



Natural Image Statistics for Natural Image Segmentation

MATTHIAS HEILER AND CHRISTOPH SCHNÖRR

*Computer Vision, Graphics, and Pattern Recognition Group, Department of Mathematics and Computer Science,
University of Mannheim, D-68131 Mannheim, Germany*

heiler@uni-mannheim.de

Received July 1, 2003; Revised June 29, 2004; Accepted June 29, 2004

First online version published in December, 2004

Abstract. We integrate a model for filter response statistics of natural images into a variational framework for image segmentation. Incorporated in a sound probabilistic distance measure, the model drives level sets toward meaningful segmentations of complex textures and natural scenes. Despite its enhanced descriptive power, our approach preserves the efficiency of level set based segmentation since each region comprises two model parameters only. Analyzing thousands of natural images we select suitable filter banks, validate the statistical basis of our model, and demonstrate that it outperforms variational segmentation methods using second-order statistics.

Keywords: image segmentation, natural image statistics, level sets

1. Introduction

1.1. Motivation

Statistical models play an increasingly decisive role in computer vision for shape modeling, segmentation, tracking and appearance-based recognition (GMBV, 2002). In the context of Bayesian inference, the nature of a statistical model defines the class of the optimization problem to be solved (Mumford, 1994). From an optimization point of view, there is a trade-off between the descriptive power of statistical models and the difficulty of the associated Bayesian inference step.

Recently, the statistics of filter outputs turned out to provide powerful and general models for image statistics and texture (Zhu and Mumford, 1997; Zhu et al., 1998; Puzicha et al., 1999; Portilla and Simoncelli, 2000; Wu et al., 2000). Unfortunately, incorporating such models into a variational approach to image seg-

mentation results in computationally intractable optimization problems which require time-consuming stochastic sampling methods to compute a minimizer (Zhu and Mumford, 1997; Zhu et al., 1998; Puzicha et al., 1999).

From the viewpoint of optimization, on the other hand, a range of variational approaches to image segmentation exist (Chan and Vese, 2001; Jehan-Besson et al., 2003) which can be regarded as efficient and computationally convenient approximations of the Mumford-Shah functional (Mumford and Shah, 1989). However, the class of problems that these models have been applied to is limited so far to cartoon-like piecewise smooth images and second-order statistics of multiple filter channel responses (Zhu and Yuille, 1996; Jehan-Besson et al., 2003).

This motivates us to use more sophisticated statistical models that describe a larger class of natural images without compromising the efficiency of level set based segmentation.

1.2. Contribution and Related Work

In this article, we study the use of a recent model of natural image statistics in an efficient level set based variational framework for image segmentation.

The statistical model that we describe in Section 2 has been used for wavelet-based image coding (Mallat, 1998; Buccigrossi and Simoncelli, 1999) and was empirically verified for a large database of natural images (Reininger and Gibson, 1983; Huang and Mumford, 1999).

Our variational approach to image segmentation follows Zhu and Yuille (1996) and Chan and Vese (2001). We considerably enhance the applicability of these methods by incorporating the above-mentioned statistical model (Section 2). Despite its greater descriptive power, our model is simple enough to be estimated *locally*: By measuring an appropriate distance to non-locally estimated models we compute forces that drive level sets to meaningful segmentations of complex scenes. In this sense, our work is similar to Paragios and Deriche (2002) who, in the context of *supervised* texture segmentation, successfully enhance the geodesic active contour model (Kichenassamy et al., 1995; Caselles et al., 1997) with a Gaussian mixture model of filter response statistics. However, our statistical model is more compact and targeted toward natural scenes. This allows natural images and unknown textures to be treated in a completely unsupervised way. We give a rigorous derivation of the equations driving the motion of region boundaries toward a segmentation in Section 3.

In Section 4 we put our approach in a broader context and evaluate it on images from publicly available databases. We compare it with second order statistics as in (Zhu and Yuille, 1996; Jehan-Besson et al., 2003)

using the same variational framework (Chan and Vese, 2001). In Section 5 we conclude and indicate further work.

Finally, we wish to point out that our results should not solely be judged from the specific viewpoint of texture segmentation. This would require to consider a wide range of possible dissimilarity measures (Puzicha et al., 1999; Randen and Husøy, 1999) many of which cannot be easily incorporated into a level set based segmentation framework. Rather, we focus in this paper on a compact parametric model related to natural image statistics for super- and unsupervised level set segmentation of scenes where texture plays an important but not an exclusive role.

2. Natural Image Statistics

We capture statistics of natural images using generalized Laplacians fitted to marginal histograms of linear filter responses. The Kullback-Leibler (KL) distance between the Laplacians then serves as a distance measure on the images (cf. Fig. 1). The following section describes the statistical model in detail.

2.1. Feature Extraction

The basis of our approach is the statistical model

$$p(z) = \frac{\alpha}{2s\Gamma(1/\alpha)} \exp(-|z/s|^\alpha) \quad (1)$$

for the filter response z of a linear filter applied to natural images. It was pointed out in the literature that for a large class of images the generalized Laplacian (1) describes the response statistics of various linear filters surprisingly well. This model was empirically verified

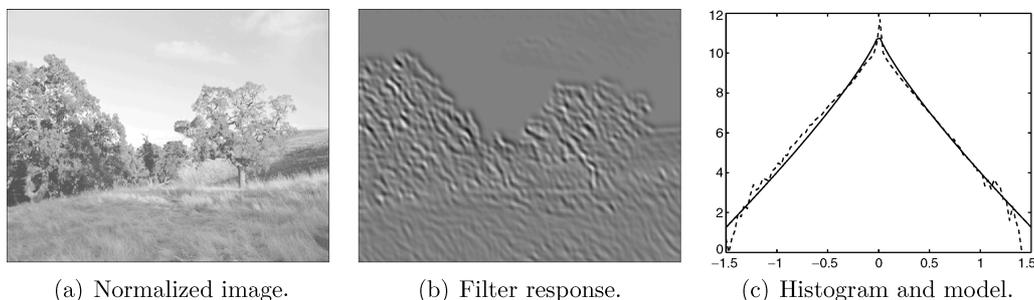


Figure 1. Overview. An image (a) is filtered by a linear filter. The marginal histogram (dotted line) of the resulting filter response image (b) is computed and a generalized Laplacian (solid line) is fitted to the histogram (c). The parameters (α, s) of the generalized Laplacian serve as image descriptors.

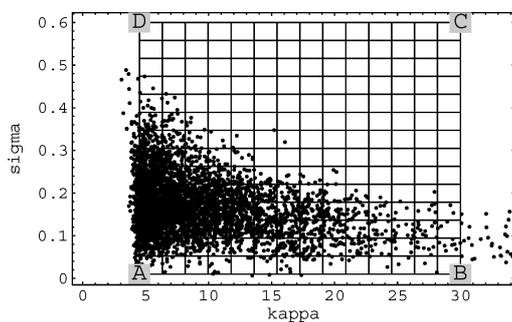
for samples of natural images (Reininger and Gibson, 1983; Huang and Mumford, 1999) and has been applied to compactly code wavelet coefficients of images (Mallat, 1998; Joshi et al., 1995; Buccigrossi and Simoncelli, 1999; LoPresto and Ramchandran, 1997) and to Bayesian image restoration (Simoncelli and Adelson, 1996). In the present paper we apply this model locally and globally within a variational framework to the segmentation of natural images.

Various linear transformations of images have been used in conjunction with the model: The discrete cosine transform (Reininger and Gibson, 1983), steerable pyramids (Freeman and Adelson, 1991; Simoncelli and Freeman, 1995; Simoncelli and Adelson, 1996), and various orthogonal wavelets (Mallat, 1998; Huang and Mumford, 1999). In this work we examine steerable pyramid filters and quadrature mirror filters as well as the well-known Haar wavelet and Daubechies wavelet of order 3. In the following, these filters are abbreviated by spn , $qmf n$, $haar$, and $daub3$, where n is an integer encoding the number of filter orientations. Experimental results to be discussed below reveal how the choice of the filter bank and the metric affect the performance of the statistical model.

2.2. Density Estimation

The generalized Laplacian model (1) has two parameters, s and α , which are related to variance σ^2 and kurtosis κ of the filter response by

$$\sigma^2 = \frac{s^2 \Gamma(3/\alpha)}{\Gamma(1/\alpha)} \quad \kappa = \frac{\Gamma(1/\alpha) \Gamma(5/\alpha)}{\Gamma^2(3/\alpha)}. \quad (2)$$



(a) Filter response parameters.

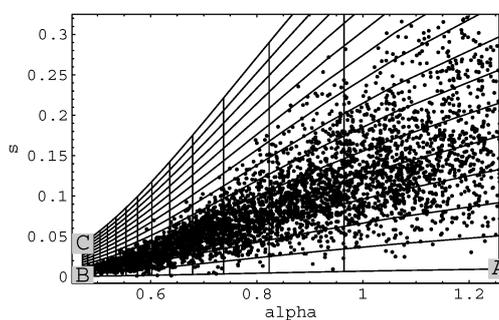
Figure 2 illustrates the nonlinear mapping from the measured statistics (σ, κ) to the model parameters (s, α) in (1). When $\kappa > 9/5$ we can solve the right equation numerically for α and determine s via the left equation. Mathematically, we cannot model distributions with $\kappa \leq 9/5$ as for $\alpha \rightarrow \infty$ the generalized Laplacian approaches the uniform distribution centered at 0, the kurtosis of which equals $9/5$. This is not a severe restriction, however: In the experimental section (Section 4 and Fig. 4) we show that such statistics are very rare in natural images.

