Review of Statistical Approaches to Level Set Segmentation: Integrating Color, Texture, Motion and Shape

Daniel Cremers and Mikael Rousson
Siemens Corporate Research
755 College Road East, Princeton, NJ 08540
{*daniel.cremers,mikael.rousson*}*@siemens.com*

Rachid Deriche
INRIA, Sophia-Antipolis, France
{*rachid.deriche*}*@sophia.inria.fr*

Abstract. Since their introduction as a means of front propagation and their first application to edge-based segmentation in the early 90's, level set methods have become increasingly popular as a general framework for image segmentation. In this paper, we present a survey of a specific class of region-based level set segmentation methods and clarify how they can all be derived from a common statistical framework.

Region-based segmentation schemes aim at partitioning the image domain by progressively fitting statistical models to the intensity, color, texture or motion in each of a set of regions. In contrast to edge-based schemes such as the classical Snakes, region-based methods tend to be less sensitive to noise. For typical images, the respective cost functionals tend to have less local minima which makes them particularly well-suited for local optimization methods such as the level set method.

We detail a general statistical formulation for level set segmentation. Subsequently, we clarify how the integration of various low level criteria leads to a set of cost functionals and point out relations between the different segmentation schemes. In experimental results, we demonstrate how the level set function is driven to partition the image plane into domains of coherent color, texture, dynamic texture or motion. Moreover, the Bayesian formulation allows to introduce prior shape knowledge into the level set method. We briefly review a number of advances in this domain.

Keywords: Image segmentation, level set methods, Bayesian inference, color, texture, motion

1. Introduction

The goal of image segmentation is to partition the image plane into meaningful areas, where *meaningful* typically refers to a separation of areas corresponding to different objects in the observed scene from the area corresponding to the background.

A large variety of segmentation algorithms have been proposed over the last few decades. While earlier approaches were often based on a set of rather heuristic processing steps (cf. [69]), optimization methods

have become established as more principled and transparent methods: Segmentations of a given image are obtained by minimizing appropriate cost functionals. Among optimization methods, one can distinguish between *spatially discrete* and *spatially continuous* representations.

In spatially discrete approaches, the pixels of the image are usually considered as the nodes of a graph, and the aim of segmentation is to find cuts of this graph which have a minimal cost. Optimization algorithms for these problems include greedy approaches such as the *Iterated Conditional Modes* (ICM) [2] and continuation methods such as *Simulated Annealing* [35] or *Graduated Non-convexity* [5]. Specific classes of graph cut approaches gained in popularity with the re-discovery of efficient global optimization methods, which are based on concepts of dynamic programming [6], on spectral methods [82, 56] or on semidefinite programming techniques [45].

In spatially continuous approaches, the segmentation of the image plane $\Omega \subset \mathbb{R}^2$ is considered as a problem of infinite-dimensional optimization. Using variational methods, one computes segmentations of a given image $I : \Omega \to \mathbb{R}$ by evolving contours in the direction of the negative energy gradient using appropriate partial differential equations (pdes). Such pde-based segmentation methods became popular with the seminal paper on *Snakes* by Kass et al. [44]. In this paper, the contour is implemented by an explicit (parametric) curve $C : [0, 1] \to \Omega$ which is evolved by locally minimizing the functional

$$E(C) = -\int |\nabla I(C)|^2 \, ds \; + \; \nu_1 \int |C_s|^2 \, ds \; + \; \nu_2 \int |C_{ss}|^2 \, ds, \quad (1)$$

where C_s and C_{ss} denote the first and second derivative with respect to the curve parameter s. The first term in (1) is the external energy which accounts for the image information, in the sense that the minimizing contour will favor locations of large image gradient. The last two terms – weighted by nonnegative parameters ν_1 and ν_2 – can be interpreted as an *internal energy* of the contour, measuring the length of the contour and its stiffness or rigidity.[1]

The Snakes approach had an enormous impact in the segmentation community (with over 3000 citations to date). Yet, it suffers from several drawbacks:

1. The implementation of contour evolutions based on an explicit parameterization requires a delicate regriding (or reparameterization) process to avoid self-intersection and overlap of control or marker points.

[1] From a survey of a number of related publications and from our personal experience, it appears that the rigidity term is not particularly important, such that one commonly sets $\nu_2 = 0$.

2. The explicit representation by default does not allow the evolving contour to undergo topological changes such that the segmentation of several objects or multiply-connected objects is not straight-forward.[2]

3. The segmentations obtained by a local optimization method are bound to depend on the initialization. The Snake algorithm is known to be quite sensitive to the initialization. For many realistic images, the segmentation algorithm tends to get stuck in undesired local minima – in particular in the presence of noise.

4. The Snakes approach lacks a meaningful probabilistic interpretation. Extensions to other segmentation criteria – such as color, texture or motion – are not straight-forward.

In the present paper, we will review recent developments in the segmentation community which aim at resolving the above problems. We will review the level set method for front propagation as a means to handle topological changes of evolving interfaces and to remove the issues of contour parameterization and control point regriding. Among the level set methods, we will focus on statistical region-based methods, where the contour is not evolved by fitting to local edge information (as in the Snakes) but rather by fitting statistical models to intensity, color, texture or motion within each of the separated regions. The respective cost functionals tend to have less local minima for most realistic images. As a consequence, the segmentation schemes are far less sensitive to noise and to varying initialization.

The outline of the paper is as follows: In Section 2, we will review the general idea of level set based boundary propagation and its first applications to image segmentation. In Section 3, we will then review a probabilistic formulation of region-based segmentation. In particular, we will make very explicit what are the assumptions underlying the derivation of appropriate cost functionals. In the subsequent sections, we then detail how to adapt the probabilistic level set framework to different segmentation criteria: In Section 4, we present probabilistic models which drive the segmentation process to group regions of homogeneous intensity, color or texture. In Section 5, we briefly present extensions of this framework to Diffusion Tensor Images. In Section 6, we discuss a further extension which allows to exploit *spatio-temporal* dynamics to drive a segmentation process, given an entire sequence of images. In particular, this approach allows to separate textures which

[2] It should be pointed out that based on various heuristics, one can successfully incorporate regriding mechanisms and topological changes into explicit representations – cf. [62, 48, 28, 25].

have identical spatial characteristics but differ in their temporal dynamics. In Section 7 we detail how to integrate motion information as a criterion for segmentation, leading to a partitioning of the image plane into areas of piecewise parametric motion. Finally, in Section 8, we briefly discuss numerous efforts to introduce statistical shape knowledge in level set based image segmentation in order to cope with missing or misleading low-level information.

2. Level Set Methods for Image Segmentation

In the variational framework, a segmentation of the image plane Ω is computed by locally minimizing an appropriate energy functional, such as the functional (1). The key idea is to evolve the boundary $\mathcal{C}$ from some initialization in direction of the negative energy gradient, which is done by implementing the gradient descent equation:

$$\frac{\partial \mathcal{C}}{\partial t} = -\frac{\partial E(\mathcal{C})}{\partial \mathcal{C}} = F \cdot n, \tag{2}$$

modeling an evolution along the normal n with a speed function F.[3]

In general, one can distinguish between explicit (parametric) and implicit representations of contours. In explicit representations – such as splines or polygons – a contour is defined as a mapping from an interval to the image domain: $\mathcal{C} : [0, 1] \rightarrow \Omega$. The propagation of an explicit contour is typically implemented by a set of ordinary differential equations acting on the control or marker points. In order to guarantee stability of the contour evolution (i.e. preserve well-defined normal vectors), one needs to introduce certain regriding mechanisms to avoid overlap of control points, for example by numerically resampling the marker points every few iterations, by imposing in the variational formulation a rubber-band like attraction between neighboring points [25], or by introducing electrostatic repulsion [91]. Moreover, in order to segment several objects or multiply connected objects, one needs to introduce numerical tests to enable splitting and remerging of contours during the evolution. Successful advances in this direction were proposed among others by [50, 62, 48, 28].

In implicit contour representations, contours are represented as the (zero) level line of some embedding function $\phi : \Omega \rightarrow I\!R$:

$$\mathcal{C} = \{x \in \Omega \mid \phi(x) = 0\}. \tag{3}$$

[3] Most meaningful contour evolutions do not contain any tangential component as the latter does not affect the contour, but only the parameterization.

There are various methods to evolve implicitly represented contours. The most popular among these is the level set method [29, 30, 65], in which a contour is propagated by evolving a time-dependent embedding function $\phi(x,t)$ according to an appropriate partial differential equation. In the following, we will briefly sketch two alternative methods to derive a level set evolution implementing the minimization of the energy $E(\mathcal{C})$.

For a contour which evolves along the normal n with a speed F – see equation (2) – one can derive a corresponding partial differential equation for the embedding function ϕ in the following way. Since $\phi(\mathcal{C}(t),t) = 0$ at all times, the total time derivative of ϕ at locations of the contour must vanish:

$$\frac{d}{dt}\,\phi(\mathcal{C}(t),t) = \nabla\phi\frac{\partial\mathcal{C}}{\partial t} + \frac{\partial\phi}{\partial t} = \nabla\phi\, F\cdot n + \frac{\partial\phi}{\partial t} = 0. \tag{4}$$

Inserting the definition of the normal $n = \frac{\nabla\phi}{|\nabla\phi|}$, we get the evolution equation for ϕ:

$$\frac{\partial\phi}{\partial t} = -|\nabla\phi|\, F. \tag{5}$$

By derivation, this equation only specifies the evolution of ϕ (and the values of the speed function F) at the location of the contour. For a numerical implementation one needs to extend the right-hand side of (5) to the image domain away from the contour.

Alternatively to the above derivation, one can obtain a level set equation from a variational formulation (cf. [98, 12]): Rather than deriving an appropriate partial differential equation for ϕ which implements the contour evolution equation (2), one can embed a variational principle $E(\mathcal{C})$ defined on the space of contours by a variational principle $E(\phi)$ defined on the space of level set functions:

$$E(\mathcal{C}) \quad\longrightarrow\quad E(\phi)$$

Subsequently, one can derive the Euler-Lagrange equation which minimizes $E(\phi)$:

$$\frac{\partial\phi}{\partial t} = -\frac{\partial E(\phi)}{\partial\phi}. \tag{6}$$

In both cases, the embedding is not uniquely defined. Depending on the chosen embedding, one can obtain slightly different evolution equations for $\phi(x,t)$.

