Unsupervised Texture Segmentation with Nonparametric Neighborhood Statistics

Suyash P. Awate, Tolga Tasdizen, and Ross T. Whitaker

Scientific Computing and Imaging Institute, School of Computing,
University of Utah, Salt Lake City, UT 84112, USA

Abstract. This paper presents a novel approach to unsupervised texture segmentation that relies on a very general nonparametric statistical model of image neighborhoods. The method models image neighborhoods directly, without the construction of intermediate features. It does not rely on using specific descriptors that work for certain kinds of textures, but is rather based on a more generic approach that tries to adaptively capture the core properties of textures. It exploits the fundamental description of textures as images derived from stationary random fields and models the associated higher-order statistics nonparametrically. This general formulation enables the method to easily adapt to various kinds of textures. The method minimizes an entropy-based metric on the probability density functions of image neighborhoods to give an optimal segmentation. The entropy minimization drives a very fast level-set scheme that uses threshold dynamics, which allows for a very rapid evolution towards the optimal segmentation during the initial iterations. The method does not rely on a training stage and, hence, is unsupervised. It automatically tunes its important internal parameters based on the information content of the data. The method generalizes in a straightforward manner from the two-region case to an arbitrary number of regions and incorporates an efficient multi-phase level-set framework. This paper presents numerous results, for both the two-texture and multiple-texture cases, using synthetic and real images that include electron-microscopy images.

1 Introduction

Image segmentation is one of the most extensively studied problems in computer vision. The literature gives numerous approaches based on a variety of criteria including intensity, color, texture, depth, and motion. This paper addresses the problem of segmenting textured images. Textured regions do not typically adhere to the piecewise-smooth or piecewise-constant assumptions that characterize most intensity-based segmentation problems. Julesz [13] pioneered the statistical analysis of textures and characterized textures as possessing regularity in the higher-order intensity statistics. This establishes the description of a textured image, or a Julesz ensemble, as one derived from stationary random fields [21]. This principle forms the foundation of the approach in this paper.

In recent years, researchers have advanced the state-of-the-art in texture segmentation in several important directions. An important direction relates to

the mechanism used to model or quantify the regularity in image textures. Researchers have developed progressively richer descriptions of local image geometry and thereby captured more complex and subtle distinctions between textures [2, 23, 24]. In another direction, researchers have expressed the dissimilarity between textures through sophisticated statistically-based metrics [4, 15, 19, 14, 22]. Furthermore, research in texture segmentation, like image segmentation in general, has focused on more robust mechanisms for enforcing geometric smoothness in the segmented-region shapes. This is usually done via the construction of a patchwork of regions that simultaneously minimize a set of geometric and statistical criteria [26].

This paper advances the state-of-the-art in texture segmentation by exploiting the principle characteristics of a texture coupled with the generality of nonparametric statistical modeling. The method relies on an information-theoretic metric on the statistics of image neighborhoods that reside in high-dimensional spaces. The nonparametric modeling of the statistics of the stationary random field imposes very few restrictions on the statistical structure of neighborhoods. This enables the method to easily adapt to a variety of textures. The method does not rely on a training stage and, hence, is unsupervised. These properties make it is easily applicable to a wide range of texture-segmentation problems. Moreover, the method incorporates relatively recent advances in level-set evolution strategies that use threshold dynamics [11, 10].

The rest of the paper is organized as follows. Section 2 discusses recent works in texture segmentation and their relationship to the proposed method. Section 3 describes the optimal formulation with an entropy-based energy on higher-order image statistics. Entropy optimization entails the estimation of probability density functions. Hence, Section 4 describes a nonparametric multivariate density estimation technique. It also describes the general problems associated with density estimation in high-dimensional spaces and provides some intuition behind the success of the proposed method in spite of these difficulties. Section 5 gives the optimization strategy using a very fast level-set scheme that uses threshold dynamics, along with the associated algorithm. Section 6 addresses several important practical issues pertaining to nonparametric statistical estimation and its application to image neighborhoods. Section 7 gives experimental results on numerous real and synthetic images, including electron-microscopy medical images. Section 8 summarizes the contributions of the paper and presents ideas for further exploration.

2   Related Work

Much of the previous work in texture segmentation employs filter banks, comprising both isotropic and anisotropic filters, to capture texture statistics. For instance, researchers have used Gabor-filter responses to discriminate between different kinds of textures [19, 23, 24]. Gabor filters are a prominent example of a very large class of oriented multiscale filters [4, 3]. This approach emphasizes the extraction of appropriate features for discriminating between specific textures,

which is typically a non-trivial task. The proposed method, on the other hand, does not rely on using specific descriptors that work for certain kinds of textures, but is based on a more generic approach that tries to adaptively capture the core properties of a wide variety of textures.