2.3. MDL-Criterion for Segmentation

Our goal is to partition the image domain Ω into two, maybe multiply-connected, regions Ω_{in} and Ω_{out} separated by a contour \mathcal{C} such that the local image statistics are “close” to the global statistics within Ω_{in} or Ω_{out} , respectively. More precisely, if p_x denotes the statistics of a small window W_x centered at image location x , and if p_{in} and p_{out} denote the statistics of the interior and exterior regions Ω_{in} and Ω_{out} , respectively, then we want to minimize

$$E_{\text{mdl}}(\Omega_{\text{in}}, \Omega_{\text{out}}) = \int_{\mathcal{C}} ds + \int_{\Omega_{\text{in}}} D(p_x \parallel p_{\text{in}}) dx + \int_{\Omega_{\text{out}}} D(p_x \parallel p_{\text{out}}) dx. \quad (3)$$

Here $D(p \parallel q) = -\int p(z) \log(p(z)/q(z)) dz$ is the Kullback-Leibler (KL) distance between densities p and q . Note that (3) fits into Zhu and Yuille’s region competition framework (Zhu and Yuille, 1996) when



(b) Density parameters.

Figure 2. The role of nonlinear parameter mapping. Standard deviation σ and kurtosis κ of the filter responses are mapped nonlinearly according to (2) to density parameters α and s . Note, that the left part of (a), where most points are located, is spread after mapping. Conversely, the area on the right of (a), where relatively few points are situated, is compressed. The points depicted are 4167 measurements collected from the van Hateren database using a linear derivative filter. The labeled grid visualizes the nonlinearity of the transformation.

$P_{\text{in/out}} \propto \exp(-D(p_x \parallel p_{\text{in/out}}))$ are the probabilities for the region models and the local image features I_x are given by the distributions p_x .

The motivation for energy (3) is that it can be linked to the length of a hypothetical image code (Wallace and Boulton, 1968; Rissanen, 1978) based on two generalized Laplacians p_{in} and p_{out} : We encode the filter response image using either p_{in} or p_{out} as models. Assuming densities are truly Laplacian, encoding a pixel x with model p_x estimated from W_x has average length $H(p_x)$, H denoting Shannon entropy. Encoding it using the model for one of the regions Ω_{in} or Ω_{out} instead requires a code of average length $H(p_x) + D(p_x \parallel p_{\text{in/out}})$. KL-distance is nonnegative, therefore (3) describes the additional coding effort we face when encoding x with the models for one of the regions. The first integral in (3) measures the length of the separating contour \mathcal{C} , ensuring that the membership relation, that is, whether a specific point x belongs to Ω_{in} or to Ω_{out} , will be inexpensive to encode (Leclerc, 1989).

The KL-distance between two generalized Laplacians p and q with parameters (s_p, α_p) and (s_q, α_q) can be computed conveniently: First, evaluate (2) for sample estimates on the left hand sides, then insert the resulting values for the parameters $(s_p, \alpha_p, s_q, \alpha_q)$ into the following expression:

$$D(p \parallel q) = \frac{\left(\frac{s_p}{s_q}\right)^{\alpha_q} \Gamma\left(\frac{1+\alpha_q}{\alpha_p}\right)}{\Gamma\left(\frac{1}{\alpha_p}\right)} + \log\left(\frac{s_q \Gamma\left(1 + \frac{1}{\alpha_q}\right)}{s_p \Gamma\left(1 + \frac{1}{\alpha_p}\right)}\right) - \frac{1}{\alpha_p}. \quad (4)$$

Note, that the hypothetical image code described above is only optimal if adjacent pixels in the filter response are statistically *independent*. Spatial correlations of filter responses at neighboring locations are not exploited. For an efficient real-world coding scheme this would be mandatory.

2.4. Combining Filter Responses

Given the statistics for a set of filter responses, how do we combine information gathered at different scales and orientations? In this work, we strive for a *generic* measure not optimized for any particular set of textures or filters, so feature selection schemes are not directly applicable.

We propose, as a first approximation, to treat the statistics of individual filter responses as statistically

independent. Under this assumption the individual KL-distances simply add up so that we can minimize the average distance collected over all linear filters i :

$$E_{\text{mdl}}(\Omega_{\text{in}}, \Omega_{\text{out}}) = \int_{\mathcal{C}} ds + \sum_i \left[\int_{\Omega_{\text{in}}} D(p_{x,i} \parallel p_{\text{in},i}) dx + \int_{\Omega_{\text{out}}} D(p_{x,i} \parallel p_{\text{out},i}) dx \right] \quad (5)$$

Here $p_{\text{in/out},i}$ denotes the probability density function modeling the response of filter i in region $\Omega_{\text{in/out}}$ and $p_{x,i}$ is the corresponding density for a window W_x centered at location x in the image plane.

It is known that in reality the independence assumption does *not* hold. For orthogonal wavelet bases normalization schemes have been proposed to remove dependencies between filter responses at different scale and orientation (Buccigrossi and Simoncelli, 1999; Wainwright et al., 2001). In this first implementation of our approach, however, we did not incorporate any such scheme. While in theory this is clearly suboptimal, our experiments (Section 4) suggest that the model is sufficiently accurate for many real-world scenes.

3. Level Set Formulation

In this section we incorporate our statistical distance measure into a level set formulation. The update equations determining the dynamics of the segmentation are rigorously derived, taking into account *all* region-dependent terms, by computing the first variation of the corresponding area integrals.

3.1. Energy Functional

We minimize energy (5) within the region-based variational framework of Chan and Vese (2001). The framework applies to energy functionals of the form

$$E(\phi) = \int_{\Omega} k^b(x) |\nabla \phi| \delta(\phi) dx + \lambda_1 \int_{\Omega} k^{\text{out}}(x, \phi) H(\phi) dx + \lambda_2 \int_{\Omega} k^{\text{in}}(x, \phi) (1 - H(\phi)) dx. \quad (6)$$

Here $\phi : \mathbb{R}^2 \rightarrow \mathbb{R}$ denotes the embedding level set function, the zero-level of which represents segmentation boundaries. $H : \mathbb{R} \rightarrow \{0, 1\}$ is the step function

and δ Dirac's delta function. $k^b(x)$, $k^{\text{in}}(x, \phi)$, and $k^{\text{out}}(x, \phi)$ represent the boundary, interior, and exterior energy contributions at a location x and for a given level set function ϕ . Finally, λ_1 and λ_2 weight the relative importance of the interior and exterior energy terms against boundary energy. In the following we usually drop the arguments ϕ and x for brevity.

Chan and Vese's original gray-value based image model (Chan and Vese, 2001) fits into this framework as a special case with

$$\begin{cases} k^b = 1 \\ k^{\text{in}} = |u_0 - c_{\text{in}}|^2 \\ k^{\text{out}} = |u_0 - c_{\text{out}}|^2, \end{cases} \quad (7)$$

whereas with

$$\begin{cases} k^b = 1 \\ k^{\text{in}} = \sum_i D(p_{x,i} \parallel p_{\text{in},i}) \\ k^{\text{out}} = \sum_i D(p_{x,i} \parallel p_{\text{out},i}) \end{cases} \quad (8)$$

energy (5) is obtained.

3.2. First Variation and Boundary Update

The variational update $\dot{\phi} = -\langle E'(\phi), \psi \rangle$, $\forall \psi$, of the level set function reads¹:

$$\begin{aligned} \frac{\partial E}{\partial \phi} &= \frac{\partial}{\partial \phi} \left[\int_{\Omega} k^b |\nabla \phi| \delta dx \right] \\ &+ \int_{\Omega} (\lambda_1 k^{\text{out}} - \lambda_2 k^{\text{in}}) \delta \psi dx \\ &+ \int_{\Omega} \left(\lambda_1 \frac{\partial k^{\text{out}}}{\partial \phi} H + \lambda_2 \frac{\partial k^{\text{in}}}{\partial \phi} (1 - H) \right) \psi dx. \end{aligned} \quad (9)$$

The third term, which is omitted in Chan and Vese (2001), originates from applying the product rule to the area integrals and thus takes into account that k^{in} and k^{out} also depend on the level set function ϕ . After some tedious calculations (Appendix A) and with the shorthands $n = \frac{\nabla \phi}{|\nabla \phi|}$ and $c = \text{div}(n)$ we arrive at

$$\begin{aligned} \frac{\partial E}{\partial \phi} &= \int_{\Omega} (-\nabla k^b n - k^b c + \lambda_1 k^{\text{out}} - \lambda_2 k^{\text{in}}) \psi ds \\ &+ \int_{\Omega} \left(\lambda_1 \frac{\partial k^{\text{out}}}{\partial \phi} H + \lambda_2 \frac{\partial k^{\text{in}}}{\partial \phi} (1 - H) \right) \psi dx. \end{aligned} \quad (10)$$

We point out that this formula was recently derived in a different way in Jehan-Besson et al. (2003) based on the calculus of shape optimal design (Sokolowski and Zolesio, 1991) which, in turn, relies on previous mathematical work like, e.g., (Simon, 1980).