The first applications of this level set formalism for the purpose of image segmentation were proposed in [10, 58, 57]. Indepdently, Caselles et al. [11] and Kichenassamy et al. [46] proposed a level set formulation

for the Snake energy (1) given by:

$$\frac{\partial \phi}{\partial t} = |\nabla \phi| \mathrm{div}\left(g(I) \frac{\nabla \phi}{\|\nabla \phi|} \right) = g(I)|\nabla \phi| \mathrm{div}\left(\frac{\nabla \phi}{\|\nabla \phi|} \right) + \nabla g(I) \cdot \nabla \phi, \quad (7)$$

where the gradient $|\nabla I|$ in functional (1) was replaced by a more general edge function $g(I)$. This approach is known as *Geodesic Active Contours*, because the underlying energy can be interpreted as the length of a contour in a Riemannian space with a metric induced by the image intensity. See [11, 46] for details.

Local optimization methods such as the Snakes have been heavily criticized because the computed segmentations depend on the initialization and because algorithms are easily trapped in undesired local minima for many realistic images. In particular in the presence of noise, numerous local minima of the cost functional (1) are created by local maxima of the image gradient. To overcome these local minima and to drive the contour toward the boundaries of objects of interest, researchers have introduced an additional balloon force [16] which leads to either a shrinking or an expansion of contours. Unfortunately this requires prior knowledge about whether the object of interest is inside or outside the initial contour. Moreover, the final segmentation will be biased toward smaller or larger segmentations.

In the following, we will review a probabilistic formulation of the segmentation problem which leads to region-based functionals rather than edge-based functionals such as the Snakes. Moreover, we will provide numerous experiments which demonstrate that such probabilistic region-based segmentation schemes do not suffer from the above drawbacks. While optimization is still done in a local manner, the respective functionals tend to have few local minima and segmentation results tend to be very robust to noise and varying initialization.

3. Statistical Formulation of Region-Based Segmentation

3.1. Image Segmentation as Bayesian Inference

Statistical approaches to image segmentation have a long tradition, they can be traced back to models of magnetism in physics, such as the Ising model [41], pioneering works in the field of image processing include spatially discrete formulations such as those of Geman and Geman [35] and Besag [2], and spatially continuous formulations such as the ones of Mumford and Shah [63] and Zhu and Yuille [100].

The probabilistic formulation of the segmentation problem presented in the following extends the statistical approaches pioneered in [49, 100,

66, 87]. In particular, this extension allows the probabilistic framework to be applied to segmentation criteria such as texture and motion, which will be detailed in subsequent sections. In [49], a segmentation functional is obtained from a Minimum Description Length (MDL) criterion. The link with the Mumford-Shah functional and the equivalence to Bayesian maximum *a posteriori* (MAP) estimation is provided in [100]. Following [66], an optimal partition $\mathcal{P}(\Omega)$ of the image plane Ω (i.e. a partition of the image plane into pairwise disjoint regions) can be computed by maximizing the *a posteriori* probability $p(\mathcal{P}(\Omega) \,|\, I)$ for a given image I.[4] The Bayes rule permits to express this conditional probability as

$$p(\mathcal{P}(\Omega) \,|\, I) \propto p(I \,|\, \mathcal{P}(\Omega))\; p(\mathcal{P}(\Omega)), \tag{8}$$

thereby separating image-based cues (first term) from geometric properties of the partition (second term). The Bayesian framework has become increasingly popular to tackle many ill-posed problems in computer vision. Firstly the conditional probability $p(I \,|\, \mathcal{P}(\Omega))$ of an observation given a model state is often easier to model than the posterior distribution, it typically follows from a generative model of the image formation process. Secondly, the term $p(\mathcal{P}(\Omega))$ in (8) allows to introduce prior knowledge stating which interpretations of the data are *a priori* more or less likely. Wherever available, such *a priori* knowledge may help to cope with missing low-level information.

One can distinguish between generic priors and object specific priors. Object specific priors can be computed from a set of sample segmentations of an object of interest. In Section 8, we will briefly review a number of recent advances regarding the incorporation of statistically learnt priors into the level set framework.

In this section, we will focus on generic (often called *geometric*) priors. The most commonly used regularization constraint is a prior which favors a short length $\mathcal{C}$ of the partition boundary:

$$p(\mathcal{P}(\Omega)) \propto e^{-\nu\,|\mathcal{C}|}, \quad \nu > 0. \tag{9}$$

Higher-order constraints may be of interest for specific applications such as the segmentation of thin elongated structures [71, 64].

To further specify the image term $p(I \,|\, \mathcal{P}(\Omega))$ in (8), we make the following hypotheses. Following [66], we assume the image partition to be composed of N regions without correlation between the labellings. This gives the simplified expression:

$$p(I \,|\, \mathcal{P}(\Omega)) = p(I \,|\, \{\Omega_1, \ldots, \Omega_N\}) = \prod_{i=1}^{N} p(I \,|\, \Omega_i), \tag{10}$$

[4] In the following, I can refer to a single image or to an entire image sequence.

where $p(I \,|\, \Omega_i)$ denotes the probability of observing an image I when Ω_i is a region of interest. Let us assume that regions of interest are characterized by a given *feature* $f(x)$ associated with each image location. This feature may be a scalar quantity (such as the image intensity), a vector quantity (such as color or the spatio-temporal image gradient), or a tensor (such as a structure tensor or a diffusion tensor).

For the features presented in this paper, we make the assumption that the values of f at different locations of the same region can be modeled as independent and identically distributed realizations of the same random process.[5] Let p_i be the probability density function (*pdf*) of this random process in Ω_i. Expression (10) then reads

$$p(I \,|\, \mathcal{P}(\Omega)) = \prod_{i=1}^{N} \prod_{x \in \Omega_i} \Big(p_i(f(x)) \Big)^{dx}, \tag{11}$$

where the bin volume dx is introduced to guarantee the correct continuum limit. Approximation (11) is not valid in general since image features (such as spatial gradients) are computed on a neighborhood structure and may therefore exhibit local spatial correlations. More importantly, one should expect to find spatial correlations of features when modeling textured regions. However, one can capture certain spatial correlations in the above model by computing appropriate features such as the structure tensor.

Maximization of the *a posteriori* probability (8) is equivalent to minimizing its negative logarithm. Integrating the regularity constraint (9) and the region-based image term (11), we end up with following energy:

$$E(\{\Omega_1, \ldots, \Omega_N\}) = -\sum_i \int_{\Omega_i} \log p_i(f(x))\, dx + \nu\, |\mathcal{C}|. \tag{12}$$

In the context of intensity segmentation (i.e. $f = I$), this energy is the basis of several works [49, 100, 80, 66]. The region statistics are typically computed interlaced with the estimation of the boundary $\mathcal{C}$ [100], yet one can also compute appropriate intensity histograms beforehand [66]. In this paper, we will focus on the case that distributions and segmentation are computed jointly. Distributions can be either modeled as parametric or non-parametric ones. Upon insertion of parametric representations for p_i with parameters θ_i, the energy (12) takes on the

[5] In Section 7, we will consider a generalization in which the underlying random processes are assumed to be space-varying. The distributions p_i in (11) then contain an explicit space dependency $p_i(f(x), x)$ which allows to model spatially varying statistical distributions of features.

form

$$E(\{\Omega_i, \theta_i\}_{i=1..N}) = -\sum_i \int_{\Omega_i} \log p(f(x) \,|\, \theta_i)\, dx + \nu\, |\mathcal{C}|. \qquad (13)$$

For particular choices of parametric densities, the optimal parameters can be expressed as functions of the corresponding domains and only the regions remain as unknowns of the new energy

$$\hat{E}(\{\Omega_i\}_i) \equiv \min_{\{\theta_i\}} E(\{\Omega_i, \theta_i\}_i) = -\sum_i \int_{\Omega_i} \log p(f(x) \,|\, \hat{\theta}_i)\, dx + \nu\, |\mathcal{C}|, \quad (14)$$

such that

$$\hat{\theta}_i = \arg\min_{\theta} \left(-\int_{\Omega_i} \log p(f(x) \,|\, \theta)\, dx \right). \qquad (15)$$

In this case, the optimal model parameters $\hat{\theta}_i$ typically depend on the regions Ω_i. As pointed out by several authors [85, 81, 1], this region-dependence can be taken into account in the computation of accurate shape gradients. Exact shape gradients can also be applied with non-parametric density estimation techniques like the Parzen window method [47, 73, 40]. In [75], it is shown that no additional terms arise in the shape gradient if the distributions p_i are assumed to be Gaussian. And in [39], the authors point out that the additional terms are negligible in the case of Laplacian distributions.[6] We will therefore neglect higher-order terms in the computation of shape gradients and simply perform an alternating minimization of the energy (13) with respect to region boundaries and region models.

3.2. Two-Phase Level Set Formulation

Let us for the moment assume that the solution to (13) is in the class of binary (two-phase) segmentations, i.e. a partitioning of the domain Ω such that each pixel is ascribed to one of two possible phases. Extending the approach of Chan and Vese [12], one can implement the functional (13) by:

$$E(\phi, \{\theta_i\}) = -\int_{\Omega} H\phi \log p(f|\theta_1) - (1 - H\phi) \log p(f|\theta_2) + \nu\, |\nabla H\phi|\, dx, \quad (16)$$

where $H\phi$ denotes the heaviside step function defined as:

$$H\phi \equiv H(\phi) = \begin{cases} 1 & \text{if } \phi \geq 0 \\ 0 & \text{else} \end{cases}. \qquad (17)$$

[6] For a recent study of various noise models on level set segmentation see [61].

The first two terms in (16) model the areas inside and outside the contour while the last term represents the length of the separating interface.

Minimization is done by alternating a gradient descent for the embedding function ϕ (for fixed parameters θ_i):

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \left(\nu \operatorname{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) + \log \frac{p(f(x) \,|\, \theta_2)}{p(f(x) \,|\, \theta_1)} \right). \tag{18}$$

with an update of the parameters θ_i according to (15). In practice, the delta function δ is implemented by a smooth approximation – cf. [12].

3.3. Multiphase Level Set Formulation

Several authors have proposed level set formulations which can handle a larger number of phases [98, 96, 66, 8]. These methods use a separate level set function for each region. This clearly increases the computational complexity. Moreover, numerical implementations are somewhat involved since the formation of overlap and vacuum regions needs to be suppressed. By interpreting these overlap regions as separate regions, Chan and Vese derived an elegant formulation which only requires $\log_2(n)$ level set functions to model n regions. Each of the n regions is characterized by the various level set functions being either positive or negative. See [93] for details.