Researchers have also investigated using more compact sets of texture features. For instance, Bigun *et al.* [2] use the structure tensor (a second-order moment matrix used, e.g., to analyze flow-like patterns [32]) to detect local orientation. Rousson *et al.* [22] refine this strategy by using vector-valued anisotropic diffusion, instead of Gaussian blurring, on the feature space formed using the components of the structure tensor. This strategy requires the structure tensors to have a sufficient degree of homogeneity within regions as well as sufficient dissimilarity between regions. However, as the coming paragraphs explain, not all images meet these criteria.

Other approaches use the intensity (or grayscale) histograms to distinguish between textures [15, 14]. However, the grayscale intensity statistics (i.e. 1D histograms), may fail to capture the *geometric* structure of neighborhoods, which is critical for distinguishing textures with similar 1D histograms. The proposed method exploits higher-order image statistics, modeled nonparametrically, to adaptively capture the geometric regularity in textures.

Figure 1(a) shows two textures that are both *irregular* (in addition to having similar means and gradient-magnitudes) that would pose a challenge for structure-tensor-based approaches such as [2, 22]. In Figure 1(b) the textures differ *only in scale*. Approaches based on structure tensors at a single scale would fail to distinguish such cases, as reported in [22]. Approaches solely using intensity histograms would also fail here. In Figure 1(c) the textures have identical histograms, identical scale, and an almost-identical set of structure-tensor matrix components. In this case, the above-mentioned approaches [2, 22] would face a formidable challenge. The proposed method, on the other hand, incorporating
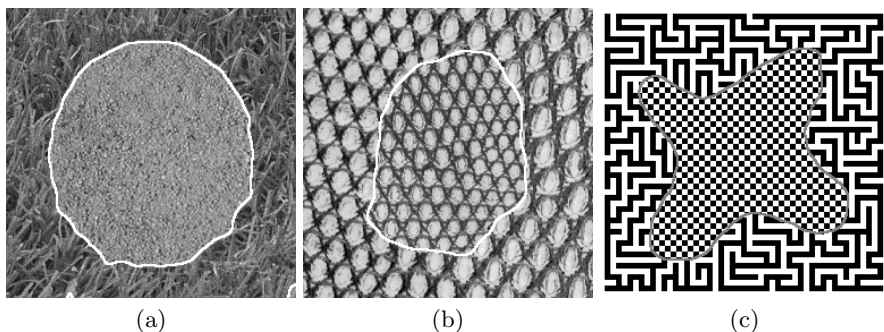

(a)   (b)   (c)

Fig. 1. Segmentations with the proposed approach (depicted by white/gray outlines) for (a) Brodatz textures for sand and grass— both *irregular* textures with similar gradient magnitudes, (b) Brodatz textures differing *only in scale*, and (c) synthetic textures with identical histograms, identical scales, and an almost-identical set of structure-tensor matrix components

a fundamentally richer texture description, produces successful segmentations (depicted by white/gray outlines) for all the images in Figure 1.

Recently, researchers have investigated more direct approaches towards modeling image statistics. For instance, the dynamic-texture segmentation approach by Doretto *et al.* [6] uses a Gauss-Markov process to model the relationships among pixels within regions and over time. However, that approach assumes a Gaussian process for image intensities, a restrictive assumption that cannot easily account for complex or subtle texture geometries [6, 22, 4]. Rousson *et al.* [22] use nonparametric statistics for one of the channels (the image-intensity histogram) in their feature space to counter this restriction and the proposed method generalizes that strategy to the complete higher-order image statistics.

Popat *et al.* [20] were among the first to use nonparametric Markov sampling in images. Their method takes a *supervised* approach for learning neighborhood relationships. They attempt to capture the higher-order nonlinear image statistics via cluster-based nonparametric density estimation and apply their technique for texture classification. Varma and Zisserman [28] used a similar training-based approach to classify textures based on a small Markov neighborhood that was demonstrably superior to filter based approaches. Indeed, researchers analyzing the statistics of 3×3 patches in images, in the corresponding high-dimensional spaces, have found the data to be concentrated in clusters and low-dimensional manifolds exhibiting nontrivial topologies [16, 5]. The proposed approach also relies on the principle that textures exhibit regularity in neighborhood structure, but this regularity is discovered for each texture individually in a nonparametric manner. The proposed method builds on the work in [1], which lays down the essentials for unsupervised learning of higher-order image statistics. That work, however, focuses on image restoration.

The literature dealing with texture synthesis also sheds some light on the proposed method. Texture synthesis algorithms rely on texture statistics from an input image to construct novel images that exhibit a qualitative resemblance to the input texture [9, 31]. This paper describes a very different application, but the texture-synthesis work demonstrates the power of neighborhood statistics in capturing the essential aspects of texture.

Lastly, this paper also borrows from a rather extensive body of work on variational methods for image segmentation [26], in particular the Mumford-Shah model [18], its extensions to motion, depth, and texture, and its implementation via level-set flows [29]. The proposed method employs the very fast approximation proposed by Esedoglu and Tsai [11, 10] based on threshold dynamics, and extends it to include multiple regions within the variational framework.

3   Neighborhood Statistics for Texture Segmentation

This section introduces the random-field texture model, along with the associated notation, and then describes the optimal segmentation formulation based on entropy minimization.