3.3. Derivation of the Model's Area Term

Let us examine more closely the area integral in (10). As mentioned above, Eq. (8), we model the local coding cost w.r.t. the interior region as

$$k^{\text{in}} = \sum_i D(p_{x,i} \parallel p_{\text{in},i}). \quad (11)$$

Recall that the probability density functions are given as generalized Laplacians with two parameters $s = s(\alpha, \sigma^2)$ and $\alpha = \alpha(\kappa)$ which depend themselves on kurtosis κ and variance σ^2 measured both locally in W_x and globally in Ω_{in} . Therefore, we may write more precisely

$$\begin{aligned} k^{\text{in}} &= \sum_i D(p(\alpha(\kappa_{x,i}), s(\alpha(\kappa_{x,i}), \sigma_{x,i}^2)) \\ &\parallel p(\alpha(\kappa_{\text{in},i}), s(\alpha(\kappa_{\text{in},i}), \sigma_{\text{in},i}^2))). \end{aligned} \quad (12)$$

Here $\kappa_{\text{in},i}$ and $\sigma_{\text{in},i}^2$ depend on the area Ω_{in} and thus vary with the level set function ϕ . Let us drop the index i in the following discussion, thus focusing on a single filter response only.

With a slight abuse of notation, the derivative then reads

$$\frac{\partial k^{\text{in}}}{\partial \phi} = \frac{\partial D}{\partial \kappa_{\text{in}}} \frac{\partial \kappa_{\text{in}}}{\partial \phi} + \frac{\partial D}{\partial \sigma_{\text{in}}^2} \frac{\partial \sigma_{\text{in}}^2}{\partial \phi}, \quad (13)$$

where the computation of the partial derivatives $\partial D / \partial \kappa_{\text{in}}$ and $\partial D / \partial \sigma_{\text{in}}^2$ is long but nevertheless elementary: Starting from the analytical formulation of the KL-distance (4) and inserting the relations (2) solved for α and s it is easily obtained.

The statistics depending on the area form a hierarchy of region-dependent terms:

$$\begin{aligned} \kappa_{\text{in}} &= \int_{\Omega_{\text{in}}} \frac{(x - \mu_{\text{in}})^4}{|\Omega_{\text{in}}| \sigma_{\text{in}}^4} dx & \sigma_{\text{in}}^2 &= \int_{\Omega_{\text{in}}} \frac{(x - \mu_{\text{in}})^2}{|\Omega_{\text{in}}|} dx \\ \mu_{\text{in}} &= \int_{\Omega_{\text{in}}} \frac{x}{|\Omega_{\text{in}}|} dx & |\Omega_{\text{in}}| &= \int_{\Omega_{\text{in}}} dx. \end{aligned} \quad (14)$$

In the level set formulation (6) we replace the integrals over Ω_{in} by integrals over Ω weighted by the step

function H . Now, taking the derivative w.r.t. ϕ yields (cf. Appendix B)

$$\begin{aligned} \frac{\partial \sigma_{\text{in}}^2}{\partial \phi} &= - \int_{\Omega} \frac{(x - \mu_{\text{in}})^2}{|\Omega_{\text{in}}|^2} H dx \int_{\Omega} \delta \psi dx \\ &\quad + \int_{\Omega} \frac{(x - \mu_{\text{in}})^2}{|\Omega_{\text{in}}|} \delta \psi dx \end{aligned} \quad (15)$$

and

$$\begin{aligned} \frac{\partial \kappa_{\text{in}}}{\partial \phi} &= \int_{\Omega} \frac{-4(x - \mu_{\text{in}})^3}{|\Omega_{\text{in}}| \sigma_{\text{in}}^4} H dx \left[\int_{\Omega} \frac{-x}{|\Omega_{\text{in}}|^2} H dx \right. \\ &\quad \times \left. \int_{\Omega} \delta \psi dx + \int_{\Omega} \frac{x}{|\Omega_{\text{in}}|} \delta \psi dx \right] \\ &\quad + 2\sigma_{\text{in}}^2 \int_{\Omega} \frac{(x - \mu_{\text{in}})^4}{|\Omega_{\text{in}}| \sigma_{\text{in}}^8} H dx \left[\int_{\Omega} \frac{(x - \mu_{\text{in}})^2}{|\Omega_{\text{in}}|^2} H dx \right. \\ &\quad \times \left. \int_{\Omega} \delta \psi dx - \int_{\Omega} \frac{(x - \mu_{\text{in}})^2}{|\Omega_{\text{in}}|} \delta \psi dx \right] \\ &\quad - \int_{\Omega} \frac{(x - \mu_{\text{in}})^4}{|\Omega_{\text{in}}|^2 \sigma_{\text{in}}^4} H dx \int_{\Omega} \delta \psi dx \\ &\quad + \int_{\Omega} \frac{(x - \mu_{\text{in}})^4}{|\Omega_{\text{in}}| \sigma_{\text{in}}^4} \delta \psi dx. \end{aligned} \quad (16)$$

With (13) these terms form the area derivatives in (10).

4. Experiments and Discussion

Now we describe extensive computational studies of the performance of our model. We validate the use of generalized Laplacian densities for steerable pyramid filter response statistics of natural images, perform experiments in texture retrieval and synthesis to understand what image features are captured by our model, and show sample segmentations on natural and artificial images. We compare our model with a standard second-order variational model for image segmentation and demonstrate that it performs well.

4.1. Filter Selection and Model Validation

Before focusing on segmentation (Section 4.3) we conducted experiments to select a suitable filter bank and to verify that the restriction on the kurtosis of the filter response to be greater than 9/5 is met in practice (Section 2.2). Following Huang and Mumford (1999) we used the van Hateren database of natural images (van Hateren and vander Schaaf, 1998) for evaluation and removed multiplicative constants from the images by first log-transforming them and then subtracting their log-means.

Table 1 summarizes our results: We display the median of the KL-distance between the filter response histograms (20 bins) and a generalized Laplacian with identical variance and kurtosis. For comparison, we also report the histograms' average entropy and the median of the quotient of these values. The results show that almost all information in the histograms is captured by the parametric model. Importantly, the same holds for densities estimated *locally* from moderately small image patches (Fig. 3). In the following, we perform all experiments using the steerable pyramid bank sp3 with four oriented sub-band filters and over three scales.

In Fig. 4(a) we show the log-histograms of the kurtosis κ for each individual filter determined for all 4167 images of the database. Two things are remarkable: First, the distribution of κ follows closely a shifted exponential distribution. Second, the minimal values of κ encountered are well above the critical value of 9/5. Thus, *distributions that violate the kurtosis-constraint of our model do not occur in natural images.*

Clearly, during segmentation we also work with small *parts* of images for which small values for kurtosis are observed. Especially very homogeneous image regions like sky or plain street occasionally lead to untypical filter response histograms (Fig. 3). To see how frequently this happens in reality, we randomly sampled over 700,000 image patches of size

Table 1. *Model fit.* Medians of KL-distances between histograms and parametric model (1) measured over 4167 pictures from the van Hateren database (van Hateren and vander Schaaf, 1998) for different sets of filters. For comparison, median entropies of the filter responses are also reported: Only a small fraction of the information present in the histograms is ignored (last row).

	sp0	sp1	sp3	sp5	qmf9	qmf12	qmf16	haar	daub3
KL-dist	0.018	0.011	0.012	0.014	0.016	0.016	0.017	0.016	0.017
Entropy	2.274	2.070	1.996	2.004	1.933	1.932	1.938	1.994	1.966
KL/entropy	0.008	0.005	0.006	0.007	0.009	0.009	0.009	0.008	0.009

patch size	sp0	sp1	sp3	sp5
10×10	0.163	0.144	0.144	0.162
20×20	0.053	0.045	0.046	0.054
30×30	0.032	0.025	0.025	0.032
40×40	0.023	0.017	0.017	0.023
50×50	0.019	0.013	0.014	0.018
100×100	0.010	0.007	0.007	0.009

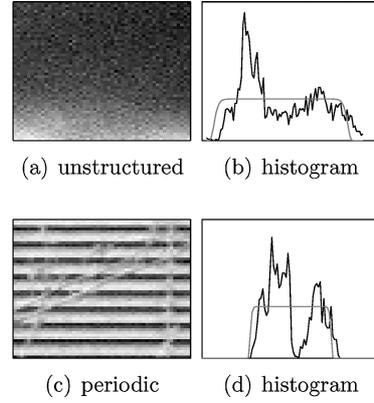


Figure 3. *Influence of window size.* How does the size of W_x influence the accuracy of the parametric model p_x in modeling the filter histogram? In the table the quotient KL-distance/entropy (Table 1) is depicted for 50.000 image patches randomly selected from the van Hateren database. The bigger the images the more accurate is the fit between histogram and Laplacian. Two 50×50 image patches representative for very bad fits are depicted on the right: Unstructured areas like sky or street (a) and areas with regular, periodic structure (c) are most problematic.