3.4. Scalar, Vector and Tensor-valued images

3.4.1. *Scalar images*

Let us consider a scalar image made up of two regions, the intensities of which are drawn from a Gaussian distribution:

$$p(I \,|\, \mu_i, \sigma_i^2) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \, e^{-\frac{(I-\mu_i)^2}{2\sigma_i^2}}, \quad i = \{1, 2\}. \tag{19}$$

This distribution can be injected in the general bi-partitioning energy (16). Given a partition of the image plane according to a level set function ϕ, optimal estimates for the mean μ_i and the variance σ_i can be computed analytically:

$$\begin{cases} \mu_1 = \frac{1}{a_1} \int H(\phi) I(x) dx & \sigma_1^2 = \frac{1}{a_1} \int H(\phi)(I(x) - \mu_1)^2 dx, \\ \mu_2 = \frac{1}{a_2} \int (1 - H(\phi)) I(x) dx, & \sigma_2^2 = \frac{1}{a_2} \int (1 - H(\phi))(I(x) - \mu_2)^2 dx. \end{cases}$$

where $a_1 = \int H(\phi) dx$ and $a_2 = \int (1 - H(\phi)) dx$ are the areas of the inside and outside region. For fixed model parameters, the gradient descent

equation for the level set function ϕ – see (18) – reads

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \left(\nu \operatorname{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) + \frac{(I-\mu_2)^2}{2\sigma_2^2} - \frac{(I-\mu_1)^2}{2\sigma_1^2} + \log \frac{\sigma_1}{\sigma_2} \right). \quad (20)$$

More details on this derivation when the parameters (μ_i, σ_i) are taken as functions of Ω_i can be found in [76]. We end up with an algorithm that alternates the estimation of the empirical intensity means and variances inside each region and the level set evolution described by equation (20). Regarding the complexity, each iteration of the level set evolution is applied only inside a narrow band around the zero-crossing because the Dirac function is equal to zero at other locations. More interesting is that the statistical parameters can also be updated with a similar complexity: new updates are functions of their previous values and of the pixels where the sign of ϕ changes. Assuming the evolving interface to visit each pixel only once, the total complexity is thus linear in the size of the image.

3.4.2. *Vector-valued images*

A direct extension to vector-valued images is to use *multivariate* Gaussian densities as region models. Region *pdf*s are then parameterized by a vector mean and a covariance matrix. Similarly to the scalar case, the optimal statistical parameters are their empirical estimates in the corresponding region. The 2-phase segmentation of an image I of any dimension can thus be obtained through the following level set evolution (cf. equation 18):

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \left(\nu \operatorname{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) + \log \frac{p(I(x) \,|\, \mu_2, \Sigma_2^2)}{p(I(x) \,|\, \mu_1, \Sigma_1^2)} \right), \quad (21)$$

with:

$$\begin{cases} \mu_i = \dfrac{1}{|\Omega_i|} \displaystyle\int_{\Omega_i} I(x)\, dx, \\ \Sigma_i = \dfrac{1}{|\Omega_i|} \displaystyle\int_{\Omega_i} (I(x) - \mu_i)(I(x) - \mu_i)^\top \, dx \quad \text{for} \quad i = 1, 2. \end{cases} \quad (22)$$

Like in the scalar case, the estimation of the statistical parameters can be optimized to avoid a full computation over the whole image domain at each iteration. Here, it becomes a bit more technical since cross-components products appear in the covariance matrices but the final complexity is identical to the one obtained in the scalar case [72].

3.4.3. *Tensor-valued images*

In order to apply the above statistical level set framework to the segmentation of tensor images, one needs to define appropriate distances on the space of tensors. Several approaches have been proposed to define distances from an information theoretic point of view by interpreting the tensors as parameterizations of 0-mean multivariate normal laws. The definition of a distance between tensors is then translated to one of a dissimilarity measure between probability distributions.

The symmetric KL divergence

Wang and Vemuri [94] applied the symmetrized Kulback-Leibler (SKL) divergence – also called J-divergence – to define the region term of the front evolution. For multivariate 0-mean normal laws with covariance matrices J_1 and J_2, the SKL divergence is given by:

$$\mathcal{D}(J_1, J_2)_{SKL} = \frac{1}{2}\sqrt{\text{trace}\left[J_1^{-1}J_2 + J_2^{-1}J_1\right] - 2n}, \tag{23}$$

where n is the dimension of the tensors. This measure has the advantage of being affine invariant and closed form expressions are available for the mean tensors which is particularly interesting to estimate region statistics. Region confidence were also incorporated in [77]. These works present several promising segmentation experiments on 2D [94] and 3D [77] real diffusion tensor images.

The Rao Distance

Another distance has been proposed in [52, 51] with the same idea of considering tensors as covariance matrices of multivariate normal distributions. Following [84], a Riemannian metric is introduced and the geodesic distance between two members of this family is given by

$$\mathcal{D}_G(J_1, J_2) = \sqrt{\frac{1}{2}\sum_{i=1}^{2}\log^2(\lambda_i)}, \tag{24}$$

where λ_i denote the eigenvalues of the matrix $J_1^{-1/2}J_2J_1^{-1/2}$. The same metric was proposed in [68] from a different viewpoint. It verifies the basic properties of a distance (positivity, symmetry and triangle inequality) and it is invariant to inversions: $\mathcal{D}(J_1, J_2) = \mathcal{D}(J_1^{-1}, J_2^{-1})$. The above metrics permit to define statistics on sets of SPD matrices which can be used to define the region term of the segmentation. It was also shown in [51] that the (asymmetric) Kullback Leibler divergence (23) is a Taylor approximation of the geodesic distance (24).

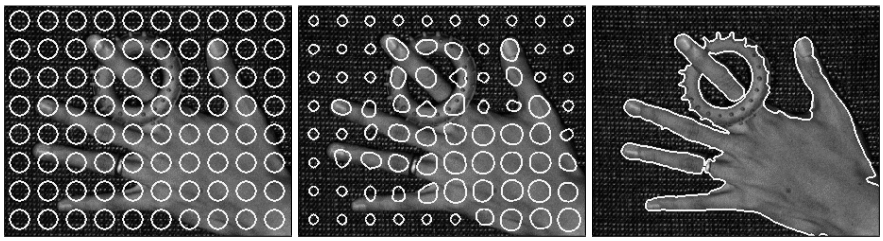


Figure 1. Curve evolution for the segmentation of a gray-level image using Gaussian intensity distributions to approximate region information.

In the following sections, we will exploit the statistical level set framework introduced above to construct segmentation schemes for color, texture, dynamic texture and motion. To this end, we will consider different choices regarding the features f – namely intensity values, color values, spatial structure tensors, spatio-temporal image gradients, or features modeling the local spatio-temporal dynamics – and respective sets of model parameters θ_i, modeling color or texture distributions or parametric motion in the separated regions. Moreover, we will consider different choices for the distributions p_i of these model parameters.

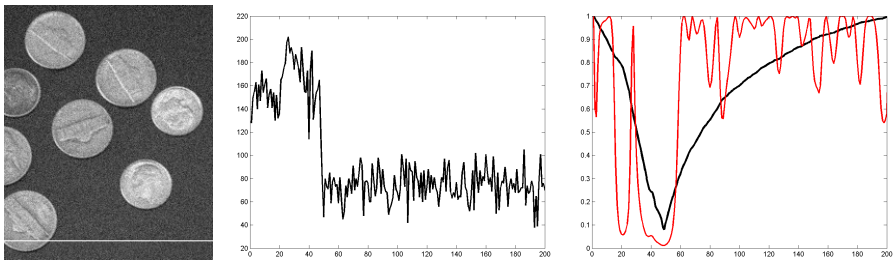
4. Intensity, Color and Texture

In the previous section, we considered Gaussian approximations for scalar and vector-values images. These models can be used to segment gray, color and texture images [12, 75, 73]. In the following, these models are applied to the segmentation of natural images. Curve evolutions are presented to illustrate the gradient descent driving to the segmentation.

4.1. Gray & color images

In Figure 1, we present the curve evolution obtained with the gray-value level set scheme of Section 3.4.1. The curve is initialized with a set of small circles and it successfully evolves toward the expected segmentation. Other initializations may be considered but using tiny circles provides a fast convergence speed and helps to detect small parts and holes. Note that changes of topology during the evolution are naturally handled by the implicit formulation.

We previously argued that the region-based formulation exhibits less local minima than approaches which solely rely on gradient information along the curve. To support this claim, we plotted in Figure 2 the



input image 1D intensity profile energy of cut

Figure 2. Comparison of edge- and region-based segmentation methods in 1D. For a 1D intensity profile of the coin image (taken along the line indicated in white), we computed the energy associated with a split of the interval at different locations. While the region-based energy exhibits a broad basin of attraction around a single minimum located at the boundary of the coin (thick black line), the energy of the edge-based approach is characterized by numerous local minima (red line). A gradient descent on the latter energy would not lead to the desired segmentation.

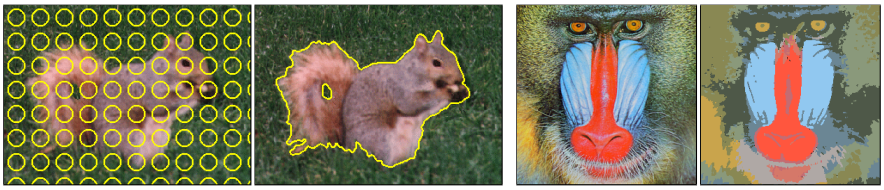


Figure 3. Binary segmentation of a color image using multivariate Gaussian distributions as region descriptor (initialization and final segmentation) and multiphase color segmentation obtained with the algorithm developed in [8].

empirical energy for the segmentation of a 1D slice of an image. In contrast to the edge-based energy, the region-based energy (thick black line) shows a single minimum corresponding to the boundary of the coin.

The region-based approach can directly be extended to color images by applying the vector-valued formulation of Section 3.4.2. The only point to be careful about is the choice of color space for the multivariate Gaussian model to make sense. The RGB space is definitely not the best one since, as can be seen from the MacAdam ellipse, the perception of color difference is nonlinear in this space. The CIE-lab space has been designed to approximate this nonlinearity by trying to mimic the logarithmic response of the eye. Figure 3 shows a two-phase and a multiphase example of vector-valued segmentation obtained on natural color images using this color space (the algorithm proposed in [8] was used for the multiphase implementation).

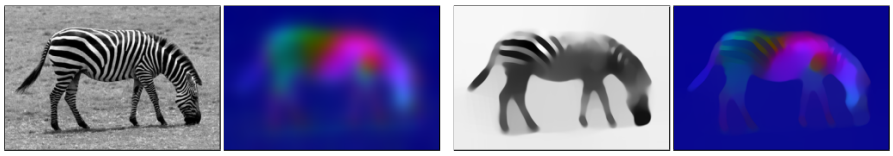


Figure 4. Left: Zebra image and color representation of its structure tensor (the components of the structure tensors are used as RGB components). Right: Intensity and structure tensor of the Zebra image after coupled nonlinear diffusion.

4.2. Texture

In gray and color image segmentation, pixel values are assumed to be spatially independent. This is not the case for textured images which are characterized by local correlations of pixel values. In the following, we will review a set of basic features which allow to capture these local correlations. More sophisticated features are conceivable as well [53].