3.1   Random Field Texture Model

A random field [7] is a family of random variables $X(\Omega; T)$, for an index set T, where, for each fixed $T = t$, the random variable $X(\Omega; t)$, or simply $X(t)$, is defined on the sample space Ω. If we let T be a set of points defined on a discrete Cartesian grid and fix $\Omega = \omega$, we have a realization of the random field called the *digital image*, $X(\omega, T)$. In this case $\{t\}_{t \in T}$ is the set of pixels in the image. For two-dimensional images t is a two-vector. We denote a specific realization $X(\omega; t)$ (the image), as a deterministic function $x(t)$.

   If we associate with T a family of pixel neighborhoods $N = \{N_t\}_{t \in T}$ such that $N_t \subset T$, and $s \in N_t$ if and only if $t \in N_s$, then N is called a neighborhood system for the set T and points in N_t are called neighbors of t. We define a random vector $Z(t) = \{X(t)\}_{t \in N_t}$, denoting its realization by $z(t)$, corresponding to the set of intensities at the neighbors of pixel t. We refer to the statistics of the random vector Z as *higher-order* statistics. Following the definition of texture as a Julesz ensemble [13, 21], we assume that the intensities in each texture region arise out of a stationary ergodic random field.

3.2   Optimal Segmentation by Entropy Minimization on Higher-Order Statistics

Consider a random variable $L(t)$, associated with each pixel $t \in T$, that gives the region the pixel t belongs to. For a good segmentation, knowing the neighborhood intensities (z) tells us the unique pixel class (k). Also, knowing the pixel class gives us a good indication of what the neighborhood is. This functional dependence is captured naturally in the concept of mutual information. Thus, the optimal segmentation is one that maximizes the mutual information between L and Z:

$$I(L, Z) = h(Z) - h(Z|L) = h(Z) - \sum_{k=1}^{K} P(L = k) h(Z|L = k), \qquad (1)$$

where $h(\cdot)$ denotes the entropy of the random variable. The entropy of the higher-order PDF associated with the entire image, $h(Z)$, is a constant for an image and is irrelevant for the optimization. Let $\{T_k\}_{k=1}^{K}$ denote a mutually-exclusive and exhaustive decomposition of the image domain T into K texture regions. Let $P_k(Z(t) = z(t))$ be the probability of observing the image neighborhood $z(t)$ given that the center pixel of the neighborhood belongs to the texture region k. We define the energy associated with the set of K texture probability density functions (PDFs), i.e.

$$E = \sum_{k=1}^{K} P(L = k) h(Z|L = k). \qquad (2)$$

The entropy

$$h(Z|L = k) = -\int_{\Re^m} P_k(Z(t_k) = z(t_k)) \log P_k(Z(t_k) = z(t_k)) dz, \qquad (3)$$

where $m = |N_t|$ is the neighborhood size and t_k is any pixel belonging to region k—for any $t_k \in T_k$ the PDF $P_k(\cdot)$ remains the same due to the assumed stationarity. Let $R_k : T \to \{0, 1\}$ denote the indicator function for region T_k, i.e. $R_k(t) = 1$ for $t \in T_k$ and $R_k(t) = 0$ otherwise. Considering the intensities in each region as derived from a stationary ergodic random field to approximate entropy, and using $P(L = k) = |T_k|/|T|$, gives

$$E \approx -\sum_{k=1}^{K} \left(\frac{P(L = k)}{|T_k|} \sum_{t \in T} R_k(t) \log P_k(z(t)) \right) \tag{4}$$

$$= \frac{-1}{|T|} \sum_{k=1}^{K} \sum_{t \in T} R_k(t) \log P_k(z(t)). \tag{5}$$

Thus, the optimal segmentation is the set of functions R_k for which E attains a minimum. The strategy in this paper is to minimize the total entropy given in (4) by manipulating the regions defined by R_k. This rather-large nonlinear optimization problem potentially has many local minima. To regularize the solution, variational formulations typically penalize the boundary lengths of the segmented regions [18]. The objective function, after incorporating this penalty using a Lagrange multiplier, now becomes

$$E + \alpha \sum_{k=1}^{K} \sum_{t \in T} \| \nabla_t R_k(t) \|, \tag{6}$$

where α is the regularization parameter and ∇_t denotes a discrete spatial-gradient operator. In this framework, the critical issue lies in the estimation of $P_k(z(t))$, and the next section focuses on addressing this issue.

4   Nonparametric Multivariate Density Estimation

Entropy optimization entails the estimation of higher-order conditional PDFs. This introduces the challenge of high-dimensional, scattered-data interpolation, even for modest sized image neighborhoods. High-dimensional spaces are notoriously challenging for data analysis (regarded as the *the curse of dimensionality* [27, 25]), because they are so sparsely populated. Despite theoretical arguments suggesting that density estimation beyond a few dimensions is impractical, the empirical evidence from the literature is more optimistic [25, 20]. The results in this paper confirm that observation. Furthermore, stationarity implies that the random vector Z exhibits identical marginal PDFs, and thereby lends itself to more accurate density estimates [25, 27]. We also rely on the neighborhoods in natural images having a lower-dimensional topology in the multi-dimensional feature space [16, 5]. Therefore, *locally* (in the feature space) the PDFs of images are lower dimensional entities that lend themselves to better density estimation.