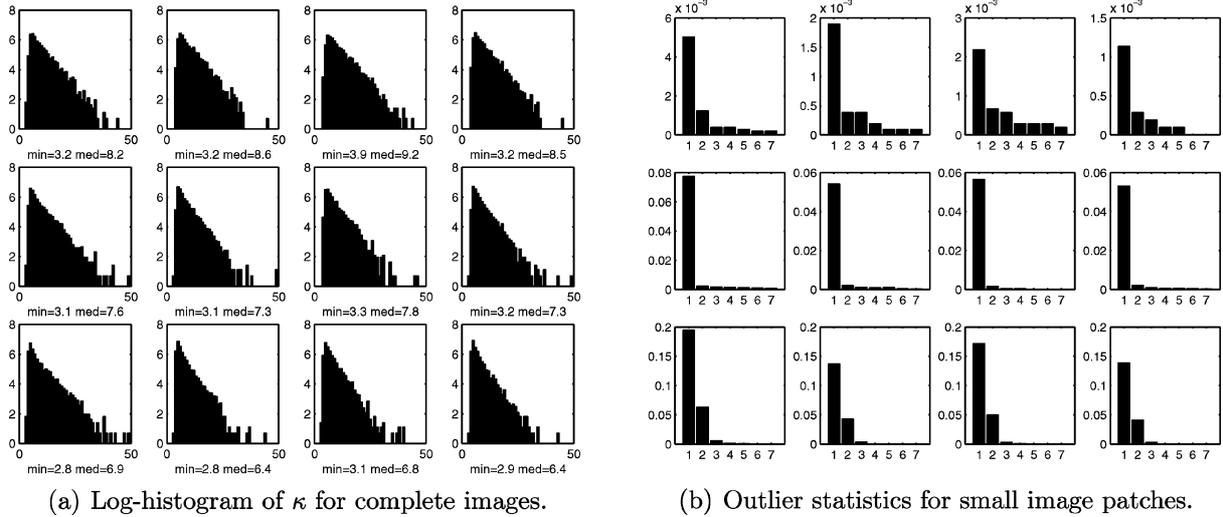


Figure 4. *Check for pathological statistics.* (a) shows the log-histogram of kurtosis κ measured over 4167 images from the van Hateren database (van Hateren and vander Schaaf, 1998) for a steerable pyramid filter bank with three scales (rows) and four orientations (columns). Minimal and median values for κ are listed in the individual image captions. The histograms are *very regular*, and for each filter κ is well above $9/5$, thus *no pathological cases* are present in the database. (b) shows the relative frequency of outliers with $\kappa < 9/5$, measured over approximately 700,000 randomly sampled image patches of size $10^2, 20^2, 30^2, 40^2, 50^2, 75^2$, and 100^2 , labeled 1 (size 10^2) to 7 (size 100^2). *Outliers are frequent with patch sizes smaller than 30×30 only.*

$10^2, 20^2, 30^2, 40^2, 50^2, 75^2$, and 100^2 pixels from the van Hateren database. For each patch size we counted how often the constraint $\kappa > 9/5$ was violated. The relative frequencies are shown in Fig. 4(b): Only for the two smallest patch sizes, corresponding to the bars labeled “1” and “2”, violations were found regularly.

For patch sizes of size 30×30 or larger violations were very rare. In the segmentation experiments reported below we treated these cases as outliers, replacing κ with a default value slightly larger than $9/5$. We found that this did not lead to a noticeable deterioration of segmentation quality.

Table 2. Assessment of KL-distance. Misclassification rate in percent on 512 randomly sampled texture image patches from the Brodatz database (Brodatz, 1966). Different filters, features, and strategies to combine individual filter responses are compared (see text for details). The minimal error rate for each filter is marked in boldface. KL-distance with L_1 metric performs best in most cases.

	L_∞ (σ, κ)	L_1 (σ, κ)	med (σ, κ)	L_∞ (α, s)	L_1 (α, s)	med (α, s)	L_∞ (KL)	L_1 (KL)	med (KL)
sp0	21.91	18.19	23.48	26.61	26.81	59.29	14.67	13.89	20.15
sp1	12.52	6.26	11.74	12.91	9.98	21.72	6.45	5.08	6.65
sp3	12.13	5.87	7.82	12.91	9.39	9.19	5.67	3.52	4.69
sp5	15.06	6.26	6.45	12.72	6.26	8.80	7.43	5.87	4.69
qmf9	16.43	4.10	6.65	16.43	8.02	16.24	3.91	3.32	4.30
qmf12	15.45	3.71	6.84	15.26	7.63	17.41	3.91	3.13	4.50
qmf16	12.52	3.71	6.26	12.91	7.43	16.04	3.13	2.34	3.32
haar	15.26	5.47	7.63	17.41	8.41	23.87	6.65	4.10	6.84
daub3	16.04	4.50	8.61	15.65	9.00	17.61	5.67	3.32	4.30

To assess our probabilistic distance measure we ran an experiment in texture retrieval on images from the Brodatz database: We extracted 16 image patches of size 100×100 pixels non-overlappingly from 32 Brodatz images. Then we took each of the $32 \cdot 16 = 512$ image patches as a query and selected from the remaining patches the one most similar to the query w.r.t. a number of distance measures. We examined vectors of filter response statistics collected over different scales and orientations and for different sets of linear filters. Additionally, we examined KL-distance. The distances for individual scales and orientations were computed independently and then combined using the max, mean, and median operator, corresponding to L_∞ , L_1 and med in the table. A retrieval was considered correct if the patch most similar to the query originated from the same Brodatz image. Otherwise it was considered wrong.

Table 2 summarizes the results which show that the mean of the KL-distance performs best for most filters. This indicates that (12) is a useful distance measure on images with texture.

4.2. Texture Synthesis Experiment

To get an intuition for which image features are captured by the generalized Laplacians we synthesized texture images using our model. For computational efficiency we did not resort to the Gibbs sampler but modified the fast pyramid-based algorithm of Heeger and Bergen (1995) instead. This greedy algorithm en-

forces filter histogram similarity between a target image and a source image initialized to random noise over different scale and orientation bands of a steerable pyramid. In contrast to Heeger and Bergen we did not fit the complete filter histograms but only their generalized Laplacians. A similar approach was taken in Srivastava et al. (2002) where Bessel K forms and the Gibbs sampler were used to synthesize texture images from a larger number of linear filter responses.

Figure 5 shows some results: While our—from the viewpoint of image synthesis overly simple—method does *not* produce realistically looking textures, it appears subjectively that some discriminative information essential for image segmentation such as predominant orientation is retained.

4.3. Supervised and Unsupervised Segmentation with Level Sets

To learn how our segmentation method performs on a set of standard images, we composed randomly selected textures from the Brodatz database and arranged them in a texture collage with a cross-shaped inlay of one texture in another (Fig. 6). We segmented 100 texture collages using (9) without area derivatives and with fixed default parameters. While in our experience the window size is an important parameter and should be chosen not too small, the choice of $\lambda_{1,2}$ is not critical. In the experiments we chose $\lambda_1 = \lambda_2 = 1$ and window size $|W| = 80 \times 80$ pixels. The texture collages were of size 512×512 . For comparison, we implemented an

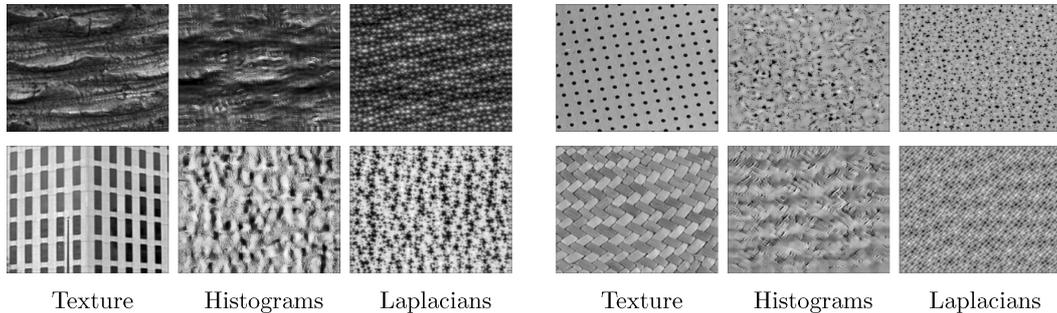


Figure 5. *Texture synthesis.* Textures from the VisTex database are reproduced using the histogram-based algorithm of Heeger and Bergen (1995) and a simplified version which uses only image features captured by our model. All images were synthesized using identical filters and the same number of iterations. This illustrates how our model captures some structure of the image.

