INSTITUT NATIONAL DE RECHERCHE EN INFORMATIQUE ET EN AUTOMATIQUE

Active Unsupervised Texture Segmentation on a Diffusion Based Feature Space

Mikaël Rousson — Thomas Brox — Rachid Deriche

N° 4695

January 2003

THÈME 3

Rapport de recherche

ISSN 0249-6399 ISRN INRIA/RR--4695--FR+ENG

Active Unsupervised Texture Segmentation on a Diffusion Based Feature Space

Mikaël Rousson , Thomas Brox , Rachid Deriche

Thème 3 — Interaction homme-machine,
images, données, connaissances
Projet Odyssee

Rapport de recherche n° 4695 — January 2003 — 20 pages

Abstract: In this report, we propose a novel and efficient approach for active unsurpervised texture segmentation. First, we show how we can extract a small set of good features for texture segmentation based on the structure tensor and nonlinear diffusion. Then, we propose a variational framework that allows to incorporate these features in a level set based unsupervised segmentation process that adaptively takes into account their estimated statistical information inside and outside the region to segment. Unlike features obtained by Gabor filters, our approach naturally leads to a significantly reduced number of feature channels. Thus, the supervised part of a texture segmentation algorithm, where the choice of good feature channels has to be learned in advance, can be omitted, and we get an efficient solution for unsupervised texture segmentation. The actual segmentation process based on the new features is an active and adaptative contour model that estimates dynamically probability density functions inside and outside a region and produces very convincing results. It is implemented using a fast level set based active contour technique and has been tested on various real textured images. The performance of the approach is favorably compared to recent studies.

Key-words: Level Set Theory, Texture Segmentation, Adaptative Image Segmentation, Nonlinear Diffusion, Structure Tensor.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles, BP 93, 06902 Sophia Antipolis Cedex (France)
Téléphone : +33 4 92 38 77 77 — Télécopie : +33 4 92 38 77 65

Segmentation active et non supervisée d'images texturées à l'aide d'une diffusion non linéaire du tenseur de structure

Résumé : Dans ce rapport, nous proposons une nouvelle approche pour la segmentation active d'images texturées. Tout d'abord, nous présentons une méthode d'extraction d'un nombre restreint de composantes pour caractériser l'information de texture présente dans l'image à segmenter. Ce processus est basé sur le tenseur de structure et la diffusion non-linéaire. Ensuite, nous proposons un cadre variationnel afin d'incorporer ces différentes caractérisitiques dans un processus de segmentation adaptatif et non supervisé, basé sur les ensembles de niveaux. Contrairement aux approches utilisant des filtres de Gabor pour extraire l'information texture, notre approche fournit naturellement un nombre réduit de composantes. Ainsi, la partie supervisée pour la segmentation d'images texturées, où le choix des bonnes caractéristiques est issu d'un processus d'apprentissage, peut être évitée et nous obtenons une solution efficace pour une segmentation non-supervisée d'images texturées. Basé sur ces nouvelles composantes, nous proposons un processus de segmentation actif et adaptatif où les densités de probabilités à l'intérieur et a' l'extérieur du contour sont estimées de manière dynamique. Nous utilisons une technique rapide basée sur les ensembles de niveaux pour la mise en oeuvre. Pour finir, mous présentons des résultats de validation sur diverses images réelles et nous les comparons avec succès à ceux obtenus à partir d'études récentes.

Mots-clés : Théorie des courbes de niveaux, segmentation d'images texturées, segmentation adaptative d'images, diffusion non-linéaire, tenseur de structure.

1 Introduction

In recent time it has become very popular to use prior knowledge in the field of image segmentation. There are many techniques proposed in the literature, some of them including priors on the shape of objects [12, 6, 19], others, in the field of texture segmentation, use learned descriptive or generative models of textures in order to retrieve them during the segmentation process [16, 13]. With no doubt the usage of appropriate prior knowledge is very important for the ability to deal with difficult image scenes. However, there is very few work so far on how to obtain this prior knowledge automatically. So far segmentation algorithms with the ability to handle difficult image scenes need a supervised initialization step, where they are told the right segmentations of a training set of images. It would be preferable to extract the prior knowledge automatically from simpler image scenes, in order to use it in succeeding scenes that are more difficult to deal with. This requires a powerful image segmentation method that does not depend on prior knowledge, but can nevertheless handle the whole set of possible objects as long as the scene is not spoiled by clutter. Basically such a segmentation process splits into two parts: The first part is the acquisition of suitable features that are powerful enough to discriminate regions that a human observer would describe as different. The second part models the statistics of these features and searches for a segmentation that fits best to this model.
Especially in the area of texture segmentation the literature does not agree so far, how adequate features should be extracted from the image. Mostly Gabor filters are used [9, 16, 22, 21] but for example also the parameters of Markov Random Fields [7] are quite popular. Gabor filters have the decisive drawback that they induce a lot of redundancy and thus lots of feature channels. This is not so much a problem for supervised segmentation, as in this case statistics can be used to reduce the dimension of feature space, yet for unsupervised segmentation it is. An interesting work that helps to solve this problem is that of Bigün et al. [4]. They used the structure tensor in order to discriminate between textures. The advantage of their method is that the structure tensor yields only three feature channels for each scale. However, the Gaussian smoothing used for the structure tensor dislocates the edges in feature space leading to inaccurate segmentation results. Brox and Weickert [5] proposed a nonlinear structure tensor based on nonlinear matrix-valued diffusion that is able to tackle this problem. A very similar outcome is achieved by the approach of van den Boomgaard and van de Weijer [24], who applied robust statistics to orientation estimation. Our approach is based on that in [5], yet we use a better adapted diffusion technique that also brings us back to vector-valued diffusion. Sagiv et al. [21] also apply vector-valued PDEs in order to smooth their feature channels. However, their

feature extraction method based on Gabor filters is completely different from ours. In Section 4 we will compare the two approaches to each other.