We use the Parzen-window nonparametric density estimation technique [8] with an n-dimensional Gaussian kernel $G_n(z, \Psi_n)$, where $n = |N_t|$. We have no

a priori information on the structure of the PDFs, and therefore we choose an isotropic Gaussian, i.e. $\Psi_n = \sigma^2 I_n$, where I_n is the $n \times n$ identity matrix. For a stationary ergodic random field, the multivariate Parzen-window estimate is

$$P_k(Z(t) = z(t)) \approx \frac{1}{|A_{k,t}|} \sum_{s \in A_{k,t}} G_n(z(t) - z(s), \Psi_n), \qquad (7)$$

where the set $A_{k,t}$ is a small subset of T_k chosen randomly, from a uniform PDF, for each t. This results in a *stochastic* estimate of the entropy that helps alleviate the effects of spurious local maxima introduced in the Parzen-window density estimate [30]. We refer to this sampling strategy as the *global-sampling* strategy. Selecting appropriate values of the kernel-width σ is important for success, and Section 6 presents a data-driven strategy for the same.

5   Fast Level-Set Optimization Using Threshold Dynamics

The level-set framework [26] is an attractive option for solving the variational problem defined by (6), because it does not restrict either the shapes or the topologies of regions. However, classical level-set evolution schemes for front-tracking based on narrow-band strategies entail some significant computational costs—in particular, the CFL condition for numerical stability [26] limits the motion of the moving wavefront (region boundaries) to one pixel per iteration.

Recently, Esedoglu and Tsai introduced a fast level-set algorithm based on threshold dynamics [11, 10] for minimizing Mumford-Shah type energies. The proposed method adopts their approach for the level-set evolution but relies on a *multiphase extension* of the basic formulation to enable multiple-texture segmentation [17, 29]. In this method, the embeddings, one for each phase, are maintained as piecewise-constant binary functions. This method, essentially, evolves the level-set by first updating the embeddings using the PDE-driven force, and then regularizing the region boundaries by Gaussian smoothing the embedding followed by re-thresholding. This approach needs to neither keep track of points near interfaces nor maintain distance transforms for embeddings. At the same time it allows new components of a region to crop up at remote locations. We have found that this last property allows for very rapid level-set evolution when the level-set location is far from the optimum.

We now let $\{R_k\}_{k=1}^{K}$ be a set of level-set functions. The segmentation for texture k is then defined as $T_k = \{t \in T | R_k(t) > R_j(t), \forall j \neq k\}$. It is important to realize that coupling (6) and (7) creates nested region integrals that introduce extra terms in the gradient flow associated with the level-set evolution [15, 22]. The shape-derivative tool [12], specifically designed to handle such situations, gives the level-set speed term for minimizing the energy defined in (6) as

$$\frac{\partial R_k(t)}{\partial \tau} = \log P_k(z(t)) + \frac{1}{|T_k|} \sum_{s \in T_k} \frac{G_n(z(s) - z(t), \Psi_n)}{P_k(z(s))} + \alpha \nabla_t \cdot \left(\frac{\nabla_t R_k(t)}{\| \nabla_t R_k(t) \|} \right),$$

$$(8)$$

where τ denotes the time-evolution variable [15, 22].

To obtain an initial segmentation $\{R_k^0\}_{k=1}^K$, the proposed method uses randomly generated regions, as shown in Section 7, based on the following algorithm.

1. Generate K images of uniform random noise, one for each R_k^0.
2. Convolve each R_k^0 with a chosen Gaussian kernel.
3. $\forall k, t$ do: if $R_k^0(t) > R_j^0(t), \forall j \neq k$ then set $R_k^0(t) = 1$, otherwise set $R_k^0(t) = 0$.

The iterations in Esedoglu and Tsai's fast level-set evolution scheme [11, 10], given a segmentation $\{R_k^m\}_{k=1}^K$ at iteration m, proceed as follows.

1. $\forall k, t$ do:
   (a) Estimate $P_k(z(t))$ nonparametrically, as described in Section 4.
   (b) $R_k'(t) = R_k^m(t) + \beta \left(\log P_k(z(t)) + \frac{1}{T_k} \sum_{s \in T_k} \frac{G_n(z(s) - z(t), \Psi_n)}{P_k(z(s))} \right)$
2. Compute $R_k'' = R_k' \otimes \mathbb{N}(0, \gamma^2)$, where $\otimes$ denotes convolution and $\mathbb{N}(0, \gamma^2)$ is a Gaussian kernel with zero mean and standard deviation γ.
3. $\forall k, t$ do: if $R_k''(t) > R_j''(t), \forall j \neq k$ then set $R_k^{m+1}(t) = 1$, otherwise set $R_k^{m+1}(t) = 0$.
4. Stop upon convergence, i.e. when $\| R_k^{m+1} - R_k^m \|_2 < \delta$, a small threshold.