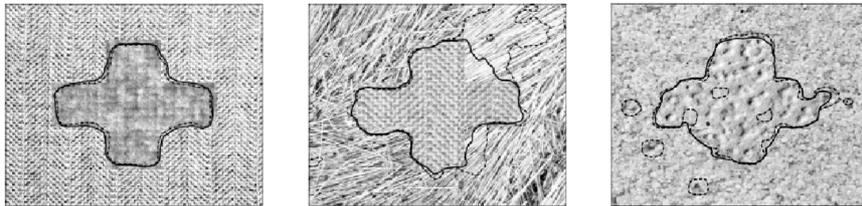


Figure 6. *Sample segmentations.* Brodatz texture collages segmented with KL-distance (solid line) and second order statistics (dotted line) with default parameters set. The KL-distance captures the cross-shaped inlay better than second order statistics. Here, we show some examples out of a large number of segmentation experiments, the statistics of which is given in Table 3. The image on the right is not successfully segmented by either method.

image model based on second order statistics (cf. (Zhu and Yuille, 1996) Eq. (20)):

$$\begin{cases} k^b = 1 \\ k^{\text{in}} = \sum_i \log(\sigma_{\text{in},i}^2) + \frac{(\mu_{x,i} - \mu_{\text{in},i})^2}{\sigma_{\text{in},i}^2} + \sigma_{x,i}^2 / \sigma_{\text{in},i}^2 \\ k^{\text{out}} = \sum_i \log(\sigma_{\text{out},i}^2) + \frac{(\mu_{x,i} - \mu_{\text{out},i})^2}{\sigma_{\text{out},i}^2} + \sigma_{x,i}^2 / \sigma_{\text{out},i}^2. \end{cases} \quad (17)$$

This model should work well for images where the mean is the most important region descriptor (Fig. 9(h)). Our Brodatz-collages are of such type: The individual texture images usually are quite homogeneous, so filter response differences are likely to origin from texture boundaries.

We ran both image models for 100 iteration steps, i.e., well after we expected convergence, on each texture collage, *using the same variational framework* (Section 3.1) for energy minimization. For increased speed we computed the image statistics on a subsampled image and interpolated the result on the whole image. This makes the region boundaries look slightly

smoother than one would expect. As both models are affected in exactly the same way this should not affect the model comparison. We finally determined the percentage of correctly segmented pixels. We found (see Table 3) that the average performance (median) as well as the performance on difficult images (25% quartile) of our model was significantly better than the performance of model (17).

We then evaluated the importance of the area derivatives, which are often omitted in variational segmentation implementations. We took the first 100 images

Table 3. *Comparison of segmentation quality.* The percentage of correctly segmented pixels on a set of 100 randomly generated Brodatz texture collages is reported for our model and for a reference model based on second order statistics. The median and both quartiles are shown. Our model clearly outperforms the reference model on average and shows much better performance on difficult images.

	Reference model	Proposed model	Improvement
median	0.65	0.81	25%
q-25	0.47	0.69	47%
q-75	0.81	0.84	4%

from the van Hateren database and computed the area derivative term from (9)

$$\lambda_1 \frac{\partial k^{\text{out}}}{\partial \phi} H + \lambda_2 \frac{\partial k^{\text{in}}}{\partial \phi} (1 - H) \quad (18)$$

for an initial segmentation consisting of equally spaced squares distributed over the whole image (Fig. 9(a)). For comparison, we computed the KL-term

$$\lambda_1 k^{\text{out}} - \lambda_2 k^{\text{in}}, \quad (19)$$

and measured the influence over the whole image.

The results (Table 4) indicate that *for our choice of distance measure the area derivatives are negligible*. This validates common practice and allows for simpler implementations. Note, however, that this might not hold in general: Recently Jehan-Besson et al. (2003) reported different results for a different choice of distance measure.

Table 4. Importance of the area term. The 10% and the 90% quantiles of Eqs. (18) and (19) evaluated on 100 images from the van Hateren database are reported. The contributions of the area term are five orders of magnitude smaller than the contributions of the KL-term, indicating that for our distance measure the area derivatives are negligible.

	q-10	q-90
KL-term	-4.0	2.6
Area-term	$-3.6 \cdot 10^{-5}$	$2.1 \cdot 10^{-5}$

Figures 7 to 9 show some examples for supervised and unsupervised segmentation of natural images. In Fig. 7 we examine an image from the Berkeley database (Martin et al., 2001). The contour was initialized to equally spaced boxes. As stopping criterion we computed the improvement of the energy functional (6)

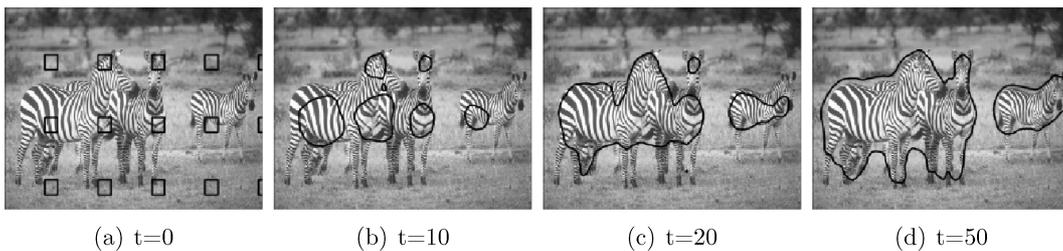


Figure 7. Unsupervised segmentation. Zebras are separated from the background. Contours were initialized to boxes, stopping was determined automatically according to $E(\phi)'$.

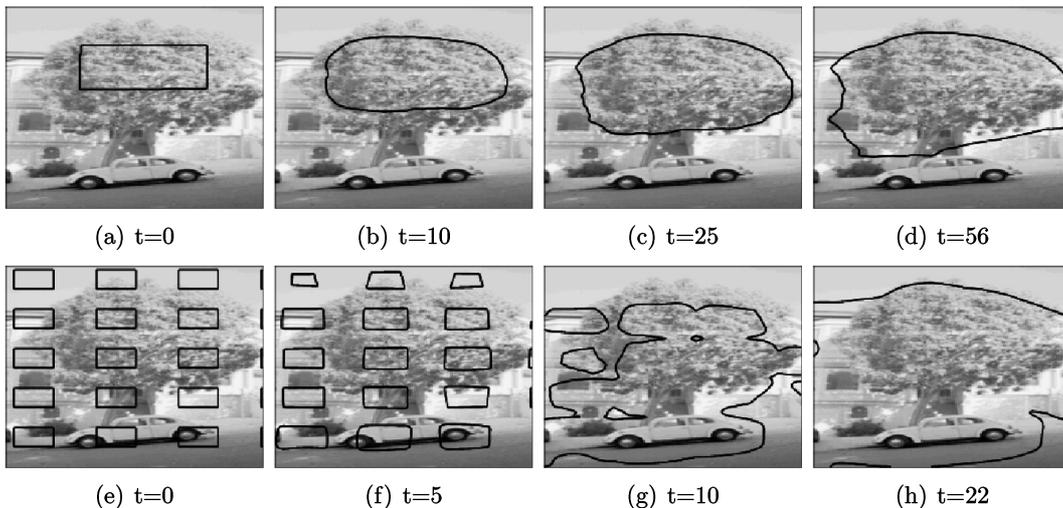


Figure 8. Supervised and unsupervised segmentation. With supervised segmentation the tree is separated from house and car. Unsupervised segmentation fails in this case: Initialization of the filter response model is too unspecific, yielding a rather uninteresting segmentation into homogeneous (sky, street) and inhomogeneous regions (car, tree, house). Note the low image contrast in the lower left part of the tree.

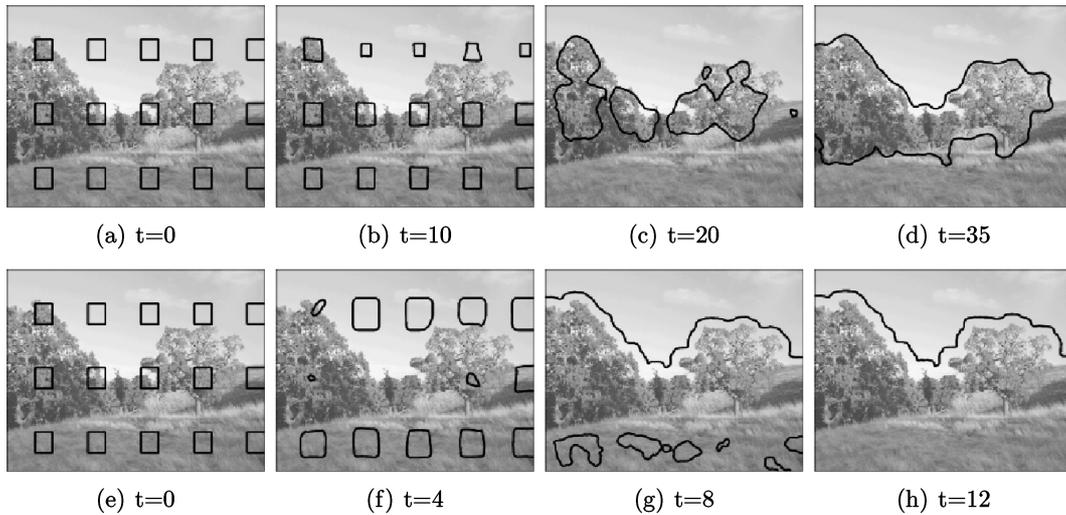


Figure 9. *Unsupervised segmentation.* Unsupervised segmentation of a natural scene from the VisTex database (Picard et al., 1995). Contours were initialized to boxes, stopping was determined automatically according to $E(\phi)'$. The contour evolution at different time steps is displayed for our model ((a)–(d)) and for second order statistics ((e)–(h)). The trees in the center of the image are the visually most dominant element which is reflected by the segmentation with our model. Second order statistics separates the bright sky from the darker rest of the image, failing to capture the visually dominant trees.

for every time step and stopped as soon as it dropped below a previously determined threshold. The same threshold was used for all experiments. The zebra pattern is captured well by our model: The contour immediately locks onto the zebra pattern and energy (6) (not shown) drops sharply until the zebras are covered.

