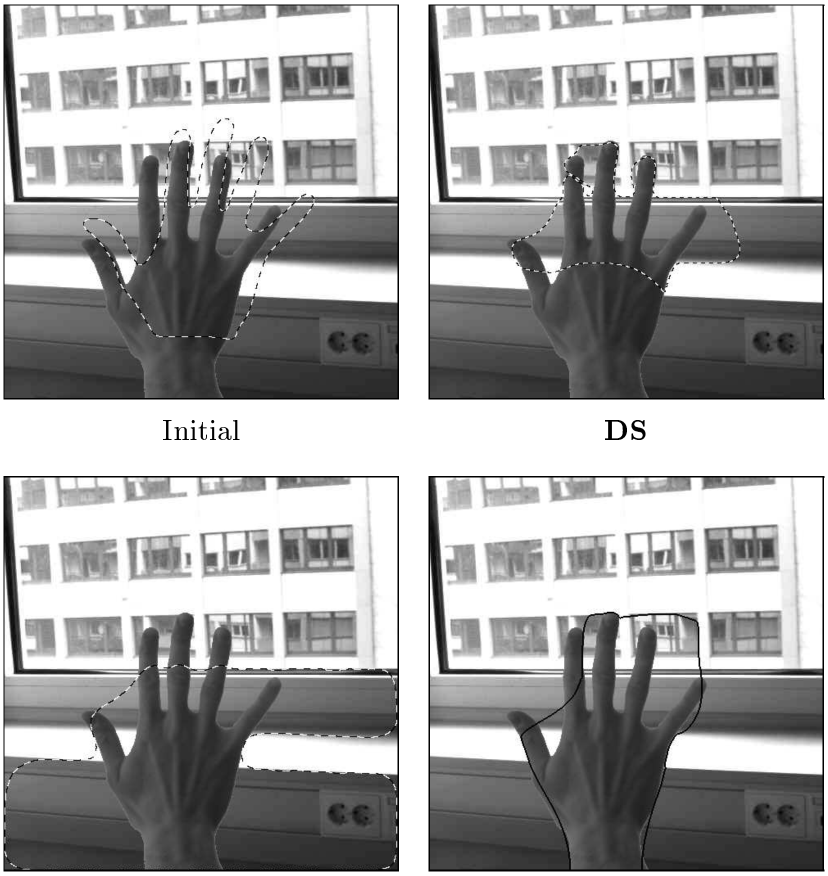
5.4. Coping with Clutter

The scene in Figure 2 contained little clutter. Therefore, segmentation results were fairly close to the desired one, 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 **DS** converges more locally than its cartoon limit, the **SDS**. This will be discussed in more detail in Section 5.5.

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 **DS**. 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.



Initial **DS** **SDS** Geodesic Active Contour

Figure 4. Segmentation with no prior in strongly cluttered background. Initial contour (top left), segmentation results obtained by the **DS**, the **SDS** and a level set scheme of geodesic active contours. Due to the similar grey value properties of object and background, none of these purely data-driven segmentation approaches produces the desired segmentation.

5.5. Comparing the Modified Mumford-Shah Model and its Cartoon Limit

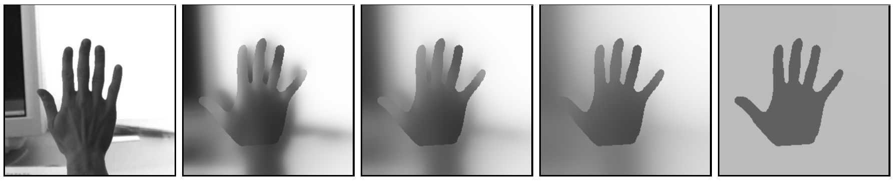
The full Mumford-Shah functional and its cartoon limit differ in their contour evolution equation in that the former collects grey value information from the area surrounding the contour by means of a diffusion process, whereas the latter does this by separately averaging over the areas which are separated by the present contour. The images in Figure

DS **SDS** Energy plots

Figure 5. **Left and middle:** Segmentation with prior in strongly cluttered environment for the modified Mumford-Shah (**DS**) and its cartoon limit (**SDS**). **Right:** Corresponding energy plots (see text). Comparison with the respective images in Figure 4 shows, that the introduction of the shape prior drives the segmenting contour towards the desired segmentation for the **DS**. However, since the **SDS** is more strongly affected by global grey value changes, it does not converge to the desired segmentation for the given initialization (Figure 4, top left).

6 show that the solution for the diffused image converges to that of the cartoon limit for $\lambda \to \infty$. 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 **DS** and its cartoon limit, the **SDS**, for a fixed contour which we simply translated in x- and y-direction. This corresponds to a suppression of shape deformation. We used the same input image as in Figure 4. The contour was optimally placed upon the hand boundaries and then shifted in 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 **SDS** (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



Original $\lambda = 10$ $\lambda = 20$ $\lambda = 40$ $\lambda = 1000$

Figure 6. From the diffusion model to the cartoon limit. Original image and diffused versions for a fixed contour. With growing values of the smoothing parameter λ, the amount of information about the local context is reduced. The contour is modeled by edgels "between pixels", such that all pixels belong to one of the two regions and are therefore affected by the diffusion process.

given – cf. Figure 4, bottom left. Even in the case of added shape prior, the hand contour is pushed to the bottom of the image in the **SDS** – cf. Figure 5, middle.