For a detailed discussion on the relationship between the new parameters γ, β, and the parameter α in the traditional level-set framework, we refer the reader to [11, 10]. In short, increasing β corresponds to increasing the PDE-driven force on the level-set evolution and increasing γ results in smoother region boundaries.

6   Important Implementation Issues

This section discusses several practical issues that are crucial for the effectiveness of the entropy reduction scheme. The work in [1] presents a detailed discussion on these issues.

Data-driven choice for the Parzen-window kernel width: Using appropriate values of the Parzen-window parameters is important for success, and that can be especially difficult in the high-dimensional spaces associated with higher-order statistics. The best choice depends on a variety of factors including the sample size $|A_{k,t}|$ and the natural variability in the data. To address this issue we fall back on our previous work for automatically choosing the optimal values [1]. In that work, which focused on image restoration, we choose σ to minimize the entropy of the associated PDF via a Newton-Raphson optimization scheme. We have found that such a σ, i.e. one minimizing the entropy of Z, can be too discriminative for the purpose of texture segmentation, splitting the image into many more regions than what may be appropriate. Hence, in this paper, we set σ to be 10 times as large. The choice of the precise value of this multiplicative factor is not critical and we have found that the algorithm is quite robust to small changes in this parameter.

Data-driven choice for the Parzen-window sample size: Our experiments show [1] that for sufficiently large $|A_{k,t}|$ additional samples do not significantly affect the estimates of entropy and σ, and thus $|A_{k,t}|$ can also be selected automatically from the input data. For the Parzen-windowing scheme we choose 500 samples, i.e. $|A_{k,t}| = 500$, uniformly distributed over each region.

Neighborhood size and shape: The quality of the results also depend on the neighborhood size. We choose the size relative to the size of the textures in the image. Bigger neighborhoods are generally more effective but increase the computational cost. To obtain rotationally invariant neighborhoods, we use a metric in the feature space that controls the influence of each neighborhood pixel so that the distances in this space are less sensitive to neighborhood rotations [1]. In this way, feature space dimensions close to the corners of the square neighborhood shrink so that they do not significantly influence the filtering. Likewise, image boundaries are handled through such anisotropic metrics so that they do not distort the neighborhood statistics of the image.

7   Experiments and Results

This section presents results from experiments with real and synthetic data. The number of regions K is a user parameter and should be chosen appropriately. The neighborhood size, in the current implementation, is also a user parameter. This can be improved by using a multi-resolution scheme for the image representation and constitutes an important area of future work. We use 9×9 neighborhoods, $\beta = 2$, and $\gamma = 3$ for all examples, unless stated otherwise. Each iteration of the proposed method takes about 3 minutes for a 256×256 image on a standard Pentium workstation. Figure 2(a) shows a level-set initialization $\{R_k^0\}_{k=1}^K$ as a randomly generated image with $K = 2$ regions.
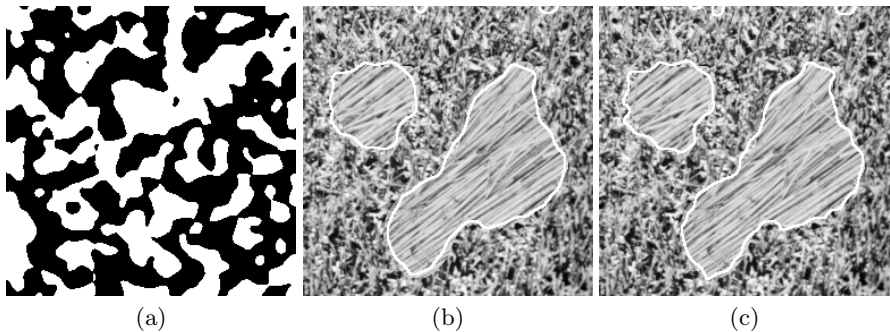

(a)   (b)   (c)

Fig. 2. Two-texture segmentation. (a) Random initial segmentation for an image having two Brodatz textures for grass and straw. The black and white intensities denote the two regions. (b) Segmentation after stage 1; *global* samples only (see text). (c) Segmentation after stage 2; *local* and *global* samples (see text).