4.2.1. *The nonlinear structure tensor as texture feature*

While texture analysis can rely on texture samples to learn accurate models [38, 26, 59, 83, 99], unsupervised image segmentation should learn these parameters on-line. Since high-order texture models introduce too many unknown parameter to be estimated in an unsupervised approach, more compact features are usually favored. Bigün et al. in [4] addressed this problem with the introduction of the structure tensor (also called second order moment matrix) which yields three different feature channels per scale. It has mainly been used to determine the intrinsic dimensionality of images in [3, 34] by providing a continuous measure to detect critical points like edges or corners. Yet, the structure tensor does not only give a scalar value reflecting the probability of an edge but it also includes the texture orientation. All these properties make this matrix a good descriptor for textures. The structure tensor [34, 4, 70, 55, 36] is given by the matrix of partial derivatives smoothed by a Gaussian kernel K_σ with standard deviation σ:

$$J_\sigma = K_\sigma * (\nabla I \nabla I^\top) = \begin{pmatrix} K_\sigma * I_{x_1}^2 & K_\sigma * I_{x_1} I_{x_2} \\ K_\sigma * I_{x_1} I_{x_2} & K_\sigma * I_{x_2}^2 \end{pmatrix}. \qquad (25)$$

For color images, all channels can be taken into account by summing the tensors of the individual channels [97].

Despite its good properties for texture discrimination, the structure tensor is invariant to intensity changes. In order to segment images with and without texture, a feature vector including the square root of

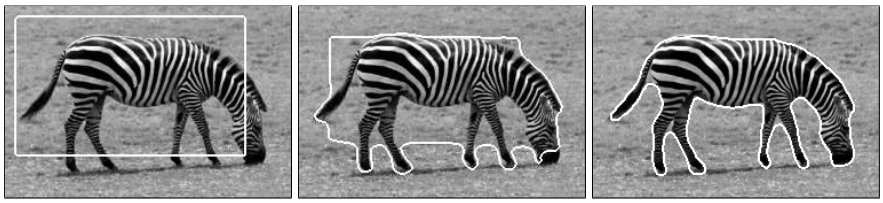


Figure 5. Curve evolution for the segmentation of a zebra image using the nonlinear structure tensor and the smoothed intensity (here, a rectangle is used as initialization but small circles also lead to a similar result).

the structure tensor and the intensity was defined in [73]:

$$f(x) = \left(I, \frac{I_{x_1}^2}{|\nabla I|}, \frac{I_{x_2}^2}{|\nabla I|}, \frac{2 I_{x_1} I_{x_2}}{|\nabla I|} \right)^T . \tag{26}$$

The major problem of the classic structure tensor is the dislocation of edges due to the smoothing with Gaussian kernels as shown in figure 4. To address this problem, Weickert and Brox proposed in [95] to replace the Gaussian smoothing by nonlinear diffusion, applying nonlinear matrix-valued diffusion schemes introduced in [89, 90]. Applied on the feature vector f, this nonlinear diffusion couples all channels by a joint diffusivity, the information of all channels is used to decide whether an edge is worth to be enhanced or not, leading to the simplification of the data, the removal of outliers, and the closing of structures. Figure 4 shows the features obtained on the *zebra* image.

The feature vectors resulting from the nonlinear diffusion form a vector-valued image which can be segmented using the vector-valued formulation presented in Section 3.4.2. Figure 5 shows a segmentation of the *zebra* image obtained with this method. For results on a wider range of texture images, we refer to [73].

The structure tensor is undoubtedly pertinent for texture discrimination but the approach developed so far still allows for further improvements. We shortly mention two recent extensions of this work in the following two paragraphs.

4.2.2. *Scale introduction via TV flow*

With the nonlinear structure tensor, one mainly considers a single scale for the whole image. Yet textures often differ from one another with respect to their intrinsic scale. In order to account for varying scale, a straightforward extension is to combine texture features at different scales. While this modification may integrate information at different scales, it also increases dramatically the number of channels and redundant information is introduced, making the second phase – the image

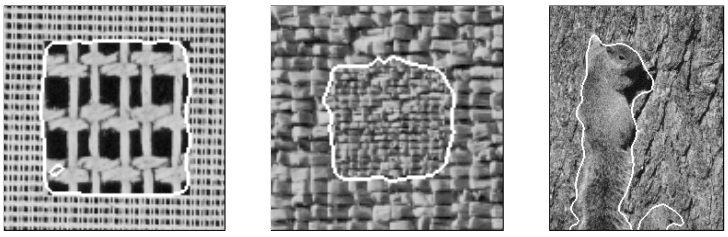


Figure 6. Segmentation results obtained by running a level set segmentation process on the 5-dimensional feature space given by the features of the structure tensor, the image intensity and a local scale measure computed from the speed of a total variation flow. Images are courtesy of Brox and Weickert [9].

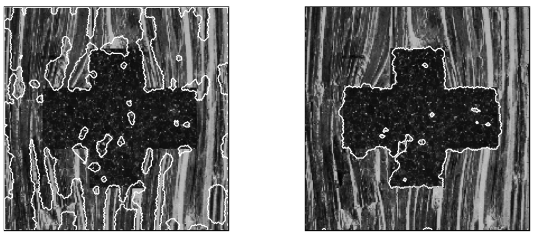


Figure 7. Segmentation of a textured image with different dissimilarity measures between tensors. The left result was obtained using the Frobenius norm while the right segmentation is based on the Rao distance (see text for details). Images courtesy of de Luís García and Deriche [27].

partitioning – more difficult. In order to work with a reduced feature space, Brox and Weickert [9] proposed an elegant and efficient extension of the above framework, which combines similar texture features as above with a local scale measure. By exploiting the linear contrast reduction property of the TV (total variation) flow:

$$\begin{cases} \dfrac{\partial f}{\partial t} = \operatorname{div}\left(\dfrac{\nabla f}{\sqrt{|\nabla f|^2 + \epsilon^2}}\right) \\ f(t=0) = I \end{cases} \tag{27}$$

the authors are able to extract a local scale measure computed from the speed of the diffusion process. Upon combining this scale with the intensity and orientation features in (26), one can perform segmentation in a 5-dimensional feature space. Figure 6 shows three representative segmentation results.

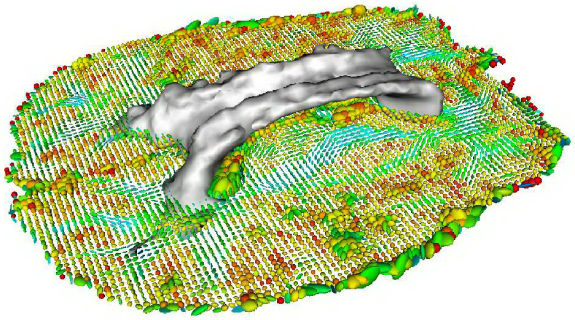


Figure 8. Segmentation of corpus callosum from a diffusion tensor image using the geodesic distance in the manifold of multivariate normal distributions. This 3D segmentation was obtained in [51] using the level set formulation presented in Section 2. Being implicit, the level set representation allows a straightforward extension to higher dimensions.

4.2.3. *Metric between tensors*

While the previous approaches construct a feature vector from the components of the structure tensor and apply a vector-valued segmentation scheme, on can directly define metrics on the space of structure tensors [27], for example the metrics defined in Section 3.4.3. The comparison in Figure 7 shows that appropriate tensor distances lead to drastic improvements in the segmentation.

5. Diffusion tensor images

The problem of segmenting tensor-valued data also appears in medical imaging with the relatively new modality of diffusion tensor magnetic resonance images. In these images, a diffusion tensor is measured at each voxel. This tensor captures the local motion of water molecules as approximated by a Gaussian law. The metric between tensors described in Section 3.4.3 can be used to segment these images. Figure 8 shows a segmentation of the corpus callosum obtained with the geodesic distance.[7]

[7] This result was obtained in [51] with data provided by J. F. Mangin and J. B. Poline.

6. Dynamic Texture

The texture segmentation framework detailed earlier is based on assigning local texture signatures to each image location. The subsequent integration into a level set framework aims at optimally grouping regions of similar signatures while imposing a length constraint on the separating boundary.

Given a video sequence of temporally varying textures – such as smoke on water – one can extend this concept to the space-time domain and group regions of similar spatio-temporal statistics. The first work addressing this problem was proposed in [32] where the authors made use of recent developments in the modeling of dynamic textures [31]. Due to the scope of this paper, we will merely review the key ideas.

Dynamic textures are models of temporally varying textures which assume the image sequence to be generated by a second-order stationary process. Experiments have demonstrated that numerous realistic image sequences, such as water waves, fluttering foliage, smoke and steam can be well synthesized by such Gauss-Markov processes [32].

More specifically, it is assumed that the temporally varying pixel intensities $\{I_i(t)\}_{i=1..m}$ can be approximated by a model $\{y_i(t)\}_{i=1..m}$ which is driven by a random process $r(t) \in I\!R^n$ as follows [31]:

$$\begin{cases} r(t+1) = Ar(t) + \sqrt{Q}\, v(t);\ r(0) = r_0 \\ y(t) = Cr(t) + \sqrt{R}\, w(t) \end{cases} \tag{28}$$

Here $y(t) \in I\!R^m$ represents the vector of intensities of all m pixels at time t, $v(t) \in I\!R^n$ and $w(t) \in I\!R^m$ are white zero-mean Gaussian processes, $A \in I\!R^{n\times n}$, $C \in I\!R^{m\times n}$ are the model parameters, and $Q \in I\!R^{n\times n}$, $R \in I\!R^{m\times m}$ are the noise covariance matrices. The model parameters A and C in (28) can be estimated from an image sequence $I(x,t)$ [31].

As suggested in [32], we can associate with each image location $x \in \Omega$ a local signature $\xi(x)$ characterizing the spatio-temporal dynamics at this location based on the model parameters A and C computed in a small spatial window. A meaningful signature cannot be directly defined on these model parameters because – as can be seen from the definition of (28) – there exists an entire equivalence class of model parameters which lead to the same dynamic texture[8]. Instead we define a local signature $\xi(x)$ by:

$$\xi(x) = \big(\cos\theta_1(x), \ldots, \cos\theta_n(x)\big), \tag{29}$$

[8] Substituting in (28) A with TAT^{-1}, C with CT^{-1}, Q with TQT^{-1}, and choosing the initial condition $Tr(0)$, where $T \in GL(n)$ is any invertible $n \times n$ matrix generates the same output covariance sequence.