5.6. Comparison with Geodesic Active Contours when Using 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 data-driven segmentation methods. For the same input images f and the same initial contours C, we minimized the energy functional [4, 20]

$$E_f(C) = \int_0^1 g\left(|\nabla f_\sigma(C(s))|^2\right) |C_s(s)|\, ds, \tag{37}$$

with Gaussian-smoothed input image f_σ, and metric [32]

$$g(s^2) = \begin{cases} 1, & \text{if } s^2 = 0 \\ 1 - \exp\left(-\frac{3.315}{(s/\lambda)^8}\right), & \text{if } s^2 > 0 \end{cases}. \tag{38}$$

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 [32].

The comparison in Figure 2 shows that the segmentation obtained by the Mumford-Shah based models and the one 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. Thus, the segmentation results in Figure 4 indicate that the hypothesis of piecewise homogeneous grey value is more strongly violated than the hypothesis that the object is defined by pronounced edges.

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 permits a straightforward implementation of shape statistics and similarity invariance and may be of advantage for further post-processing of the contour information.

5.7. Invariance to Similarity Transformations

By construction the shape energy (28) is invariant with respect to translation, rotation and scaling. Figure 7 shows a minimization by gradient descent from the initial contour (top left) to the final segmentation, with a shape prior constructed from a set of 10 binarized hand images (bottom right). During its evolution, the contour is effectively restricted to the subspace of familiar contours, but translation, rotation and scaling are permitted.

Due to the closed-form solution for eliminating translation, scale and rotation from the shape energy, no additional parameters enter the minimization. Other authors (cf. [6]) propose to explicitly model a translation, an angle and a scale and minimize with respect to these quantities (by gradient descent or other). In our opinion this has several drawbacks:

- The introduction of four *explicit* pose parameters makes numerical minimization more complicated – corresponding parameters to balance the gradient descent must be chosen. In practice, this is

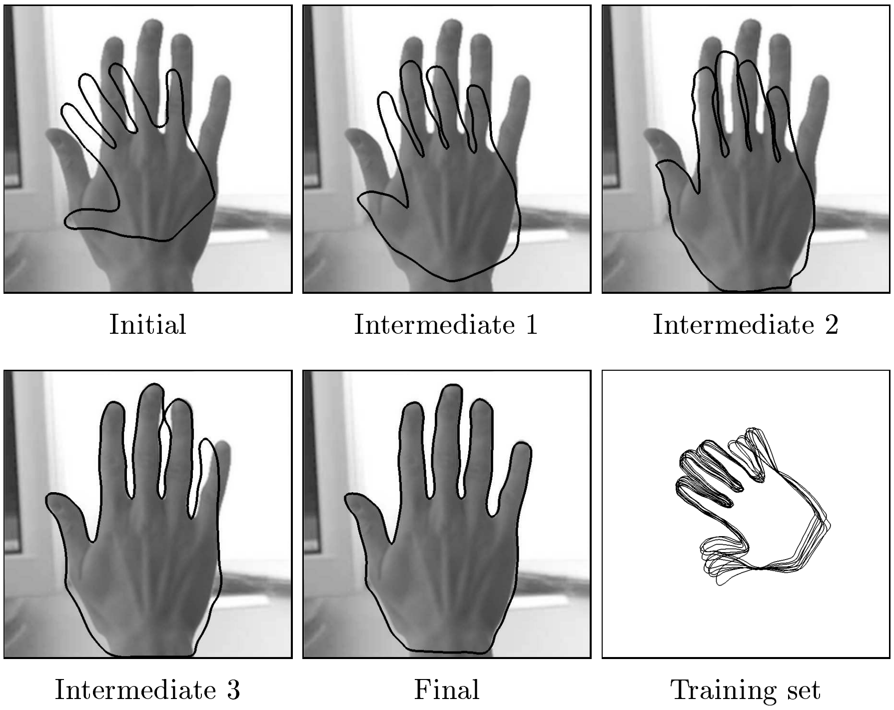


Figure 7. Invariance with respect to similarity transformation. Minimization by gradient descent from the initial contour to the final segmentation. Due to the closed-form solution (28), no additional parameters enter the minimization to account for scale, rotation and translation. Due to the intrinsic alignment of the evolving contour, the relative position, scale and rotation of the training set (bottom right) are of no importance to the knowledge-driven segmentation process.

not only tedious, but it may cause numerical instabilities if the corresponding step sizes are chosen too large.

- The joint minimization of pose and shape parameters mixes the degrees of freedom corresponding to scale and rotation with those corresponding to shape deformation.

- Potential local minima may be introduced by the additional pose parameters. In a given application, this may prevent the convergence of the contour towards the desired segmentation.

On several segmentation tasks we could confirm these effects by comparing the two approaches. For example, for the problem presented in Figure 7, we did not manage to balance the minimization in a way that it would converge to the desired segmentation when using the latter approach of minimizing explicit pose parameters.