The level-set scheme using threshold dynamics, coupled with the *global-sampling* strategy as explained in Section 4, makes the level sets evolve very fast towards the optimal segmentation. We have found that, starting from the random initialization, just a few iterations (less than 10) are sufficient to reach an almost-optimal segmentation. However, this sampling strategy is sometimes unable to give very accurate boundaries. This is because, in practice, the texture boundaries present neighborhoods overlapping both textures and exhibiting subtleties that may not be captured by the global sampling. Moreover, the joining of intricate textures may inherently make the boundary location significantly fuzzy so that it may be impossible even for humans to define the true segmentation. Figure 2(b) depicts this behavior. In this case, for each point t, selecting a larger portion of the samples in $A_{k,t}$ from a region close to t would help. Hence, we propose a second stage of level-set evolution that incorporates *local* sampling, in addition to global sampling, and is initialized with the segmentation resulting from the first stage. We found that such a scheme consistently yields better segmentations. Figure 2(c) shows the final segmentation. We have used about 250 local samples taken from a Gaussian distribution, with a variance of 900, centered at the concerned pixel. Furthermore, we have found that the method performs well for any choice of the variance such that the Gaussian distribution encompasses more than several hundred pixels. Note that given this variance, both $|A_{k,t}|$ and the Parzen-window σ are computed automatically in a data-driven manner, as explained before in Section 6.

Figure 3 gives examples dealing with multiple-texture segmentation. Figure 3(a) shows a randomly generated initialization with three regions that leads to the final segmentation in Figure 3(b). In this case the proposed algorithm uses a multi-phase extension of the fast threshold-dynamics based scheme [11, 10]. Figure 3(c) shows another multiple-texture segmentation with four textures.

Figure 4 shows electron-microscopy images of cellular structures. Because the original images severely lacked contrast, we preprocessed them using adaptive
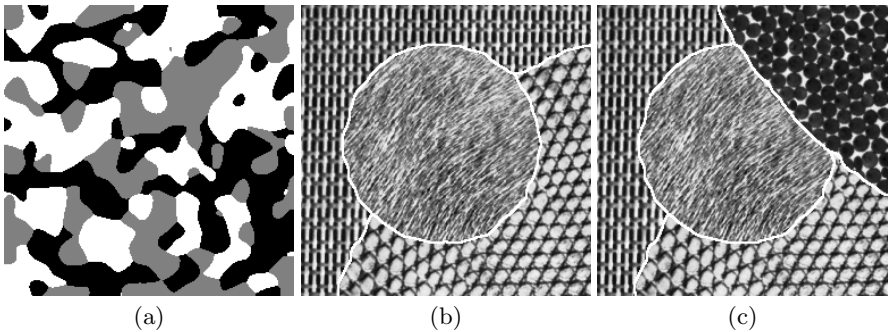

(a)   (b)   (c)

Fig. 3. Multiple-texture segmentation. (a) Random initial segmentation containing three regions for the image in (b). (b) Final segmentation for an image with three Brodatz textures, including both irregular and regular textures. (c) Final segmentation for an image with four Brodatz textures.
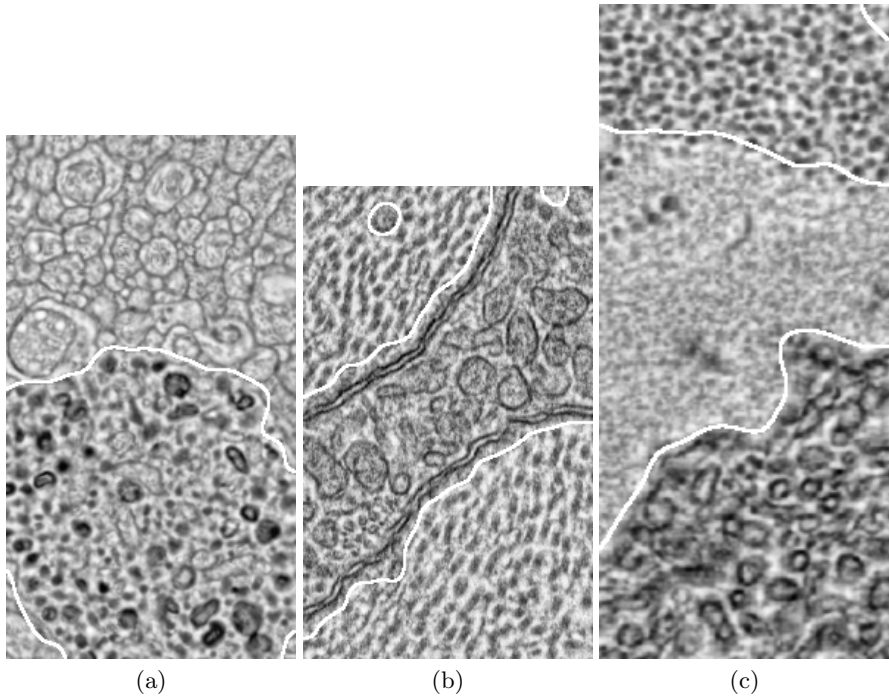
Fig. 4. Final segmentations for electron-microscopy images of rabbit retinal cells for (a),(b) the two-texture case, and (c) the three-texture case

histogram equalization before applying the proposed texture-segmentation method. Figure 4 shows the enhanced images. These images are challenging to segment using edge or intensity information because of reduced textural homogeneity in the regions. The discriminating feature for these cell types is their subtle textures formed by the arrangements of sub-cellular structures. To capture the large-scale structures in the images we used larger neighborhood sizes of 13×13. We combine this with a higher γ for increased boundary regularization. Figure 4(a) demonstrates a successful segmentation. In Figure 4(b) the two cell types are segmented to a good degree of accuracy; however, notice that the membranes between the cells are grouped together with the middle cell. A third texture region could be used for the membrane, but this is not a trivial extension due to the thin, elongated geometric structure of the membrane and the associated difficulties in the Parzen-window sampling. The hole in the region on the top left forms precisely because the region contains a large elliptical patch that is identical to such patches in the other cell. Figure 4(c) shows a successful three-texture segmentation for another image.