Figure 8 shows a more difficult case: A tree standing in front of a house, casting a sharp shadow on the house. With this image, unsupervised segmentation merely separates the irregular regions from the homogeneous sky and parts of the streets (Fig. 8(e)–(h)). In contrast, if the contour is initialized in a supervised way (Fig. 8(a)) the model captures the visually dominant tree. However, in the final segmentation (Fig. 8(d)) relatively large parts of the shadowed house are captured as well.

In Fig. 9 we compare our model with second order statistics (18) on an image from the MIT VisTex database (Picard et al., 1995). The MDL criterion (3) separates the trees from the image fore- and background. This is sensible: The trees form an image region which is relatively expensive to encode while sky and grassland are comparatively homogeneous. Using one probability model for the trees and one for the rest of the image thus minimizes the expected coding length of the image. Second order statistics simply separates the bright sky from the rest of the image, yielding a less appealing segmentation.

4.4. Relation to Established Segmentation Approaches

The image model we employ can potentially be useful within alternative segmentation frameworks based on graph cuts (Shi and Malik, 2000; Keuchel et al., 2003), density clustering (Puzicha et al., 1999), or within the image parsing framework (Tu and Zhu, 2002). For graph cut methods, the KL-distances $D(p_x \parallel p_{x'})$ between generalized Laplacians p_x and $p_{x'}$ is easily translated into a similarity value $w_{xx'} = \exp(-D(p_x \parallel p_{x'})/c)$ which can then be treated, for instance, by normalized cut. In Fig. 10 we show, as a mere proof of concept, results obtained when such similarity graphs are partitioned by normalized cut. The scaling constant c was chosen as the mean KL-distance observed in the images, and the images were subsampled to reduce the size of the eigenvalue problem to approximately 1000×1000 matrix entries. The results roughly resemble those of the level set implementation and could further be improved by integrating a smoothing term in the similarity measure and by implementing a more sophisticated approximation method (Fowlkes et al., 2004) to reduce block-artifacts.

Note, however, that in the graph cut framework there are no explicit models p_{in} and p_{out} for the complete interior and exterior image regions Ω_{in} and Ω_{out} . Only similarities between *locally* estimated image models p_x

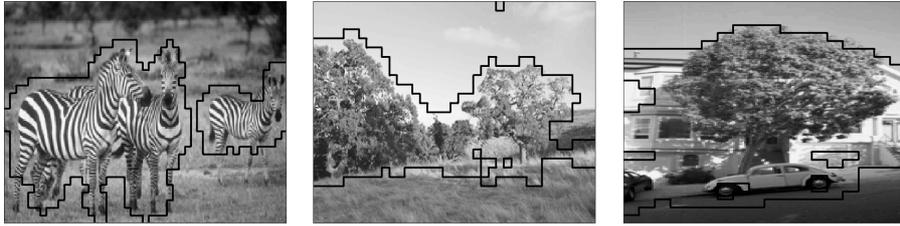


Figure 10. *KL-Segmentation with Normalized Cut*. A similarity matrix, derived from KL-distances between locally fitted generalized Laplacians (see text), was treated within the classical normalized cut framework. No effort was undertaken to enforce particularly smooth partitions. The blocky segmentation boundaries are an artifact of our particular implementation.

are used. In connection with our approach this might be a drawback as the global estimation over the larger image regions Ω_{in} and Ω_{out} is usually more reliable than the local models estimated from small image windows: During a typical PDE evolution the region models p_{in} and p_{out} are refined iteratively until they represent their corresponding image regions quite accurately. This is not possible for non-iterative optimization methods.

Conversely, our framework can benefit from employing alternative image models instead of parametric generalized Laplacians. Bessel K forms, derived from an image model based on weighted superposition of transparent objects, are suggested to represent broad image classes (Srivastava et al., 2002). Similarly, Weibull and power-law distributions were recently proposed and evaluated on thousands of natural images (Geusebroek and Smeulders, 2003). The beauty of these parametric models is that, while depending on few parameters only, they apply in rather broad contexts, and in some cases statistical goodness-of-fit tests are readily available.

When more flexibility in image modeling is needed, in particular for images with regular textures, mixture models (Belongie et al., 1998; Paragios and Deriche, 2002; Wu et al., 2003) and non-parametric models (Tang and Ma, 2001; Rousson et al., 2003) come into play. These can, in principle, model empirical densities to arbitrary precision. However, in order to avoid overfitting within unsupervised settings, care must be taken that model complexity is kept under control. Also, KL-distances can in general no longer be evaluated analytically.

Empirical densities represented by histograms of filter responses also provide greater modeling capacity (Zhu et al., 1998). They fit into framework (6) when the parametric KL-distance in (8) is replaced by the discrete KL-distance between histograms. However, this solution might not be optimal as results can be sensitive to the chosen histogram bin-size. Therefore,

more robust statistical measures, such as earth movers distance, χ^2 , Kolmogorov-Smirnov, or the Anderson-Darling statistics seem more promising (Puzicha et al., 1999; Rubner et al., 2000; Liu and Wang, 2000; Geusebroek and Smeulders, 2003).

Ideally, the user would not be required to decide for a particular image model or for the number of different image regions to expect *a priori*, but multiple models of different complexity would compete to explain the image during the course of optimization. This leads to a model selection problem which can in principle be treated within an MDL framework (Leclerc, 1989; Hansen and Yu, 2001). While extensions of the Chan and Vese framework to multiple image regions (Vese and Chan, 2002) and models (Cremers et al., 2004) have been proposed, it is unclear if they generalize to a full MDL approach with multiple image models of different modeling capacity. Currently, methods from non-convex optimization are employed to handle such problems (Leclerc, 1989; Zhu et al., 2000).

5. Conclusions and Further Work

In this paper we proposed a segmentation approach based on natural image statistics and the gradient-less level set segmentation method introduced by Chan and Vese (2001). Exploiting the fact that a simple parametric model accurately describes the statistics of a wide class of filter responses on natural images we constructed an energy functional justified by a minimum description length argument.

We ran evaluations on thousands of images checking that pathological cases not captured by our model do not occur in real world images (Fig. 4(a)), that the empirically observed histograms are accurately represented (Table 1), and that the minimum description length formulation does contribute to the descriptive power of our model (Table 2).

We conducted experiments to evaluate the performance of our segmentation method in comparison to a second order model which has been used successfully for image segmentation before (Zhu and Yuille, 1996). The results indicate that for segmentation tasks where image structure is more important than brightness contrasts our model compares favorably (Figs. 6, 8, 9, and Table 3).

Finally, we examined the importance of an area derivative term emerging during the derivation of the first variation of our energy functional. We found (Table 4) that for our functional the area derivative's contribution is not significant, thus validating the common practice of ignoring it. Omitting the area derivatives removes the requirement that the area descriptors k^{in} and k^{out} must be differentiable and greatly simplifies the implementation.

An interesting line of research for the future is to examine how image features can be captured with more involved image probability densities without overly compromising model simplicity: Our energy functional can in principle be applied with arbitrary probability densities. However, model validation issues as well as performance arguments make image models desirable which are easy to train even on small image patches.