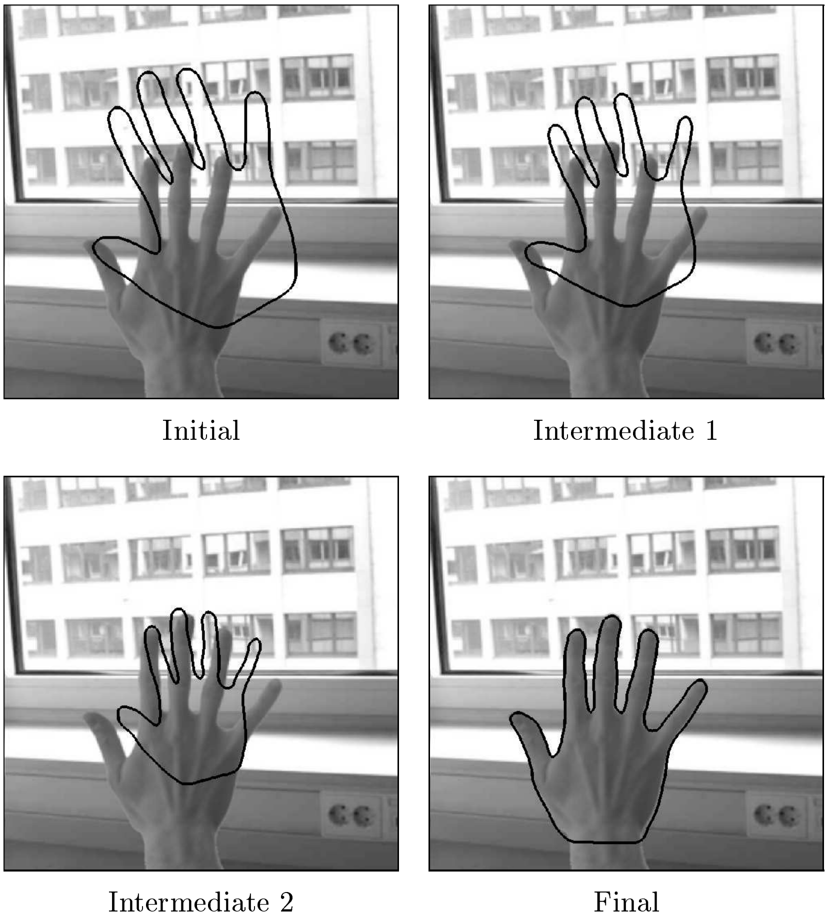
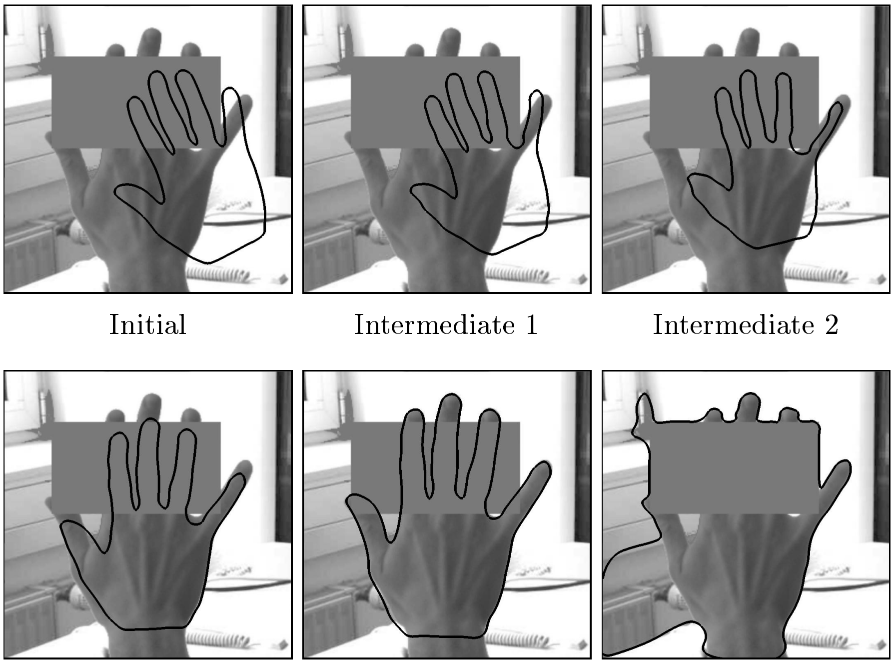


Figure 8. Segmentation process for the **SDS** with a similarity invariant shape prior in a strongly cluttered background. Note that the contour initially shrinks in size, before expanding again towards the desired segmentation. Since the shape energy is invariant with respect to scaling, rotation and translation, the contour will in particular undergo scale changes in order to maximize the grey value homogeneity criterion. This property induces an increased flexibility of the evolving contour.

In many applications, we found the similarity invariance to facilitate the convergence of the evolving contour towards the desired segmentation. Figure 8 shows the result of a segmentation process with a similarity invariant shape prior on the same image as in Figure 4, however with a slightly different initialization. Interestingly, it appears to be energetically favorable for the evolving contour to initially shrink in size (in order to fulfill the grey value homogeneity requirement) before expanding again towards the correct segmentation.



Initial | Intermediate 1 | Intermediate 2

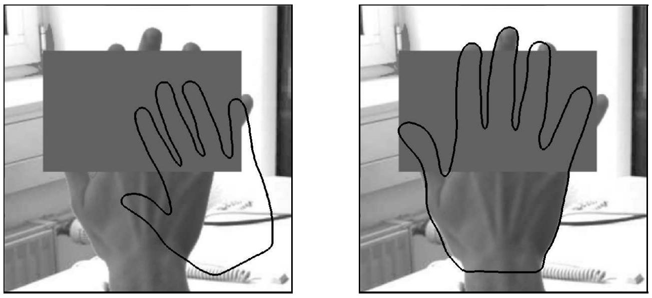
Intermediate 3 | Final with prior | Final without prior

Figure 9. Dealing with occlusion. Contour evolution for the **SDS** with a similarity invariant shape prior. While the segmentation without shape prior (bottom right) fails to capture the object of interest (for the same initialization), the statistical shape prior permits a reconstruction of the occluded parts of the silhouette (bottom center).