Figure 5(a) shows a zebra example that occurs quite often in the texture-segmentation literature, e.g. [23, 22]. Figures 5(b) and 5(c) show other zebras. Here, the proposed method performs well to differentiate the striped patterns, with varying orientations and scales, from the irregular grass texture. The grass
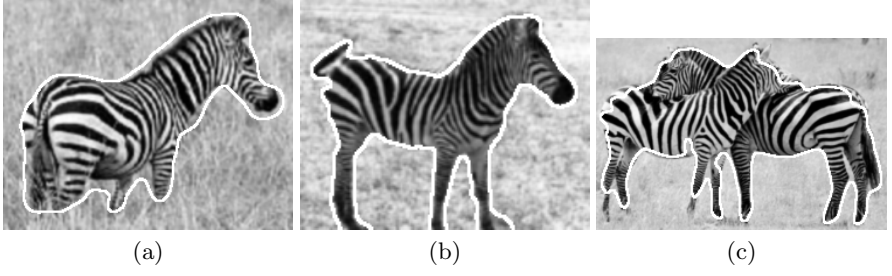
(a)   (b)   (c)

Fig. 5. Final segmentations for real images of Zebras

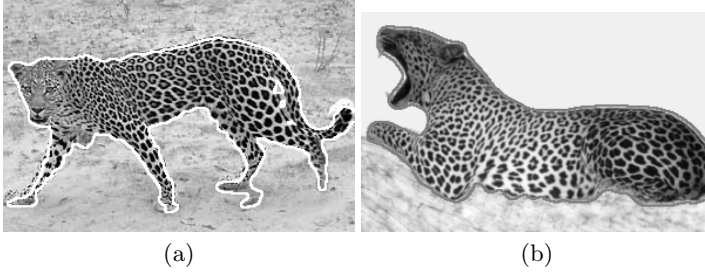(a)   (b)

Fig. 6. Final segmentations for real images of Leopards. Note: The segmentation outline for image (b) is shown in gray.

texture depicts homogeneous statistics. The striped patterns on the Zebras' body, although incorporating many variations, change gradually from one part of the body to another. Hence, neighborhoods from these patterns form one continuous manifold in the associated high-dimensional space, which is captured by the method as a single texture class.

Figure 6(a) shows the successful segmentation of the Leopard with the random sand texture in the background. Figure 6(b) shows an image that actually contains three different kinds of textures, where the background is split into two textures. Because we constrained the number of regions to be two, the method grouped two of the background textures into the same region.

8   Conclusions and Discussion

This paper presents a novel approach for texture segmentation exploiting the higher-order image statistics that principally define texture. The proposed method adaptively learns the image statistics via nonparametric density estimation and does not rely on specific texture descriptors. It relies on the information content of input data for setting important parameters, and does not require significant parameter tuning. Moreover, it does not rely on any kind of training and, hence, is easily applicable to a wide spectrum of texture segmentation tasks. The paper applies the proposed method to segment different cell

types in electron-microscopy medical images, giving successful segmentations. It also demonstrates the effectiveness of the method on real images of Zebras and Leopards, as well as numerous examples with Brodatz textures. The method incorporates a very fast multiphase level-set evolution framework using threshold dynamics [11, 10].

The algorithmic complexity of the method is $O(K|T||A_{k,t}|S^D)$ where D is the image dimension and S is the extent of the neighborhood along a dimension. This grows exponentially with D, and our current results are limited to 2D images. The literature suggests some improvements, e.g. reduction in the computational complexity via the improved fast-gauss transform [33]. In the current implementation, the neighborhood size is chosen manually and this is a limitation. This can be improved by defining a feature space comprising neighborhoods at multiple scales. These are important areas for future work.