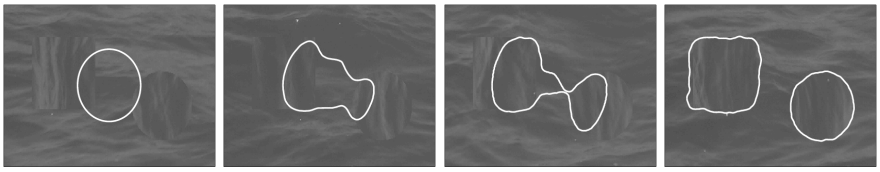


Figure 9. ***Segmentation by texture orientation:*** *Segmentation of two dynamic textures which differ only in orientation but share the same dynamics and general appearance (intensity values).*

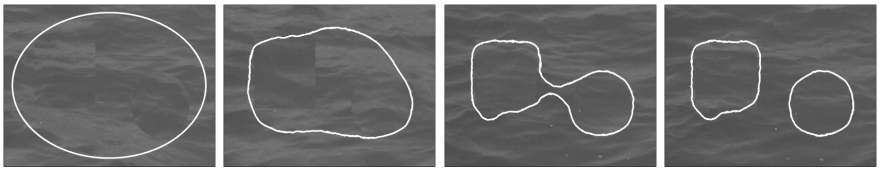


Figure 10. ***Segmentation by changing dynamics:*** *The two dynamic textures are identical in appearance, but differ in the dynamics. This particular segmentation problem is quite difficult, even for human observers. Segmentation is obtained exclusively on the basis of the temporal properties of the textures.*

where $\{\theta_i\}_{i=1..n}$ are the subspace angles associated with the equivalence classes of the model at location x and some reference model. More precisely, if A and B are two measurement matrices, then $\{\theta_i\}_{i=1..n}$ are given by the principal angles [15] between range(A) and range(B). For details on the computation of these angles, we refer to [32].

Assuming that the spatio-temporal signatures defined in (29) correspond to two Gaussian distributions, one can apply the vector-valued segmentation scheme introduced in Section 3.4.2 to group areas of similar spatio-temporal dynamics.

Due to the scope of this survey paper, we will merely show two complementary results obtained by the above segmentation scheme. Figure 6 shows the separation based on spatial orientation of a moving texture. The image data shows a water sequence for which we simply rotated two areas by 90 degrees. By construction, intensity characteristics and dynamics of the separated regions are identical, yet due to the different orientation they can be separated. Figure 6 shows a segmentation result which is complementary to the previous one. We generated a sequence containing regions which only differ with respect to their dynamics (but have identical spatial texture) by overlapping the ocean sequence in the regions corresponding to the disc and square over an ocean sequence slowed down by a factor of 2.

7. Motion

7.1. Motion as a criterion for segmentation

The central question underlying the construction of segmentation methods is to identify what properties characterize objects and distinguish them from other objects and from the background. In the previous sections, we reviewed level set methods which exploit low level properties such as color, texture or even dynamical texture. The respective image segmentation algorithms essentially group regions of similar low level properties.

Many objects in our environment are characterized by the fact that they move in a coherent manner. Figure 11, top row, shows the intensity-based segmentation of a single frame taken from an image sequence of two cars driving down the street.[9] The two cars and the background are moving in different directions. Clearly the individual cars are not homogeneous regarding their intensity or texture. A purely intensity-based segmentation therefore fails to separate the objects from the background.

In the following, we will detail how the statistical segmentation scheme presented above can be adapted to incorporate motion information given two consecutive frames from an image sequence. Minimization of the resulting cost functional leads to a segmentation of the scene in terms of piecewise parametric motion.[10] The present formulation was proposed in [17, 23] with an earlier (explicit contour) formulation in [21]. Related approaches were also proposed in [60, 67]. The central idea is that we do not precompute local motion vectors. Instead we jointly estimate the segmentation and the motion models for each of a set of regions by minimizing the proposed functional. In the notation introduced in Section 3, this means that – in contrast to the texture schemes of the previous Sections – the model parameters θ_i (the motion models of the separate regions) will not correspond to simple aggregates of the local feature vectors $f(x)$ (the space time gradients), but rather they will be derived quantities. Due to the scope of this article, we will constrain the presentation to the key ideas. For further details and a discussion of related approaches we refer the reader to [23]. For an extension of the proposed framework to the segmentation of space-time volumes given an entire video sequence, we refer to [22].

[9] `http://i21www.ira.uka.de/image_sequences/`

[10] In this paper, we are only concerned with 2D motion models. Such 2D motion models allow in particular to separate the different depth layers of a static scene filmed by a moving camera (cf. [22]). In terms of 3D motion, such a static scene instead corresponds to a *single* motion model.

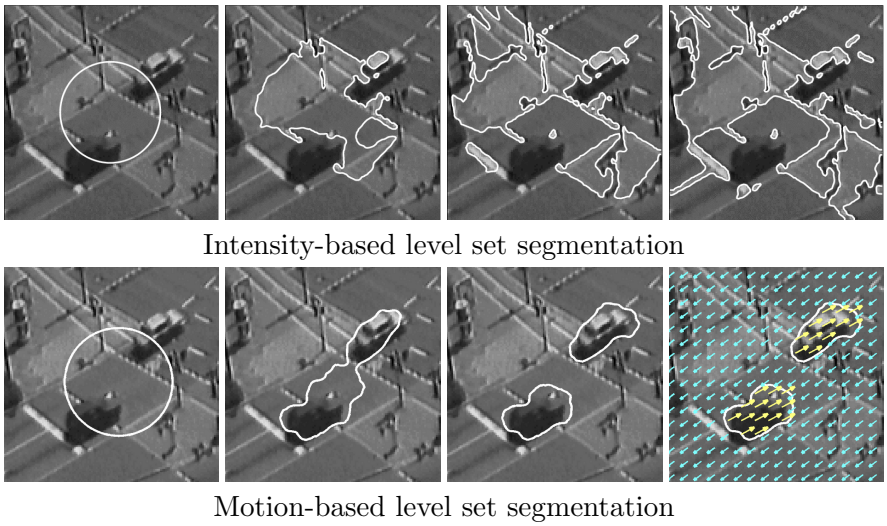


Figure 11. Intensity versus motion segmentation. Since cars and background are not well-defined in terms of homogeneous intensity, color or texture, unsupervised low-level segmentation schemes based on a single frame are unable to separate objects and background (top row). By minimizing the motion competition functional (38) with $\nu = 1.5$, one obtains a fairly accurate segmentation of the two cars and an estimate of the motion of cars and background.

7.2. Motion Competition

Let $I : \Omega \times \mathbb{R} \to \mathbb{R}$ be a gray value image sequence. Denote the spatio-temporal image gradient of $I(x,t)$ by

$$\nabla_3 I = \left(\frac{\partial I}{\partial x_1}, \frac{\partial I}{\partial x_2}, \frac{\partial I}{\partial t} \right)^\top . \tag{30}$$

Let $v : \Omega \to \mathbb{R}^3$, $v(x) = (u(x), w(x), 1)^\top$ be the velocity vector at a point x in homogeneous coordinates.[11]

Let us assume that the intensity of a moving point remains constant throughout time.[12] Expressed in differential form, this gives a relation between the spatio-temporal image gradient and the homogeneous velocity vector, known as the *optic flow constraint*:

$$\frac{dI}{dt} = \frac{\partial I}{\partial t} + \frac{\partial I}{\partial x_1}\frac{dx_1}{dt} + \frac{\partial I}{\partial x_2}\frac{dx_2}{dt} = v^\top \nabla_3 I = 0. \tag{31}$$

[11] Since we are only concerned with two consecutive frames from a sequence, we will drop the time coordinate in the notation of the velocity field.

[12] To allow for variation of the global illumination, one can alternatively assume constancy of higher-order derivatives (cf. [7]).

For the sake of segmentation, we will assume that the velocity in each of a set of regions can be modeled by a parametric motion of the form

$$v(x) = S(x) \cdot q, \tag{32}$$

with a space dependent matrix S and a parameter vector q. In particular, this includes the case of translational motion where S is the 3×3 unit matrix and $q = (u, w, 1)$ the vector of constant velocity in homogeneous coordinates. The parametric formulation (32) also includes the more general *affine motion model* with:

$$S(x) = \begin{pmatrix} x_1 & x_2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_1 & x_2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \text{ and } \; q = (a, b, c, d, e, f, 1)^\top \tag{33}$$

In the context of segmentation in space *and time*, this parametric formulation can be extended to incorporate temporally varying motion (such as acceleration and deceleration) – for details we refer to [22]. Inserting the parametric model (32) into the optic flow constraint (31) leads to a constraint on the relation between the parameter vector q and the space-time gradient $\nabla_3 I$ at a specific location:

$$\nabla_3 I^\top \, S(x) \, q = 0. \tag{34}$$

Neglecting the case that the space-time gradient vanishes, this constraint states that the two vectors q and $S(x)^\top \nabla_3 I(x)$ must be orthogonal. We therefore model the conditional probability to encounter a certain gradient measurement given a velocity model as a function of the angle α between the two vectors:

$$P(\nabla_3 I \,|\, q \,; x) \; \propto \; \exp\left(-\frac{(q^\top \, S^\top \nabla_3 I)^2}{|q|^2 \, |S^\top \nabla_3 I|^2} \right). \tag{35}$$

This expression is maximal if the two vectors are indeed orthogonal, it is minimal if the two vectors are parallel. Yet it does not depend on the length of the two vectors. Note that due to the introduction of a spatially parametric model, this conditional probability becomes space-dependent – this is in contrast to the space-independent conditional probabilities considered in the Sections 4, 4.2 and 6 on color and texture segmentation. Analogous parametric generalizations of intensity-based level set segmentation approaches have been proposed by Vese [92]. And corresponding extensions of the above texture segmentation schemes are certainly conceivable.

Based on the Bayesian formulation introduced in Section 3, we can integrate the conditional probability for a measurement given certain model parameters into a variational framework for segmentation.

Inserting equation (35) into energy (13), we obtain the functional

$$E(\mathcal{C}, \{q_i\}) = \sum_{i=1}^{n} \int_{\Omega_i} \frac{q_i^\top \, T(x) \, q_i}{|q_i|^2} \, dx \; + \nu \, |\mathcal{C}|, \tag{36}$$

where, for notational simplification, we have introduced the matrix

$$T(x) = \frac{\nabla_3 I \; S^\top S \; \nabla_3 I^\top}{|S^\top \nabla_3 I|^2}, \tag{37}$$

The corresponding two-phase level set implementation – cf. equation (16) – is given by

$$E(q_1, q_2, \phi) = \int_{\Omega} \frac{q_1^t T q_1}{|q_1|^2} H\phi + \frac{q_2^t T q_2}{|q_2|^2} \Big(1 - H\phi\Big) + \nu |\nabla H\phi| dx, \tag{38}$$

The first two terms in (38) enforce a homogeneity of the estimated motion in the two phases, while the last term enforces a minimal length of the region boundary given by the zero level set of ϕ. As discussed in 3.2, the functional (38) is optimized by alternating the estimate of the motion models q_1 and q_2 and an update of the level set function ϕ defining the motion boundaries.