5.8. Occlusion Revisited

The main idea of introducing the shape prior is that it is to cope for missing or misleading information. In the case of occlusion, for example, we expect the statistical shape prior to induce a reconstruction of the shape silhouette in parts where the object is not visible. This is demonstrated in Figure 9. A partially occluded image of a hand was segmented by the **SDS** with a similarity invariant shape prior. The silhouette of the object is reconstructed according to the shape prior in areas where it is occluded.

Naturally, the quality of the final segmentation slowly degrades as the size of the occlusion is increased. This is demonstrated in Figure 10, which shows segmentation results obtained for the **SDS** with a similarity invariant statistical shape prior. A comparison with Figure 9 shows that due to the increased occlusion, there is not sufficient information left to correctly guide the evolving contour. Such experiments demon-



Initial Segmentation with prior

Figure 10. With increasing occlusion, the information contained in the image may not be sufficient to correctly guide the segmentation process. Therefore the final segmentation eventually degrades: In this example, the last two fingers are not correctly segmented.

strate that with increasing occlusion, the final segmentation slowly degrades.

5.9. Dealing with Noise

A different case of missing information is given when the image containing the object of interest is corrupted by noise. Depending on the amount of noise, there may be very little information to drive the evolving contour towards the desired segmentation. Again, the statistical shape prior can improve segmentation, because it effectively reduces the dimension of the search space in such a way that segmentations which do not correspond to familiar shapes are ruled out *a priori*.

Figure 11, top left, shows the same input image as in Figure 3. However, this time, 75% of the pixels were replaced by an arbitrary grey value sampled from a uniform distribution over the interval $[0, 255]$. This means that only one out of four pixels contains information about the input image. Figure 11 shows four steps in the contour evolution for the **SDS** with a similarity invariant shape prior. For the given initialization, the segmentation process *with* prior is able to converge to the desired segmentation. In contrast, for the same initialization, the segmentation process *without* the shape prior fails to segment the object of interest, as shown in Figure 11, bottom right.

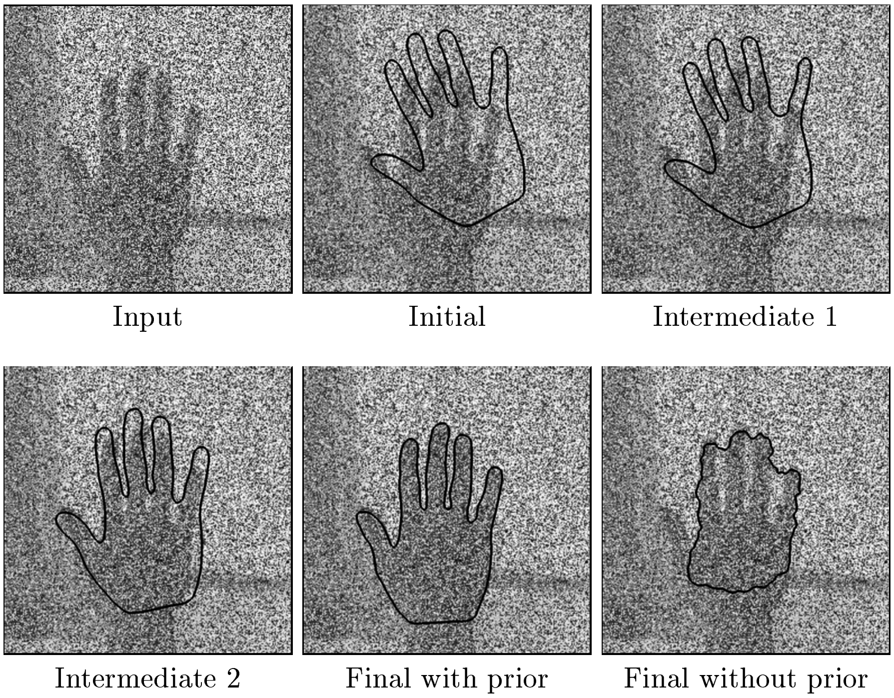


Figure 11. Segmentation of an image corrupted by noise. The input image is the one shown in Figure 3, with 75% of the pixels replaced by grey values sampled from a uniform distribution on the interval $[0, 255]$. Four frames from the gradient descent minimization indicate the contour evolution for the cartoon model with a similarity invariant shape prior. The frame on the bottom right shows the final segmentation obtained for the same initialization without a shape prior. By effectively suppressing unfamiliar shape deformations, the statistical prior facilitates convergence to the desired segmentation.

6. 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 employ the L_2-norm as a measure of the contour length. This enforces an equidistant spacing of control points which turns out to be vital for a stable evolution of the explicit contour. The resulting *diffusion snake* is a hybrid model combining the external energy of the Mumford-Shah functional with the internal energy of the snake.

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. Minimization of a single functional by gradient descent results in a segmentation process which takes account of both the low-level image information and the higher-level shape dissimilarity.