Appendix A: First Variation of the Energy Functional

The exposition follows (Chan and Vese, 2001), suitably generalized and adapted to our approach. Starting from (9)

$$\begin{aligned} \frac{\partial E}{\partial \phi} &= \frac{\partial}{\partial \phi} \left[\int_{\Omega} k^b |\nabla \phi| \delta dx \right] + \int_{\Omega} \delta (\lambda_1 k^{\text{out}} - \lambda_2 k^{\text{in}}) \psi dx \\ &\quad + \int_{\Omega} \left(\lambda_1 \frac{\partial k^{\text{out}}}{\partial \phi} H + \lambda_2 \frac{\partial k^{\text{in}}}{\partial \phi} (1 - H) \right) \psi dx \end{aligned}$$

we take a closer look at the first term which equals

$$\int_{\Omega} k^b \left[\delta' |\nabla \phi| \psi + \delta \frac{\nabla \phi}{|\nabla \phi|} \nabla \psi \right] dx \quad (20)$$

by product rule. With Green's first theorem the second part becomes

$$\begin{aligned} &\int_{\Omega} k^b \delta \frac{\nabla \phi}{|\nabla \phi|} \nabla \psi dx \\ &= - \int_{\Omega} \nabla \left(k^b \delta \frac{\nabla \phi}{|\nabla \phi|} \right) \psi dx + \int_{\partial \Omega} \frac{k^b \delta \partial \phi}{|\nabla \phi| \partial n} \psi ds \end{aligned}$$

$$\begin{aligned} &= - \int_{\Omega} \left[\nabla k^b \delta \frac{\nabla \phi}{|\nabla \phi|} + k^b \nabla \delta \frac{\nabla \phi}{|\nabla \phi|} \right. \\ &\quad \left. + k^b \delta \nabla \left(\frac{\nabla \phi}{|\nabla \phi|} \right) \right] \psi dx + \int_{\partial \Omega} \frac{k^b \delta \partial \phi}{|\nabla \phi| \partial n} \psi ds \end{aligned} \quad (21)$$

which in connection with $\nabla \delta \frac{\nabla \phi}{|\nabla \phi|} = \delta' |\nabla \phi|$ and (20) yields

$$\begin{aligned} &\frac{\partial}{\partial \phi} \left[\int_{\Omega} k^b |\nabla \phi| \delta dx \right] \\ &= - \int_{\Omega} \left[\nabla k^b \delta \frac{\nabla \phi}{|\nabla \phi|} + k^b \delta \nabla \left(\frac{\nabla \phi}{|\nabla \phi|} \right) \right] \psi dx \\ &\quad + \int_{\partial \Omega} \frac{k^b \delta \partial \phi}{|\nabla \phi| \partial n} \psi ds. \end{aligned} \quad (22)$$

Note that we can replace each area integral containing the Dirac impulse into an integral over the region boundary $\mathcal{C} = \{x : \phi(x) = 0\}$:

$$\int_{\Omega} f(x, \phi) \delta(\phi) dx = \int_{\mathcal{C}} f(x, 0) ds. \quad (23)$$

Hence we can write

$$\begin{aligned} \frac{\partial E}{\partial \phi} &= \int_{\mathcal{C} \cap \partial \Omega} \frac{k^b \partial \phi}{|\nabla \phi| \partial n} \psi ds + \int_{\mathcal{C}} \left[- \nabla k^b \frac{\nabla \phi}{|\nabla \phi|} \right. \\ &\quad \left. - k^b \text{div} \left(\frac{\nabla \phi}{|\nabla \phi|} \right) + \lambda_1 k^{\text{out}} - \lambda_2 k^{\text{in}} \right] \psi ds \\ &\quad + \int_{\Omega} \left(\lambda_1 \frac{\partial k^{\text{out}}}{\partial \phi} H + \lambda_2 \frac{\partial k^{\text{in}}}{\partial \phi} (1 - H) \right) \psi dx. \end{aligned} \quad (24)$$

Assuming $\mathcal{C} \cap \partial \Omega = \emptyset$, this leads to (10).

Appendix B: Derivation of the Area Terms

We start from the relations (14) and replace the integrals over Ω_{in} by integrals over Ω weighted by the step function H . Taking the derivative w.r.t. ϕ yields

$$\begin{aligned} \frac{\partial \sigma_{\text{in}}^2}{\partial \phi} &= \frac{\partial}{\partial \phi} \int_{\Omega} \frac{(x - \mu_{\text{in}})^2}{|\Omega_{\text{in}}|} H dx \\ &= \int_{\Omega} \frac{\partial}{\partial \phi} \left[\frac{(x - \mu_{\text{in}})^2}{|\Omega_{\text{in}}|} \right] H + \frac{(x - \mu_{\text{in}})^2}{|\Omega_{\text{in}}|} \delta \psi dx \end{aligned} \quad (25)$$

and

$$\frac{\partial}{\partial \phi} \left[\frac{(x - \mu_{\text{in}})^2}{|\Omega_{\text{in}}|} \right] = 2 \frac{\mu_{\text{in}} - x}{|\Omega_{\text{in}}|} \frac{\partial \mu_{\text{in}}}{\partial \phi} - \frac{(x - \mu_{\text{in}})^2}{|\Omega_{\text{in}}|^2} \frac{\partial |\Omega_{\text{in}}|}{\partial \phi} \quad (26)$$

and finally

$$\frac{\partial |\Omega_{\text{in}}|}{\partial \phi} = \int_{\Omega} \frac{\partial H}{\partial \phi} dx = \int_{\Omega} \delta \psi dx. \quad (27)$$

Collecting these terms and using $\int_{\Omega} (\mu_{\text{in}} - x) H dx = 0$ yields (15).

The derivation of $\partial \kappa / \partial \phi$ proceeds in the very same manner:

$$\begin{aligned} \frac{\partial \kappa_{\text{in}}}{\partial \phi} &= \frac{\partial}{\partial \phi} \int_{\Omega} \frac{(x - \mu_{\text{in}})^4}{|\Omega_{\text{in}}| \sigma_{\text{in}}^4} H dx \\ &= \int_{\Omega} \frac{\partial}{\partial \phi} \left[\frac{(x - \mu_{\text{in}})^4}{|\Omega_{\text{in}}| \sigma_{\text{in}}^4} \right] H + \frac{(x - \mu_{\text{in}})^4}{|\Omega_{\text{in}}| \sigma_{\text{in}}^4} \delta \psi dx \\ &= \int_{\Omega} \frac{-4(x - \mu_{\text{in}})^3}{|\Omega_{\text{in}}| \sigma_{\text{in}}^4} \frac{\partial \mu_{\text{in}}}{\partial \phi} H - \frac{(x - \mu_{\text{in}})^4}{|\Omega_{\text{in}}|^2 \sigma_{\text{in}}^8} \\ &\quad \times \left(\sigma_{\text{in}}^4 \frac{\partial |\Omega_{\text{in}}|}{\partial \phi} + |\Omega_{\text{in}}| \frac{\partial \sigma_{\text{in}}^4}{\partial \phi} \right) H \\ &\quad + \frac{(x - \mu_{\text{in}})^4}{|\Omega_{\text{in}}| \sigma_{\text{in}}^4} \delta \psi dx \end{aligned} \quad (28)$$

where

$$\begin{aligned} \frac{\partial \mu_{\text{in}}}{\partial \phi} &= \frac{\partial}{\partial \phi} \int_{\Omega} \frac{x}{|\Omega_{\text{in}}|} H dx \\ &= \int_{\Omega} \frac{-x}{|\Omega_{\text{in}}|^2} \frac{\partial |\Omega_{\text{in}}|}{\partial \phi} H + \frac{x}{|\Omega_{\text{in}}|} \delta \psi dx \\ &= - \int_{\Omega} \frac{x}{|\Omega_{\text{in}}|^2} H dx \int_{\Omega} \delta \psi dx + \int_{\Omega} \frac{x}{|\Omega_{\text{in}}|} \delta \psi dx \end{aligned} \quad (29)$$

and

$$\begin{aligned} \frac{\partial \sigma_{\text{in}}^4}{\partial \phi} &= \frac{\partial}{\partial \phi} \left(\int_{\Omega} \frac{(x - \mu_{\text{in}})^2}{|\Omega_{\text{in}}|} H dx \right)^2 \\ &= 2 \sigma_{\text{in}}^2 \int_{\Omega} \frac{\partial}{\partial \phi} \left[\frac{(x - \mu_{\text{in}})^2}{|\Omega_{\text{in}}|} \right] H \\ &\quad + \frac{(x - \mu_{\text{in}})^2}{|\Omega_{\text{in}}|} \delta \psi dx. \end{aligned} \quad (30)$$

Inserting the various terms into each other, this yields (16).

Acknowledgments

The authors thank Eero P. Simoncelli for making his steerable pyramid toolbox publicly available. They also acknowledge the comments of the anonymous reviewers which helped to improve this paper. Of course, any remaining errors are the sole responsibility of the authors.