For fixed level set function ϕ, i.e. fixed regions Ω_i, minimizing this functional with respect to the motion parameters $\{q_i\}$ results in a set of eigenvalue problems of the form:

$$q_i = \arg\min_q \frac{q^\top \, T_i \, q}{q^\top q}, \quad \text{with } T_i = \int_{\Omega_i} T(x) \, dx. \tag{39}$$

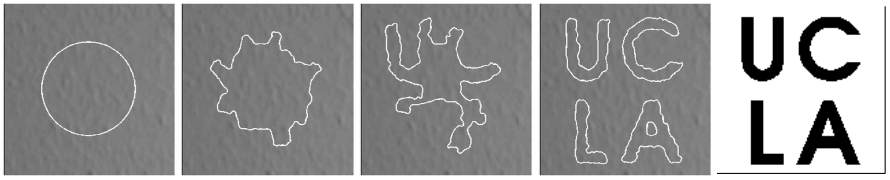
The parametric motion model q_i for each region Ω_i is therefore given by the eigenvector corresponding to the smallest eigenvalue of T_i. It is normalized, such that the third component is 1. Similar eigenvalue problems arise in motion estimation due to normalization with respect to the velocity magnitude (cf. [4, 43]).

Conversely, for fixed motion models q_i, a gradient descent on the energy (38) for the boundary $\mathcal{C}$ results an evolution equation – cf. (18) – of the form:

$$\frac{\partial \phi}{\partial t} = \delta(\phi) \left(\nu \, \mathrm{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) + e_2 - e_1 \right). \tag{40}$$

where

$$e_i = -\log P(\nabla_3 I \,|\, q_i \,; x) = \frac{q_i^t \, T \, q_i}{q_i^t q_i} = \frac{q_i^t \, \nabla_3 I \, S^\top S \, \nabla_3 I^\top \, q_i}{|q_i|^2 \, |S^\top \nabla_3 I|^2} \tag{41}$$



Contour evolution for motion segmentation **True regions**

Figure 12. Contour evolution obtained with functional (38) for $\nu = 0.06$, superimposed on one of the two input frames. The input images show the text region (right image) of the wallpaper moving right and the remainder moving left. The moving regions are accurately reconstructed, although the input images exhibit little in terms of salient features. The contour evolution took 10 seconds in Matlab.

are the motion energy densities associated with the respective regions.

Note that – as in the previous sections – we have neglected in the evolution equation (40) higher-order terms which account for the dependence of the motion parameters q_i on the level set function ϕ. An Eulerian accurate shape optimization scheme as presented for example in [42] is the focus of ongoing research.

The two terms in the contour evolution (40) have the following intuitive interpretation: The first term aims at minimizing the length of the separating motion boundary. The second term is proportional to the difference of the energy densities e_i in the regions adjoining the boundary: The neighboring regions compete for the boundary in terms of their motion energy density, thereby maximizing the motion homogeneity. For this reason we refer to this process as *Motion Competition*.

7.3. Experimental results

All image segmentation models are based on a number of more or less explicitly stated assumptions about the properties which define the objects of interest. The motion competition model is based on the assumption that objects are defined in terms of homogeneously moving regions. It extends the Mumford-Shah functional of piecewise constant intensity to a model of piecewise parametric motion. Despite this formal similarity, the segmentations generated by the motion competition framework are very different from those of its gray value analogue. Figure 11, bottom row, shows the boundary evolution obtained by minimizing the motion segmentation functional (38) and the corresponding motion estimates superimposed on the first frame. In contrast to its gray value analogue, the energy minimization simultaneously generates a fairly accurate segmentation of the two cars and an estimate of the motion of cars and background.

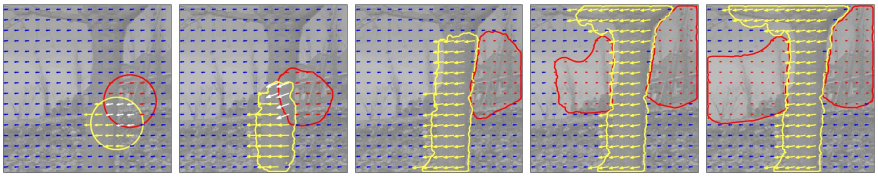


Figure 13. Multiphase motion segmentation. Contour evolution for a multiphase implementation of motion competition on two consecutive frames from the flower garden sequence. A static scene filmed by a moving camera is partitioned into layers of different depth. See [23] for details.

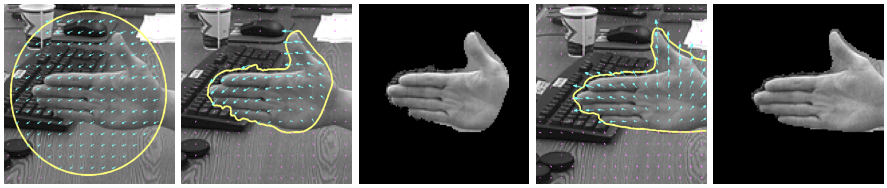


Figure 14. Piecewise affine motion segmentation. Segmentations obtained by minimizing functional (38) with $\nu = 8 \cdot 10^{-5}$ for two image pairs showing a hand rotating (top) and moving toward the camera (bottom).

Figure 12 shows segmentation the contour evolution generated by minimizing functional (38) for two wall paper images with the text region (right image) moving to the right and the remainder of the image plane moving to the left. Even for human observers the differently moving regions are difficult to detect – similar to a camouflaged lizard moving on a similarly-textured ground. The gradient descent evolution superimposed on one of the two frames gradually separates the two motion regions without requiring salient features such as edges or Harris corner points. Figure 13 shows results obtained with a multiphase implementation of the motion competition functional. The static scene filmed by a moving camera is segmented into layers of different depth.

The functional (38) allows to segment piecewise affine motion fields. In particular, this class of motion models includes rotation and expansion/contraction. Figure 14 shows segmentations obtained for a hand in a cluttered background rotating (in the camera plane) and moving toward the camera. In this example the object of interest can be extracted from a fairly complex background based exclusively on their motion.

8. Statistical Shape Priors for Level Set Segmentation

In the previous sections, we reviewed a number of approaches which allow to drive the level set segmentation based on various low-level assumptions regarding the intensity, color, texture or motion of objects and background. In numerous real-world applications, these approaches may fail to generate the desired segmentations, because the respective assumptions about the low-level properties are either insufficient or even violated. In certain medical images for example, object and background may exhibit very similar intensity characteristics. Moreover, the observed intensity or color of a 3D object may not be uniform due to directional lighting and cast shadows. And finally, misleading low-level information may arise due to noise or partial occlusion of the objects of interest. While the generic constraint (9) on the length of the segmenting boundary helps to cope with a certain amount of noise, it does introduce a bias toward contours of smaller length, thereby rounding corners or suppressing small scale details.

Beyond simple geometric regularity, the Bayesian formulation of the image segmentation problem allows to introduce higher-level prior knowledge about the shape of expected objects. This idea was pioneered by Grenander and coworkers [37]. In the following, we will briefly list some of the key contributions in the field of shape priors for level set segmentation.

The first application of shape priors for level set segmentation was developed by Leventon et al. [54] who propose to perform principal component analysis on a set of signed distance function embedding a set of sample shapes. The distance functions are sampled on a regular grid to obtain a vector representation. A term is added to the contour evolution equation to drive the embedding function to the most likely shape of the estimated distribution. Tsai et al. [86, 88] proposed a very efficient implementation of shape-driven level set segmentation by directly optimizing in the linear subspace spanned by the principal components. A detailed analysis of various shape distances and statistical shape analysis in the level set formulation can be found in [13]. Figure 15 shows the effect of variation along the first principal component on the embedding function and the implicitly represented contour.

The use of principal component analysis to model level set based shape distributions has two limitations: Firstly, the space of signed distance functions is not a linear space, i.e. arbitrary linear combinations of signed distance functions will in general not correspond to a signed distance function. Secondly, while the first few principal components capture (by definition) the most variation on the space of embedding functions, they will not necessarily capture the variation on

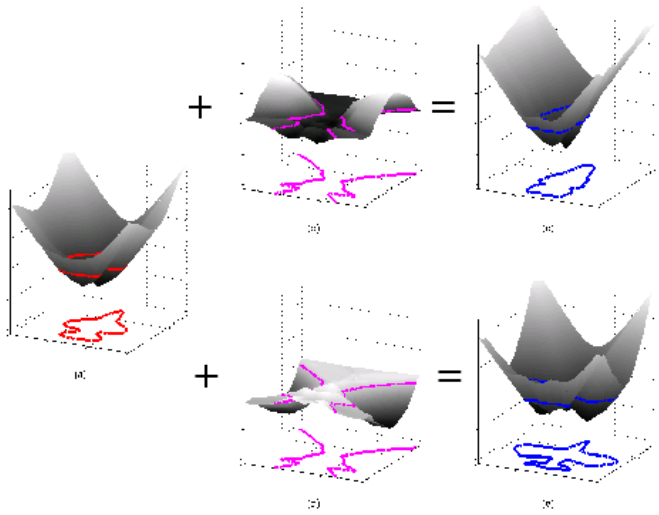


Figure 15. Visualization of principal component analysis on the level set function. The images show the mean level set function (obtained on a set of airplane shapes), and its deformation along the first eigenmode. Image data courtesy of [88].

the space of the embedded contours. As a consequence, one may need to include a larger number of eigenmodes (compared to PCA on explicit contours) in order to capture certain details of the modeled shape. Nevertheless, we found the PCA representation to work fairly well in practical applications. An alternative linear shape representation on the basis of harmonic embedding has been studied in [33]. Chen et al. [14] proposed to impose shape information on the zero crossing (rather than on the level set function). Rousson et al. proposed variational integrations of the shape prior [78, 79] based on the assumption of a Gaussian distribution. The use of nonparametric density estimation to model larger classes of level set based shape distributions was developed in [20, 74]. This approach allows to model distributions of shape which are not Gaussian – such as the various views of a 3D object [19] or the silhouettes of a walking person [20]. Moreover, in the limit of large sample size, the nonparametric estimator constrains the distribution to the vicinity of the training shapes, such that the distribution favors shapes which are signed distance functions. A method to simultaneously impose shape information about several objects into level set based segmentation and to induce a recognition-driven segmentation through the competition of shape priors was developed in [24]. Dynamical statistical shape priors for implicit shape representations were proposed in [18]. The latter approach takes into account that in the context of image sequence segmentation, the probability of a contour

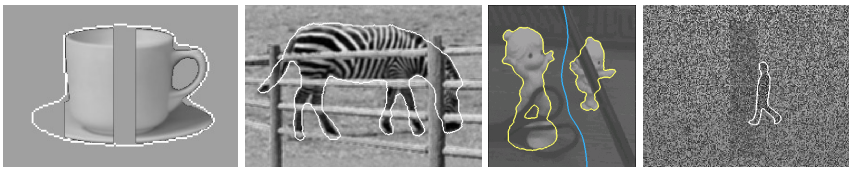


Figure 16. Sample segmentations using statistical shape priors. From left to right, the shape priors are static (a single shape), uniformly distributed in the PCA subspace, automatically selected from multiple shape instances [24] and dynamical [18].

will depend on which contours have been observed in previous frames. The respective shape models capture the temporal correlations among silhouettes which characterize many deforming shapes.