We introduced invariance with respect to similarity transformations of the contour into the variational approach. To this end, we proposed a closed-form, parameter free integration on the basis of the spline control point representation. Essentially, the statistical energy is applied to the contour upon alignment with respect to the training shapes. The resulting energy can be minimized by gradient descent and no additional parameters need to be introduced to account for translation, scale and rotation.

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.

We included sections on the numerical implementation of the diffusion process underlying the curve evolution, both a simple explicit scheme and a more sophisticated multigrid implementation.

Current work focuses on models of nonlinear shape probability density beyond the Gaussian approximation [8], and on extensions to different low-level segmentation cues such as motion [9].

Appendix

A. Implementation of the Diffusion Process

In the following, we will present two different schemes to approximate the diffusion process which is underlying the evolution of the diffusion snake (see Section 3.2): A simple explicit approximation to the diffusion equation

$$\frac{\partial u}{\partial t} \,=\, \nabla \cdot (w_c \, \nabla u) \,+\, \frac{1}{\lambda^2}\,(f-u), \tag{39}$$

and a more sophisticated multigrid scheme solving the corresponding steady state equation

$$\frac{1}{\lambda^2}(u-f) \;-\; \nabla \cdot (w_c \, \nabla u) = 0. \tag{40}$$

Let us start with the explicit scheme.

A.1. A Simple Numerical Scheme

We approximate equation (39) by finite differences. Let $\tau > 0$ denote the step size in t direction and let u_i^k be an approximation to $u(x,t)$ in some pixel i at $t = k\tau$. In a similar way, w_i^k and f_i serve as approximations to $w_c(x,t)$ and $f(x)$, respectively. Moreover, let $\mathcal{N}(i)$ denote the 4-neighborhood of pixel i. If we assume square pixels of size 1, a consistent discretization of (39) is given by

$$\frac{u_i^{k+1} - u_i^k}{\tau} \;=\; \sum_{j \in \mathcal{N}(i)} \sqrt{w_j^k w_i^k}\,(u_j^k - u_i^k) \;+\; \frac{1}{\lambda^2}\,(f_i - u_i^{k+1}) \qquad \forall\, i. \tag{41}$$

The proposed discretization of the indicator function w_c prevents diffusion across the curve C.

Assuming that u_i^k and its neighbors $\{u_j^k \,|\, j \in \mathcal{N}(i)\}$ are already known from the k-th iteration step, we can solve this equation explicitly for the unknown u_i^{k+1}:

$$u_i^{k+1} \;=\; \frac{\left(1 - \tau \sum\limits_{j \in \mathcal{N}(i)} \sqrt{w_j^k w_i^k}\right) u_i^k \;+\; \tau \sum\limits_{j \in \mathcal{N}(i)} \sqrt{w_j^k w_i^k}\; u_j^k \;+\; \frac{\tau}{\lambda^2}\, f_i}{1 + \frac{\tau}{\lambda^2}}. \tag{42}$$

This constitutes our simple iteration scheme for all pixels i and all iteration levels k.

Let us now investigate its stability. Equation (42) computes u_i^{k+1} as a weighted average of u_i^k, its four neighbors $\{u_j^k \,|\, j \in \mathcal{N}(i)\}$, and f_i. Note that the weights sum up to 1. Stability of this process can be guaranteed if all weights are nonnegative. Negative weights, however, can only appear in the first term, if τ is chosen too large. Since

$$0 \leq \sqrt{w_j^k w_i^k} \leq 1, \tag{43}$$

we end up with the stability restriction

$$\tau \leq \frac{1}{4}. \tag{44}$$

In this case we have a convex combination which guarantees that

$$\min_j (f_j, u_j^k) \;\leq\; u_i^{k+1} \;\leq\; \max_j (f_j, u_j^k) \qquad \forall\, i, k. \tag{45}$$

By initializing $u_j^0 := f_j$ and iterating over k, this simplifies to the discrete maximum-minimum principle

$$\min_j f_j \;\leq\; u_i^k \;\leq\; \max_j f_j \qquad \forall\, i, k. \tag{46}$$

This guarantees that the filtered image remains within the bounds of the original image.

A.2. A Multigrid Scheme for Diffusion Snakes

In the following, we will detail a multigrid scheme for solving the steady state equation (40). We discretize equation (40) by finite differences to obtain a linear system with natural (Neumann) boundary conditions:

$$\boldsymbol{A}u = f, \quad \text{and} \;\; \partial_n u = 0 \text{ on } \partial\Omega. \tag{47}$$

The contour is represented by edgels "between pixels" (micro-edges), such that all image pixels are affected by the diffusion process.

Solving equation (47) with standard solvers like Gauss-Seidel or Jacobi takes a long time, as low frequencies in the error vanish slowly. Therefore, we propose a multigrid implementation, which consists in recursively transferring the problem from a grid with size h to a coarser grid of size $2h$, and solving there to obtain good a initialization for the solution on the fine grid.