Note

1. To save horizontal space we abbreviate $\langle E'(\phi), \psi \rangle$ by $\partial E / \partial \phi$.

References

- Belongie, S., Carson, C., Greenspan, H., and Malik, J. 1998. Color- and texture-based image segmentation using EM and its application to content-based image retrieval. In *Proc. of the ICCV*.
- Brodatz, P. 1966. *Textures: A Photographic Album for Artists and Designers*. Dover Publications: New York.
- Buccigrossi, R.W. and Simoncelli, E.P. 1999. Image compression via joint statistical characterization in the wavelet domain. *IEEE Trans. on Image Proc.*, 8(12):1688–1700.
- Caselles, V., Kimmel, R., and Sapiro, G. 1997. Geodesic active contours. *Int. J. of Comp. Vision*, 22(1):61–79.
- Chan, T.F. and Vese, L.A. 2001. Active contours without edges. *IEEE Trans. on Image Proc.*, 10(2):266–277.
- Cremers, D., Sochen, N., and Schnörr, C. 2004. Dynamic labeling for variational recognition-driven image segmentation. In *Proc. of the ECCV* (to appear).
- Fowlkes, C., Belongie, S., Chung, F., and Malik, J. 2004. Spectral grouping using the Nyström method. *IEEE Trans. Patt. Anal. Mach. Intell.*, 26(3):214–225.
- Freeman, W.T. and Adelson, E.H. 1991. The design and use of steerable filters. *IEEE Trans. Patt. Anal. Mach. Intell.*, 13(9):891–906.
- Geusebroek, J.M. and Smeulders, A.W.M. 2003. Fragmentation in the vision of scenes. In *Proc. of the ICCV*, vol. 1, pp. 130–135.
- GMBV. 2002. First international workshop on generative-model-based vision (GMBV 2002), in conjunction with ECCV 2002. http://www.diku.dk/publikationer/tekniske_rapporter/2002/02-01/, Copenhagen, Denmark.
- Hansen, M.H. and Yu, B. 2001. Model selection and the principle of minimum description length. *J. of the Americ. Stat. Assoc.*, 96(454):746–774.
- Heeger, D.J. and Bergen, J.R. 1995. Pyramid-based texture analysis/synthesis. In *Proc. of SIGGRAPH*, pp. 229–238.
- Huang, J. and Mumford, D. 1999. Statistics of natural images and models. In *Proc. of the ICCV*, vol. 1, pp. 541–547.
- Jehan-Besson, S., Barlaud, M., and Aubert, G. 2003. DREAM²S: Deformable regions driven by an Eulerian accurate minimization method for image and video segmentations. *Int. J. of Comp. Vision*, 53(1):45–70.
- Joshi, R.L., Crump, V.J., and Fischer, T.R. 1995. Image subband coding using arithmetic coded trellis coded quantization. *IEEE Trans. on Circuits and Systems for Video Technology*, 5(6):515–523.

- Keuchel, J., Schnörr, C., Schellewald, C., and Cremers, D. 2003. Binary partitioning, perceptual grouping, and restoration with semidefinite programming. *IEEE Trans. Patt. Anal. Mach. Intell.*, 25(11):1364–1379.
- Kichenassamy, S., Kumar, A., Olver, P.J., Tannenbaum, A., and Yezzi, A.J. 1995. Gradient flows and geometric active contour models. In *Proc. of the ICCV*, pp. 810–815.
- Leclerc, Y.G. 1989. Constructing simple stable descriptions for image partitioning. *Int. J. of Comp. Vision*, 3(1):73–102.
- Liu, X. and Wang, D. 2000. Texture classification using spectral histograms. Technical report, Dept. of Comp. and Inform. Science, Ohio State University, Columbus, OH, USA.
- LoPresto, S.M. and Ramchandran, K. 1997. Image coding based on mixture modeling of wavelet coefficients and a fast estimation-quantization framework. In *Proc. of the Data Compr. Conf. (DCC)*, pp. 221–230.
- Mallat, S.G. 1998. A theory of multiresolution signal decomposition: The wavelet representation. *IEEE Trans. Patt. Anal. Mach. Intell.*, 11(7):674–693.
- Martin, D., Fowlkes, C., Tal, D., and Malik, J. 2001. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proc. of the ICCV*, vol. 2, pp. 416–423.
- Mumford, D. 1994. Bayesian rationale for the variational formulation. In *Geometry-Driven Diffusion in Computer Vision*, B. ter Haar Romeny (Ed.), Kluwer Acad. Publ.: Dordrecht, pp. 135–146.
- Mumford, D. and Shah, J. 1989. Optimal approximations by piecewise smooth functions and associated variational problems. *Comm. Pure Appl. Math.*, 42:577–685.
- Paragios, N. and Deriche, R. 2002. Geodesic active regions and level set methods for supervised texture segmentation. *Int. J. of Comp. Vision*, 46(3):223–247.
- Picard, R., Graczyk, C., Mann, S., Wachman, J., Picard, L., and Campbell, L. 1995. VisTex vision texture database. MIT Media Lab.
- Portilla, J. and Simoncelli, E.P. 2000. A parametric texture model based on joint statistics of complex wavelet coefficients. *Int. J. of Comp. Vision*, 40(1):49–71.
- Puzicha, J., Hofmann, T., and Buhmann, J.M. 1999. Histogram clustering for unsupervised segmentation and image retrieval. *Pattern Recog. Letters*, 20:899–909.
- Puzicha, J., Rubner, Y., Tomasi, C., and Buhmann, J.M. 1999. Empirical evaluation of dissimilarity measures for color and texture. In *Proc. of the ICCV*, vol. 2, pp. 1165–1172.
- Randen, T. and Husøy, J.H. 1999. Filtering for texture classification: A comparative study. *IEEE Trans. Patt. Anal. Mach. Intell.*, 21(4):291–310.
- Reininger, R.C. and Gibson, J.D. 1983. Distributions of the two-dimensional DCT coefficients for images. *IEEE Trans. on Communications*, 31(6):835–839.
- Rissanen, J. 1978. Modelling by the shortest data description. *Automatica*, 14:465–471.
- Rousson, M., Brox, T., and Deriche, R. 2003. Active unsupervised texture segmentation on a diffusion based feature space. In *Proc. of the IEEE Conf. on Comp. Vis. a. Patt. Recog.*, vol. 2, pp. 699–704.
- Rubner, Y., Tomasi, C., and Guibas, L.J. 2000. The earth movers distance as a metric for image retrieval. *Int. J. of Comp. Vision*, 40(2):99–121.
- Shi, J. and Malik, J. 2000. Normalized cuts and image segmentation. *IEEE Trans. Patt. Anal. Mach. Intell.*, 22(8):888–905.
- Simon, J. 1980. Differentiation with respect to the domain in boundary value problems. *Numerical Functional Analysis and Optimization*, 2(7):649–687.
- Simoncelli, E.P. and Adelson, E.H. 1996. Noise removal via Bayesian wavelet coring. In *Third Int'l Conf on Image Proc.*, Lausanne, IEEE Sig. Proc. Soc, pp. 379–382.
- Simoncelli, E.P. and Freeman, W.T. 1995. The steerable pyramid: A flexible architecture for multi-scale derivative computation. In *Second Int'l Conf. on Image Proc.*, Washington, DC.
- Sokolowski, J. and Zolesio, J.-P. 1991. *Introduction to Shape Optimization*. Springer: New York.
- Srivastava, A., Liu, X., and Grenander, U. 2002. Universal analytical forms for modeling image probabilities. *IEEE Trans. Patt. Anal. Mach. Intell.*, 24(9):1200–1214.
- Tang, M. and Ma, S. 2001. General scheme of region competition based on scale space. *IEEE Trans. Patt. Anal. Mach. Intell.*, 23(12):1366–1378.
- Tu, Z. and Zhu, S.C. 2002. Image segmentation by data-driven Markov Chain Monte Carlo. *IEEE Trans. Patt. Anal. Mach. Intell.*, 24(5):657–673.
- van Hateren, J.H. and van der Schaaf, A. 1998. Independent component filtering of natural images compared with simple cells in primary visual cortex. In *Proc. of the Royal Soc.*, London, pp. 359–366.
- Vese, L.A. and Chan, T.F. 2002. A multiphase level set framework for image segmentation using the mumford and shah model. *Int. J. of Comp. Vision*, 50(3):271–293.
- Wainwright, M.J., Simoncelli, E.P., and Willsky, A.S. 2001. Random cascades on wavelet trees and their use in analyzing and modeling natural images. *Applied and Computational Harmonic Analysis*, 11(1):89–123.
- Wallace, C.S. and Boulton, D.M. 1968. An information measure for classification. *Comp. J.*, 11(2):185–194.
- Wu, Y., Chan, K.L., and Huang, Y. 2003. Image texture classification based on finite gaussian mixture models. In *3rd Int. Workshop on Text. Anal. and Synth.*, ICCV, M. Chantler (Ed.), pp. 107–112.
- Wu, Y., Zhu, S.C., and Liu, X. 2000. Equivalence of Julesz ensembles and FRAME models. *Int. J. of Comp. Vision*, 38(3):247–265.
- Zhu, S.C., Liu, X., and Wu, Y.N. 2000. Exploring texture ensembles by efficient Markov Chain Monte Carlo-toward a 'trichromacy' theory of texture. *IEEE Trans. Patt. Anal. Mach. Intell.*, 22(6):554–569.
- Zhu, S.C. and Mumford, D. 1997. Prior learning and Gibbs reaction-diffusion. *IEEE Trans. Patt. Anal. Mach. Intell.*, 19(11):1236–1250.
- Zhu, S.C., Wu, Y., and Mumford, D. 1998. Filters, random fields and maximum entropy (FRAME). *Int. J. of Comp. Vision*, 27(2):1–20.
- Zhu, S.C. and Yuille, A.L. 1996. Region competition: Unifying snakes, region growing, and Bayes/MDL for multiband image segmentation. *IEEE Trans. Patt. Anal. Mach. Intell.*, 18(9):884–900.