In Figure 16, we show a selection of segmentations obtained with some of the above methods. For further details we refer the reader to the respective publications.

9. Conclusion

We presented a survey of the class of region-based level set segmentation methods and detailed how they can be derived from a common statistical framework. The common goal of these approaches is to identify boundaries such that the color, texture, dynamic texture or motion in each of the separated regions is optimally approximated by simple statistical models.

Given a set of features or measurements $f(x)$ at each image location, minimization of the respective cost functionals leads to an estimation of a boundary C and a set of parameter vectors $\{\theta_i\}$ associated with each of the separated regions. Depending on the chosen segmentation criterion, the features f may be the pixel colors, the local structure tensors or the spatio-temporal intensity gradients, while the parameter vectors $\{\theta_i\}$ model distributions of intensity, color, texture or motion. The model parameters $\{\theta_i\}$ can be either simple aggregates of the features (as in the cases of color, texture or dynamic texture presented here) or derived quantities – as in the case of motion which is computed from the aggregated space-time gradients. The boundary $C \subset \Omega$ is implemented as the zero-crossing of an embedding function $\phi : \Omega \rightarrow I\!R$. Energy minimization leads to a gradient descent evolution of the embedding function interlaced with an update of the parameter vectors $\{\theta_i\}$ modeling the statistical distributions in the separated regions.

In numerous experimental results, we demonstrate that this class of level set methods allows to partition images into domains of coherent color, texture, dynamic texture or motion. In particular, we show that – in contrast to the traditional edge-based segmentation schemes, these region-based approaches are quite robust to noise and to varying initialization, making them well-suited for local optimization methods such as the level set method. We ended by reviewing some recent advances regarding the introduction of statistical shape knowledge into level set based segmentation schemes.

Acknowledgments

We thank Christoph Schnörr for helpful comments on the manuscript. We also thank Zhizhou Wang for fruitful discussions, and Thomas Brox, Rodrigo de Luís García, Christophe Lenglet, Gianfranco Doretto, Paolo Favaro, Stefano Soatto and Anthony Yezzi for providing image examples of their methods.

References

1. G. Aubert, M. Barlaud, O. Faugeras, and S. Jehan-Besson. Image segmentation using active contours: Calculus of variations or shape gradients? *SIAM Journal of Applied Mathematics*, 63(6):2128–2154, 2003.
2. J. Besag. On the statistical analysis of dirty pictures. *J. Roy. Statist. Soc., Ser. B.*, 48(3):259–302, 1986.
3. J. Bigün and G. Granlund. Optimal orientation detection of linear symmetry. In *Proceedings of the 1st International Conference on Computer Vision*, pages 433–438, London, England, June 1987. IEEE Computer Society Press.
4. J. Bigün, G. H. Granlund, and J. Wiklund. Multidimensional orientation estimation with applications to texture analysis and optical flow. *IEEE PAMI*, 13(8):775–790, 1991.
5. A. Blake and A. Zisserman. *Visual Reconstruction*. MIT Press, 1987.
6. Y. Boykov, O. Veksler, and R. Zabih. Fast approximate energy minimization via graph cuts. *IEEE PAMI*, 23(11):1222–1239, 2001.
7. T. Brox, A. Bruhn, N. Papenberg, and J. Weickert. High accuracy optical flow estimation based on a theory for warping. In T. Pajdla and V. Hlavac, editors, *European Conf. on Computer Vision*, volume 3024 of *LNCS*, pages 25–36, Prague, 2004. Springer.
8. T. Brox and J. Weickert. Level set based image segmentation with multiple regions. In *26th DAGM*, pages 415–423, Tübingen, Germany, August 2004. In Pattern Recognition, Springer LNCS 3175, C.-E. Rasmussen, H. Bülthoff, M. Giese, and B. Schölkopf (Eds.).
9. T. Brox and J. Weickert. A TV flow based local scale measure for texture discrimination. In T. Pajdla and V. Hlavac, editors, *European Conf. on Computer Vision*, volume 3022 of *LNCS*, pages 578–590, Prague, 2004. Springer.

10. V. Caselles, F. Catté, T. Coll, and F. Dibos. A geometric model for active contours in image processing. *Numer. Math.*, 66:1–31, 1993.
11. V. Caselles, R. Kimmel, and G. Sapiro. Geodesic active contours. In *Proc. IEEE Intl. Conf. on Comp. Vis.*, pages 694–699, Boston, USA, 1995.
12. T.F. Chan and L.A. Vese. Active contours without edges. *IEEE Trans. Image Processing*, 10(2):266–277, 2001.
13. G. Charpiat, O. Faugeras, and R. Keriven. Approximations of shape metrics and application to shape warping and empirical shape statistics. *Journal of Foundations Of Computational Mathematics*, 5(1):1–58, 2005.
14. Y. Chen, H. Tagare, S. Thiruvenkadam, F. Huang, D. Wilson, K. S. Gopinath, R. W. Briggs, and E. Geiser. Using shape priors in geometric active contours in a variational framework. *Int. J. of Computer Vision*, 50(3):315–328, 2002.
15. K. De Cock and B. De Moor. Subspace angles between linear stochastic models. In *Int. Conf. on Decision and Control*, volume 2, pages 1561–1566, Dec 2000.
16. L. D. Cohen and I. Cohen. Finite-element methods for active contour models and balloons for 2-d and 3-d images. *IEEE PAMI*, 15(11):1131–1147, 1993.
17. D. Cremers. A variational framework for image segmentation combining motion estimation and shape regularization. In C. Dyer and P. Perona, editors, *IEEE Conf. on Comp. Vis. and Patt. Recog.*, volume 1, pages 53–58, June 2003.
18. D. Cremers and G. Funka-Lea. Dynamical statistical shape priors for level set based tracking. In N. Paragios, editor, *Intl. Workshop on Variational and Level Set Methods*, LNCS. Springer, 2005. To appear.
19. D. Cremers, T. Kohlberger, and C. Schnörr. Nonlinear shape statistics in Mumford–Shah based segmentation. In A. Heyden et al., editors, *Europ. Conf. on Comp. Vis.*, volume 2351 of *LNCS*, pages 93–108, Copenhagen, May 2002. Springer.
20. D. Cremers, S. J. Osher, and S. Soatto. Kernel density estimation and intrinsic alignment for knowledge-driven segmentation: Teaching level sets to walk. In C. E. Rasmussen, editor, *Pattern Recognition*, volume 3175 of *LNCS*, pages 36–44. Springer, 2004.
21. 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.
22. D. Cremers and S. Soatto. Variational space-time motion segmentation. In B. Triggs and A. Zisserman, editors, *IEEE Int. Conf. on Computer Vision*, volume 2, pages 886–892, Nice, Oct. 2003.
23. D. Cremers and S. Soatto. Motion Competition: A variational framework for piecewise parametric motion segmentation. *Int. J. of Computer Vision*, 62(3):249–265, May 2005.
24. D. Cremers, N. Sochen, and C. Schnörr. A multiphase dynamic labeling model for variational recognition-driven image segmentation. *Int. J. of Computer Vision*, 2005. To appear.
25. D. Cremers, F. Tischhäuser, J. Weickert, and C. Schnörr. Diffusion Snakes: Introducing statistical shape knowledge into the Mumford–Shah functional. *Int. J. of Computer Vision*, 50(3):295–313, 2002.
26. G. Cross and A. Jain. Markov random field texture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 5:25–39, 1983.

27. R. de Luís García and R. Deriche. Tensor processing for texture and colour segmentation. Research report (to appear), INRIA, 2005.
28. H. Delingette and J. Montagnat. New algorithms for controlling active contours shape and topology. In D. Vernon, editor, *Proc. of the Europ. Conf. on Comp. Vis.*, volume 1843 of *LNCS*, pages 381–395. Springer, 2000.
29. A. Dervieux and F. Thomasset. A finite element method for the simulation of Raleigh-Taylor instability. *Springer Lect. Notes in Math.*, 771:145–158, 1979.
30. A. Dervieux and F. Thomasset. Multifluid incompressible flows by a finite element method. *Lecture Notes in Physics*, 11:158–163, 1981.
31. G. Doretto, A. Chiuso, Y. N. Wu, and S. Soatto. Dynamic textures. *Int. Journal of Computer Vision*, 51(2):91–109, February 2003.
32. G. Doretto, D. Cremers, P. Favaro, and S. Soatto. Dynamic texture segmentation. In B. Triggs and A. Zisserman, editors, *IEEE Int. Conf. on Computer Vision*, volume 2, pages 1236–1242, Nice, Oct. 2003.
33. A. Duci, A. Yezzi, S. Mitter, and S. Soatto. Shape representation via harmonic embedding. In *ICCV*, pages 656–662, 2003.
34. M. A. Förstner and E. Gülch. A fast operator for detection and precise location of distinct points, corners and centers of circular features. In *Proceedings of the Intercommission Workshop of the International Society for Photogrammetry and Remote Sensing*, Interlaken, Switzerland, 1987.
35. S. Geman and D. Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE PAMI*, 6(6):721–741, 1984.
36. G. H. Granlund and H. Knutsson. *Signal Processing for Computer Vision*. Kluwer Academic Publishers, 1995.
37. U. Grenander, Y. Chow, and D. M. Keenan. *Hands: A Pattern Theoretic Study of Biological Shapes*. Springer, New York, 1991.
38. M. Hassner and J. Sklansky. The use of Markov random fields as models of texture. *Computer Graphics and Image Processing*, 12:357–370, 1980.
39. M. Heiler and C. Schnörr. Natural image statistics for natural image segmentation. In *IEEE Int. Conf. on Comp. Vis.*, pages 1259–1266, 2003.
40. A. Herbulot, S. Jehan-Besson, M. Barlaud, and G. Aubert. Shape gradient for multi-modal image segmentation using mutual information. In *Int. Conf. on Image Processing*, 2004.
41. E. Ising. Beitrag zur Theorie des Ferromagnetismus. *Zeitschrift für Physik*, 23:253–258, 1925.
42. S. Jehan-Besson, M. Barlaud, and G. Aubert. DREAM2S: Deformable regions driven by an eulerian accurate minimization method for image and video segmentation. *Int. J. of Computer Vision*, 53(1):45–70, 2003.
43. A. Jepson and M.J. Black. Mixture models for optic flow computation. In *Proc. IEEE Conf. on Comp. Vision Patt. Recog.*, pages 760–761, 1993.
44. M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *Int. J. of Computer Vision*, 1(4):321–331, 1988.
45. J. Keuchel, C. Schnörr, C. Schellewald, and D. Cremers. Binary partitioning, perceptual grouping, and restoration with semidefinite programming. *IEEE PAMI*, 25(11):1364–1379, 2003.
46. S. Kichenassamy, A. Kumar, P. J. Olver, A. Tannenbaum, and A. J. Yezzi. Gradient flows and geometric active contour models. In *IEEE Intl. Conf. on Comp. Vis.*, pages 810–815, 1995.
47. J. Kim, J. W. Fisher, A. Yezzi, M. Cetin, and A. Willsky. Nonparametric methods for image segmentation using information theory and curve evolution. In *Int. Conf. on Image Processing*, volume 3, pages 797–800, 2002.