A.2.1. *Hierarchical Representation at Multiple Scales*

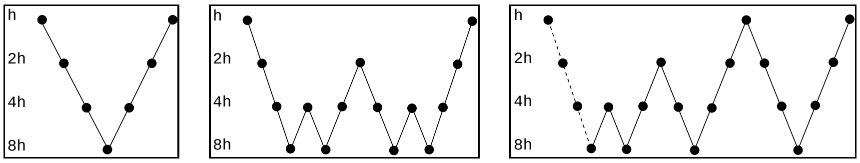
Standard implementation of some numerical multigrid scheme, like the one in [28], may easily lead to a poor implementation of the steady state diffusion equation (40) due to the strongly inhomogeneous term w_c. The hierarchical representation of this term at multiple scales is even more difficult. For the diffusion snake to work, smoothing across the curve C must be prevented at all scales.

Let v be an approximation to the solution u. Denote the error by $e = u - v$, and the residual by $r = f - \boldsymbol{A}v$. With these notations we obtain for every grid h the following equivalence:

$$\boldsymbol{A}^h u^h = f^h \quad \Longleftrightarrow \quad \boldsymbol{A}^h e^h = r^h. \tag{48}$$

In order to solve the latter problem on the fine grid h, we transfer the residual r^h and the matrix $\boldsymbol{A}^h$ to the coarser grid $2h$, solve

$$\boldsymbol{A}^{2h} e^{2h} = r^{2h} \tag{49}$$



V-cycle W-cycle Full multigrid

Figure 12. Schematic diagrams of multigrid cycles. An elegant recursive definition of different multigrid cycles can be found in [3].

$$
\begin{bmatrix} & 0 & \\ 0 & 1+\frac{2\lambda^2}{h^2} & -\frac{\lambda^2}{h^2} \\ & -\frac{\lambda^2}{h^2} & \end{bmatrix}
\begin{bmatrix} & 0 & \\ -\frac{\lambda^2}{h^2} & 1+\frac{3\lambda^2}{h^2} & -\frac{\lambda^2}{h^2} \\ & -\frac{\lambda^2}{h^2} & \end{bmatrix}
\begin{bmatrix} & 0 & \\ -\frac{\lambda^2}{h^2} & 1+\frac{2\lambda^2}{h^2} & 0 \\ & -\frac{\lambda^2}{h^2} & \end{bmatrix}
$$

$$
\begin{bmatrix} & -\frac{\lambda^2}{h^2} & \\ 0 & 1+\frac{3\lambda^2}{h^2} & -\frac{\lambda^2}{h^2} \\ & -\frac{\lambda^2}{h^2} & \end{bmatrix}
\begin{bmatrix} & -\frac{\lambda^2}{h^2} & \\ -\frac{\lambda^2}{h^2} & 1+\frac{4\lambda^2}{h^2} & -\frac{\lambda^2}{h^2} \\ & -\frac{\lambda^2}{h^2} & \end{bmatrix}
\begin{bmatrix} & -\frac{\lambda^2}{h^2} & \\ -\frac{\lambda^2}{h^2} & 1+\frac{3\lambda^2}{h^2} & 0 \\ & -\frac{\lambda^2}{h^2} & \end{bmatrix}
$$

$$
\begin{bmatrix} & -\frac{\lambda^2}{h^2} & \\ 0 & 1+\frac{2\lambda^2}{h^2} & -\frac{\lambda^2}{h^2} \\ & 0 & \end{bmatrix}
\begin{bmatrix} & -\frac{\lambda^2}{h^2} & \\ -\frac{\lambda^2}{h^2} & 1+\frac{3\lambda^2}{h^2} & -\frac{\lambda^2}{h^2} \\ & 0 & \end{bmatrix}
\begin{bmatrix} & -\frac{\lambda^2}{h^2} & \\ -\frac{\lambda^2}{h^2} & 1+\frac{2\lambda^2}{h^2} & 0 \\ & 0 & \end{bmatrix}
$$

Figure 13. Stencils for diffusion snakes. Each stencil defines the action of operator $\boldsymbol{A}$ at a given pixel location, where the zeros denote the effect of a diffusion boundary. Displayed are (from left to right, top to bottom) the stencils $[\boldsymbol{A}]_{(0,0)}$, $[\boldsymbol{A}]_{(i,0)}$, $[\boldsymbol{A}]_{(M,0)}$, $[\boldsymbol{A}]_{(0,j)}$, $[\boldsymbol{A}]_{(i,j)}$, $[\boldsymbol{A}]_{(M,j)}$, $[\boldsymbol{A}]_{(0,N)}$, $[\boldsymbol{A}]_{(i,N)}$, $[\boldsymbol{A}]_{(M,N)}$.

for e^{2h}, interpolate back to the fine grid and add e^h to the fine grid solution v^h. This idea is recursively extended to more than two grids, which leads to different multigrid cycles, some of which are depicted in Figure 12. We found that w-cycles showed the best performance in our experiments.

A.2.2. *Interpolation, Restriction and Coarse Grid Representation*

Starting with the matrix $\boldsymbol{A}^h$, we need to construct appropriate prolongation operators $\boldsymbol{P}$ and restriction operators $\boldsymbol{R}$, which define the transition from the coarse to the fine grids and vice versa. For this purpose we introduce the stencil notation, where the stencils shown in Figure 13 represent the action of the operator $\boldsymbol{A}$ on a pixel and its 3×3-neighborhood. This notation allows to intuitively understand the effect of the operator $\boldsymbol{A}$ at a given location. The effect of the boundary conditions imposed by the contour and the image edges is given by the zeros in the stencils in Figure 13.

The implementation of the contour as a diffusion border prohibits any restriction or prolongation across this border. We therefore use matrix-dependent prolongation and restriction operators, as described in [34]. Similar approaches were proposed in [1, 12, 36].