Acknowledgments

References

1. S. P. Awate and R. T. Whitaker. Unsupervised, Information-Theoretic, Adaptive Image Filtering for Image Restoration. *IEEE Trans. Pattern Anal. Mach. Intell. (PAMI)*, 28(3):364–376, March 2006.
2. J. Bigun, G. H. Granlund, and J. Wiklund. Multidimensional orientation estimation with applications to texture analysis and optical flow. *IEEE Trans. Pattern Anal. Mach. Intell.*, 13(8):775–790, 1991.
3. R. Boomgaard and J. Weijer. Robust estimation of orientation for texture analysis. In *2nd Int. Workshop on Texture Analysis and Synthesis*, 2002.
4. J. S. de Bonet and P. Viola. Texture recognition using a non-parametric multi-scale statistical model. In *Proc. IEEE Conf. on Comp. Vision and Pattern Recog.*, pages 641–647, 1998.
5. V. de Silva and G. Carlsson. Topological estimation using witness complexes. *Symposium on Point-Based Graphics*, 2004.
6. G. Doretto, D. Cremers, P. Favaro, and S. Soatto. Dynamic texture segmentation. In *Proc. Int. Conf. Computer Vision*, pages 1236–1242, 2003.
7. E. Dougherty. *Random Processes for Image and Signal Processing*. Wiley, 1998.
8. R. Duda, P. Hart, and D. Stork. *Pattern Classification*. Wiley, 2001.
9. A. A. Efros and T. K. Leung. Texture synthesis by non-parametric sampling. In *Int. Conf. Computer Vision*, pages 1033–1038, 1999.
10. S. Esedoglu, S. Ruuth, and R. Tsai. Threshold dynamics for shape reconstruction and disocclusion. In *Proc. Int. Conf. Image Processing*, pages 502–505, 2005.
11. S. Esedoglu and Y.-H. R. Tsai. Threshold dynamics for the piecewise constant mumford-shah functional. Number CAM-04-63, 2004.

12. S. Jehan-Besson, M. Barlaud, and G. Aubert. Dream2s: Deformable regions driven by an eulerian accurate minimization method for image and video segmentation. In *Proc. European Conf. on Computer Vision-Part III*, pages 365–380, 2002.

13. B. Julesz. Visual pattern discrimination. *IRE Trans. Info. Theory*, IT(8):84–92, 1962.

14. T. Kadir and M. Brady. Unsupervised non-parametric region segmentation using level sets. In *Proc. of IEEE Int. Conf. Comp. Vision*, pages 1267–1274, 2003.

15. J. Kim, J. W. Fisher, A. Yezzi, M. Cetin, and A. S. Willsky. Nonparametric methods for image segmentation using information theory and curve evolution. In *Proc. IEEE Int. Conf. on Image Processing*, pages 797–800, 2002.

16. A. Lee, K. Pedersen, and D. Mumford. The nonlinear statistics of high-contrast patches in natural images. *Int. J. Comput. Vision*, 54(1-3):83–103, 2003.

17. B. Merriman, J. K. Bence, and S. Osher. Motion of multiple junctions: A level set approach,. Technical Report CAM-93-19, Dept. Mathematics, UCLA, 1993.

18. D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *Com. Pure and App. Math.*, 42:577–685, 1989.

19. N. Paragios and R. Deriche. Geodesic active regions and level set methods for supervised texture segmentation. *Int. J. Comput. Vision*, 46(3):223–247, 2002.

20. K. Popat and R. Picard. Cluster based probability model and its application to image and texture processing. *IEEE Trans. Image Processing*, 6(2):268–284, 1997.

21. J. Portilla and E. Simoncelli. A parametric texture model based on joint statistics of complex wavelet coefficients. *Int. J. Comput. Vision*, 40(1):49–70, 2000.

22. M. Rousson, T. Brox, and R. Deriche. Active unsupervised texture segmentation on a diffusion based feature space. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pages 699–706. IEEE Computer Society, 2003.

23. C. Sagiv, N. A. Sochen, and Y. Y. Zeevi. Texture segmentation via a diffusion segmentation scheme in the gabor feature space. In *2nd Int. Workshop on Texture Analysis and Synthesis*, 2002.

24. B. Sandberg, T. Chan, and L. Vese. A level-set and gabor-based active contour algorithm for segmenting textured images. Technical Report CAM-02-39, Dept. Mathematics, UCLA, 2002.

25. D. W. Scott. *Multivariate Density Estimation*. Wiley, 1992.

26. J. Sethian. *Level Set Methods and Fast Marching Methods*. Cambridge Univ. Press, 1999.

27. B. Silverman. *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, 1986.

28. M. Varma and A. Zisserman. Texture classification: Are filter banks necessary ? In *Proc. IEEE Conf. on Comp. Vision and Pattern Recog.*, pages 691–698, 2003.

29. L. Vese and T. Chan. A multiphase level set framework for image segmentation using the mumford and shah model. Technical Report CAM-01-25, Dept. Mathematics, UCLA, 2001.

30. P. Viola and W. Wells. Alignment by maximization of mutual information. In *Int. Conf. Comp. Vision*, pages 16–23, 1995.

31. L. Wei and M. Levoy. Order-independent texture synthesis. *Stanford University Computer Science Department Tech. Report TR-2002-01*, 2002.

32. J. Weickert. Coherence-enhancing diffusion filtering. *Int. J. Comp. Vis.*, 31:111–127, 1999.

33. C. Yang, R. Duraiswami, N. Gumerov, and L. Davis. Improved fast gauss transform and efficient kernel density estimation. In *Int. Conf. Comp. Vision*, pages 464–471, 2003.