48. J.-O. Lachaud and A. Montanvert. Deformable meshes with automated topology changes for coarse-to-fine three-dimensional surface extraction. *Medical Image Analysis*, 3(2):187–207, 1999.
49. Y. G. Leclerc. Constructing simple stable description for image partitioning. *The International Journal of Computer Vision*, 3(1):73–102, 1989.
50. F. Leitner and P. Cinquin. Complex topology 3d objects segmentation. In *SPIE Conf. on Advances in Intelligent Robotics Systems*, volume 1609, Boston, November 1991.
51. C. Lenglet, M. Rousson, and R. Deriche. Toward segmentation of 3D probability density fields by surface evolution: Application to diffusion MRI. INRIA Research Report, June 2004.
52. C. Lenglet, M. Rousson, R. Deriche, and O. Faugeras. Statistics on multivariate normal distributions: A geometric approach and its application to diffusion tensor MRI. INRIA Research Report, June 2004.
53. T. Leung and J. Malik. Representing and recognizing the visual appearance of materials using three-dimensional textons. *Int. J. of Computer Vision*, 43(1):29–44, 2001.
54. M. Leventon, W. Grimson, and O. Faugeras. Statistical shape influence in geodesic active contours. In *CVPR*, volume 1, pages 316–323, Hilton Head Island, SC, 2000.
55. T. Lindeberg. *Scale-Space Theory in Computer Vision*. Kluwer Academic Publishers, 1994.
56. J. Malik, S. Belongie, T. Leung, and J. Shi. Contour and texture analysis for image segmentation. *Int. J. of Computer Vision*, 43(1):7–27, 2001.
57. R. Malladi, J. A. Sethian, and B. C. Vemuri. Evolutionary fronts for topology-independent shape modeling and recovery. In *Europ. Conf. on Computer Vision*, volume 1, pages 3–13, 1994.
58. R. Malladi, J. A. Sethian, and B. C. Vemuri. A topology independent shape modeling scheme. In *SPIE Conf. on Geometric Methods in Comp. Vision II*, volume 2031, pages 246–258, 1994.
59. S. Mallat. Multiresolution approximations and wavelet orthonormal bases of $L^2(R)$. *Trans. Amer. Math. Soc.*, 315:69–87, 1989.
60. A. Mansouri, A. Mitiche, and R. Feghali. Spatio-temporal motion segmentation via level set partial differential equations. In *Proc. of the 5th IEEE Southewst Symposium on Image Analysis and Interpretation (SSIAI)*, Santa Fe, 2002.
61. P. Martin, P. Refregier, F. Goudail, and F. Guerault. Influence of the noise model on level set active contour segmentation. *IEEE PAMI*, 26(6):799–803, June 2004.
62. T. McInerney and D. Terzopoulos. Topologically adaptable snakes. In *Proc. 5th Int. Conf. on Computer Vision*, pages 840–845, Los Alamitos, California, June 20–23 1995. IEEE Comp. Soc. Press.
63. D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. Pure Appl. Math.*, 42:577–685, 1989.
64. D. Nain, A. Yezzi, and G. Turk. Vessel segmentation using a shape driven flow. In *MICCAI*, pages 51–59, 2003.
65. S. J. Osher and J. A. Sethian. Fronts propagation with curvature dependent speed: Algorithms based on Hamilton–Jacobi formulations. *J. of Comp. Phys.*, 79:12–49, 1988.

66. N. Paragios and R. Deriche. Geodesic active regions: a new paradigm to deal with frame partition problems in computer vision. *Journal of Visual Communication and Image Representation*, 13(1/2):249–268, 2002.
67. N. Paragios and R. Deriche. Geodesic active regions and level set methods for motion estimation and tracking. *Computer Vision and Image Understanding*, 97(3):259–282, 2005.
68. X. Pennec, P. Fillard, and N. Ayache. A riemannian framework for tensor computing. *International Journal of Computer Vision*, 65(1), October 2005.
69. W. A. Perkins. Area segmentation of images using edge points. *IEEE PAMI*, 2(1):8–15, 1980.
70. A. R. Rao and B. G. Schunck. Computing oriented texture fields. *CVGIP: Graphical Models and Image Processing*, 53:157–185, 1991.
71. M. Rochery, I. Jermyn, and J. Zerubia. Higher order active contours. *Int. J. of Computer Vision*, 2005. To appear.
72. M. Rousson. *Cues Integrations and Front Evolutions in Image Segmentation*. PhD thesis, Université de Nice-Sophia Antipolis, December 2004.
73. M. Rousson, T. Brox, and R. Deriche. Active unsupervised texture segmentation on a diffusion based feature space. In *Proc. IEEE Conf. on Comp. Vision Patt. Recog.* , pages 699–704, Madison, WI, 2003.
74. M. Rousson and D. Cremers. Efficient kernel density estimation of shape and intensity priors for level set segmentation. In *MICCAI*, 2005. To appear.
75. M. Rousson and R. Deriche. A variational framework for active and adaptative segmentation of vector valued images. In *Proc. IEEE Workshop on Motion and Video Computing*, pages 56–62, Orlando, Florida, December 2002.
76. M. Rousson and R. Deriche. A variational framework for active and adaptative segmentation of vector valued images. RR 4515, INRIA, July 2002.
77. M. Rousson, C. Lenglet, and R. Deriche. Level set and region based surface propagation for diffusion tensor MRI segmentation. In *Computer Vision Approaches to Medical Image Analysis (CVAMIA) and Mathematical Methods in Biomedical Image Analysis (MMBIA) Workshop*, Prague, May 2004.
78. M. Rousson and N. Paragios. Shape priors for level set representations. In A. Heyden et al., editors, *Proc. of the Europ. Conf. on Comp. Vis.*, volume 2351 of *LNCS*, pages 78–92, Copenhagen, May 2002. Springer, Berlin.
79. M. Rousson, N. Paragios, and R. Deriche. Implicit active shape models for 3d segmentation in MRI imaging. In *MICCAI*, pages 209–216, 2004.
80. C. Samson, L. Blanc-Féraud, G. Aubert, and J. Zérubia. A variational model for image classification and restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(5):460–472, May 2000.
81. C. Schnörr. Computation of discontinuous optical flow by domain decomposition and shape optimization. *Int. J. of Computer Vision*, 8(2):153–165, 1992.
82. J. Shi and J. Malik. Normalized cuts and image segmentation. In *Proc. IEEE Conf. on Comp. Vision Patt. Recog. (CVPR'97)*, San Juan, Puerto Rico, 1997.
83. P. Simoncelli, W. Freeman, H. Adelson, and Heeger J. Shiftable multiscale transforms. *IEEE trans. on Information Theory*, 38:587–607, 1992.
84. L. T. Skovgaard. A Riemannian geometry of the multivariate normal model. *Scand. J. Statistics*, 11:211–223, 1984.
85. J. Sokolowski and J.P. Zolesio. *Introduction to shape optimization*. Computational Mathematics. Springer Verlag, 1991.

86. A. Tsai, A. Yezzi, W. Wells, C. Tempany, D. Tucker, A. Fan, E. Grimson, and A. Willsky. Model–based curve evolution technique for image segmentation. In *Comp. Vision Patt. Recog.*, pages 463–468, Kauai, Hawaii, 2001.
87. A. Tsai, A. J. Yezzi, and A. S. Willsky. Curve evolution implementation of the Mumford-Shah functional for image segmentation, denoising, interpolation, and magnification. *IEEE Trans. on Image Processing*, 10(8):1169–1186, 2001.
88. A. Tsai, A. J. Yezzi, and A. S. Willsky. A shape-based approach to the segmentation of medical imagery using level sets. *IEEE Trans. on Medical Imaging*, 22(2):137–154, 2003.
89. D. Tschumperlé and R. Deriche. Diffusion tensor regularization with constraints preservation. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Kauai Marriott, Hawaii, December 2001.
90. D. Tschumperlé and R. Deriche. Regularization of orthonormal vector sets using coupled PDE's. In *IEEE Workshop on Variational and Level Set Methods*, pages 3–10, Vancouver, Canada, July 2001.
91. G. Unal, H. Krim, and A. Y. Yezzi. Information-theoretic active polygons for unsupervised texture segmentation. *Int. J. of Computer Vision*, May 2005.
92. L. A. Vese. Multiphase object detection and image segmentation. In S.J. Osher and N. Paragios, editors, *Geometric Level Set Methods in Imaging, Vision and Graphics*, pages 175–194, New York, 2003. Springer.
93. L. A. Vese and T. F. Chan. A multiphase level set framework for image segmentation using the Mumford and Shah model. *The International Journal of Computer Vision*, 50(3):271–293, 2002.
94. Z. Wang and B. C. Vemuri. An affine invariant tensor dissimilarity measure and its application to tensor-valued image segmentation. In *IEEE Conference on Computer Vision and Pattern Recognition*, Washington, DC., June 2004.
95. J. Weickert and T. Brox. Diffusion and regularization of vector and matrix-valued images. In *Contemporary Mathematics*, volume 313, pages 251–268, 2002.
96. A. Yezzi, A. Tsai, and A. Willsky. A statistical approach to snakes for bimodal and trimodal imagery. In *Proceedings of the 7th International Conference on Computer Vision*, volume II, pages 898–903, Kerkyra, Greece, 1999.
97. S. Di Zenzo. A note on the gradient of a multi-image. *Computer Vision, Graphics, and Image Processing*, 33:116–125, 1986.
98. H.-K. Zhao, T. Chan, B. Merriman, and S. Osher. A variational level set approach to multiphase motion. *J. of Comp. Phys.*, 127:179–195, 1996.
99. S. C. Zhu, Y. Wu, and D. Mumford. Filters, random fields and maximum entropy (frame). *The International Journal of Computer Vision*, 27(2):1–20, 1998.
100. S. C. Zhu and A. Yuille. Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation. *IEEE PAMI*, 18(9):884–900, 1996.