In the following, we will define the prolongation operator, which performs the transition from the coarse grid $2h$ to the fine grid h. According to [34], two constraints have to be fulfilled for the prolongation operator in the one-dimensional case:

$$u_{2i}^h = [\boldsymbol{P}]_{i,0}^{2h} \cdot u_i^{2h} = u_i^{2h}, \tag{50}$$

$$(\boldsymbol{A}\boldsymbol{P}u^{2h})_{2i+1} = 0, \tag{51}$$

where the first lower index at the stencil denotes the pixel number and the second lower index denotes the position within the stencil, which can be -1, 0 or 1 for left, middle and right in the 1D case.

The first constraint ensures, that all coarse grid points are transfered directly to the finer grid, the second one ensures, that the prolongation operator is adapted to the matrix $\boldsymbol{A}$. The odd pixels of the finer grid can then be obtained from equation (51):

$$u_{2i+1}^h = -\frac{[\boldsymbol{A}]_{(2i+1),-1}^h \cdot u_i^{2h} + [\boldsymbol{A}]_{(2i+1),1}^h \cdot u_{i+1}^{2h}}{[\boldsymbol{A}]_{(2i+1),0}^h}. \tag{52}$$

A similar solution for the prolongation $\boldsymbol{P}$ is obtained in the two-dimensional case [34]. The stencils for the restriction correspond to the prolongation stencils, normalized so that they sum up to 1. Further details can be found in [30].

With these definitions of prolongation $\boldsymbol{P}$ and restriction $\boldsymbol{R}$ from the matrix $\boldsymbol{A}^h$, we construct the coarse grid matrix $\boldsymbol{A}^{2h}$ by using Galerkin coarsening:

$$\boldsymbol{A}^{2h} = \boldsymbol{R}\boldsymbol{A}^h\boldsymbol{P}. \tag{53}$$

To avoid a full matrix multiplication, we exploit the stencil notation as done in the efficient algorithm CALRAP [34]. Given $\boldsymbol{A}^{2h}$, we can then construct prolongation and restriction for the next coarser level and so on.

A.2.3. *Results of the Multigrid Implementation*

Figure 14 shows that using multigrid methods for solving the linear system (47) leads to a performance gain of several orders of magnitude compared to the use of standard algorithms. Using a w-cycle with three descending v-cycles and one step for presmoothing and postsmoothing on each level, one reaches the level of precision of a standard computer in only a few multigrid steps.

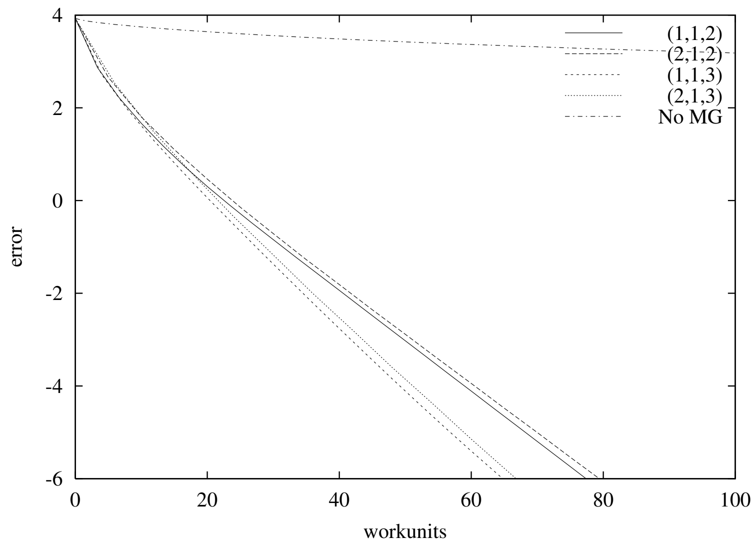


Figure 14. Comparison of different multigrid implementations and the symmetric Gauss-Seidel as a standard solver. The error is defined in logarithmic scale as $\log_{10} ||e||_2$. The numbers of presmoothing, postsmoothing and v-cycles on each level are given in brackets.

Analogous to the performance of standard solvers for the common model problems, we found the computation time of the multigrid implementation to be fairly independent of the size of the smoothing parameter λ. This proves the robustness of our hierarchical scheme with respect to the strongly inhomogeneous diffusivity w_c. Moreover, the additional storage requirements are negligible. Further details can be found in [30].

Acknowledgments

We thank J. Keuchel and C. Kervrann for helpful discussions.

References

1. R. E. Alcouffe, A. Brandt, J. E. Dendy, Jr., and J. W. Painter. The multigrid method for the diffusion equation with strongly discontinuous coefficients. *SIAM J. Sci. Stat. Comp.*, 2(4):430–454, 1981.
2. A. Blake and M. Isard. *Active Contours*. Springer, London, 1998.
3. W. L. Briggs, V. E. Henson, and S. F. McCormick. *A Multigrid Tutorial*. Siam, Philadelphia, second edition, 2000.

4. V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. In *Proc. IEEE Internat. Conf. on Comp. Vis.*, pages 694–699, Boston, USA, 1995.
5. T. Chan and L. Vese. A level set algorithm for minimizing the Mumford–Shah functional in image processing. In *IEEE Workshop on Variational and Level Set Methods*, pages 161–168, Vancouver, CA, 2001.
6. Y. Chen, S. Thiruvenkadam, H. Tagare, F. Huang, D. Wilson, and E. Geiser. On the incorporation of shape priors into geometric active contours. In *IEEE Workshop on Variational and Level Set Methods*, pages 145–152, Vancouver, CA, 2001.
7. 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.
8. D. Cremers, T. Kohlberger, and C. Schnörr. Nonlinear shape statistics in Mumford–Shah based segmentation. In A. Heyden et al., editors, *Proc. of the Europ. Conf. on Comp. Vis.*, volume 2351 of *LNCS*, pages 93–108, Copenhagen, May, 28–31 2002. Springer, Berlin.
9. D. Cremers and C. Schnörr. Motion competition: Variational integration of motion segmentation and shape regularization. In L. van Gool, editor, *Pattern Recognition*, volume 2449 of *LNCS*, pages 472–480, Zürich, Sept. 2002. Springer.
10. D. Cremers, C. Schnörr, and J. Weickert. Diffusion snakes: Combining statistical shape knowledge and image information in a variational framework. In *IEEE First Workshop on Variational and Level Set Methods*, pages 137–144, Vancouver, 2001.
11. 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 *LNCS*, pages 164–174, Kiel, Germany, Sept. 10–11, 2000. Springer.
12. J. E. Dendy. Black box multigrid. *J. Comp. Phys.*, 48:366–386, 1982.
13. I. L. Dryden and K. V. Mardia. *Statistical Shape Analysis*. Wiley, Chichester, 1998.
14. G. Farin. *Curves and Surfaces for Computer–Aided Geometric Design*. Academic Press, San Diego, 1997.
15. C. Goodall. Procrustes methods in the statistical analysis of shape. *J. Roy. Statist. Soc., Ser. B.*, 53(2):285–339, 1991.
16. U. Grenander, Y. Chow, and D.M. Keenan. *Hands: A Pattern Theoretic Study of Biological Shapes*. Springer, New York, 1991.
17. M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *Int. J. of Comp. Vis.*, 1(4):321–331, 1988.
18. 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.
19. C. Kervrann and F. Heitz. Statistical deformable model-based segmentation of image motion. *IEEE Trans. on Image Processing*, 8:583–588, 1999.
20. S. Kichenassamy, A. Kumar, P. J. Olver, A. Tannenbaum, and A. J. Yezzi. Gradient flows and geometric active contour models. In *Proc. IEEE Internat. Conf. on Comp. Vis.*, pages 810–815, Boston, USA, 1995.
21. 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, SC, June 13–15, 2000.
22. C. Mantegazza. *Su Alcune Definizioni Deboli di Curvatura per Insiemi Non Orientati*. PhD thesis, Dept. of Mathematics, SNS Pisa, Italy, 1993.

23. B. Moghaddam and A. Pentland. Probabilistic visual learning for object detection. In *Proc. IEEE Internat. Conf. on Comp. Vis.*, pages 786–793, 1995.
24. 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.
25. D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. Pure Appl. Math.*, 42:577–685, 1989.
26. N. Paragios and R. Deriche. Coupled geodesic active regions for image segmentation: a level set approach. In D. Vernon, editor, *Proc. of the Europ. Conf. on Comp. Vis.*, volume 1843 of *LNCS*, pages 224–240. Springer, 2000.
27. S. Roweis. EM algorithms for PCA and SPCA. In M. Jordan, M. Kearns, and S. Solla, editors, *Advances in Neural Information Processing Systems 10*, pages 626–632, Cambridge, MA, 1998. MIT Press.
28. D. Terzopoulos. Multilevel computational processes for visual surface reconstruction. *Comp. Vis., Graph., and Imag. Proc.*, 24:52–96, 1983.
29. M. E. Tipping and C. M. Bishop. Probabilistic principal component analysis. Technical Report Woe-19, Neural Computing Research Group, Aston University, UK, 1997.
30. F. Tischhäuser. Development of a multigrid algorithm for diffusion snakes. Diploma thesis (in German), Department of Mathematics and Computer Science, University of Mannheim, Mannheim, Germany, 2001.
31. 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.
32. J. Weickert. Applications of nonlinear diffusion filtering in image processing and computer vision. *Acta Mathematica Universitatis Comenianae*, LXX(1):33–50, 2001.
33. M. Werman and D. Weinshall. Similarity and affine invariant distances between 2d point sets. *IEEE Trans. on Patt. Anal. and Mach. Intell.*, 17(8):810–814, 1995.
34. P. Wesseling. *An Introduction to Multigrid Methods.* John Wiley & Sons, Chichester, 1992.
35. A. Yezzi, S. Soatto, A. Tsai, and A. Willsky. The Mumford–Shah functional: From segmentation to stereo. *Mathematics and Multimedia*, 2002. To appear.
36. P. M. de Zeeuw. Matrix–dependent prolongations and restrictions in a blackbox multigrid solver. *J. Comp. Appl. Math.*, 33:1–27, 1990.
37. S. C. Zhu and D. Mumford. Prior learning and Gibbs reaction–diffusion. *IEEE Trans. on Patt. Anal. and Mach. Intell.*, 19(11):1236–1250, 1997.
38. S. C. Zhu and A. Yuille. Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation. *IEEE Trans. on Patt. Anal. and Mach. Intell.*, 18(9):884–900, 1996.