Once the correct features are extracted, a segmentation process that accounts for these informations must be defined. Recently, segmentation methods based on level sets and variational formulations have been able to integrate different cues like boundary information [14], region information [16] and shape prior [12, 6, 19]. The use of level set functions to represent evolving curves gives many good properties to the segmentation process: the curve is represented implicitly, topological changes are *naturally* possible, it can be used in any dimension and efficient techniques for numerical implementation exist. Thus, we decided to define a method based on active contours using the level set representation. One key point of our method is the unsupervised aspect: the information must be incorporated in the segmentation process without adding particular knowledge on the image. For this purpose, an adaptative and active segmentation is proposed, following the idea of [18]. The minimization of the proposed energy gives the maximum a posteriori segmentation under some hypotheses which can be made in our case.

The remainder of this paper is organized as follows: In the next section we derive our feature extraction method from the structure tensor and nonlinear diffusion. Section 3 then deals with the adaptive segmentation based on the extracted features. In Section 4 we show some results and compare them to previous approaches. The paper is concluded by a brief summary.

2 Feature Extraction

Our approach to extract the features is based on the classical *structure tensor* [4, 8, 25]

$$J_\rho = K_\rho * (\nabla I \nabla I^\top) = \begin{pmatrix} K_\rho * I_x^2 & K_\rho * I_x I_y \\ K_\rho * I_x I_y & K_\rho * I_y^2 \end{pmatrix} \tag{1}$$

where K_ρ is a Gaussian kernel with standard deviation ρ and subscripts denote partial derivatives. Obviously the structure tensor yields three feature channels for each scale. In order to keep things simple, we will only consider one scale in this paper. One could incorporate further scales by adding additional feature channels very easily. Comparing the number of features obtained by the structure tensor to that of Gabor filters reveals that the degree of freedom for the orientation known from Gabor filters is replaced by the smoothed versions of the image derivatives. It should be noted that the image derivatives include the whole orientation information, so the components of the structure tensor are as powerful for the discrimination of different

textures as a whole set of Gabor filters for a fixed scale.
The major problem of the classic structure tensor is the dislocation of edges due to the smoothing with Gaussian kernels. This leads to inaccurate segmentation results near region boundaries. The basic idea in [5] to address this problem is the replacement of the Gaussian smoothing by nonlinear diffusion. We stick to this idea but enhance the technique for its application to texture segmentation.

Nonlinear diffusion is based on the early work of Perona and Malik [17]. The main idea is to reduce the smoothing in the presence of edges. The resulting diffusion equation is

$$\partial_t u = \operatorname{div}\left(g(|\nabla u|)\nabla u\right) \tag{2}$$

with $u(t=0)$ being the image I and g a decreasing function.
So far this equation can only be used with scalar-valued data like a gray value image. Gerig et al. [10] introduced a version of nonlinear diffusion for vector-valued data

$$\partial_t u_i = \operatorname{div}\left(g\left(\sum_{k=1}^{N} |\nabla u_k|^2\right)\nabla u_i\right) \qquad \forall i \tag{3}$$

where u_i is an evolving vector channel and N the total number of vector channels. Note that in this approach all channels are coupled by a joint diffusivity, so an edge in one channel also inhibits smoothing in the others.
If we regard the components of a matrix as components of a vector, what is reasonable, since the Frobenius norm of a matrix equals the Euclidean norm of the resulting vector, it is possible to diffuse a matrix, such as the structure tensor, with the above-mentioned scheme. This yields a version of a nonlinear structure tensor that is a bit different from that mentioned in [5]. The matrix-valued diffusion in fact equals a scheme proposed by Tschumperlé and Deriche [23]. Note that according to Weickert and Brox [26] the coupling of the channels ensures the preservation of semipositive definiteness.

A rather critical issue is the appropriate choice of the diffusivity function g. For the application to texture features *TV flow* [20, 2, 15] seems to suit very well, since it removes oscillations and leads to piecewise constant results. This is very important, because the structure tensor contains first derivatives, which have very local responses. The task of the smoothing process is actually to close the areas between these local phenomena while preserving the important edges. This is exactly what TV flow does. Furthermore, it has the nice property of not causing any additional parameters. Since TV flow leads to numerical problems when the gradient gets close to zero, it is necessary to circumvent this case. This is mostly done by adding a small positive constant ϵ to the gradient magnitude.

$$g(|\nabla u|) = \frac{1}{|\nabla u| + \epsilon} \tag{4}$$

We deviate a little from the actual structure tensor by adding the image gray value, which is certainly a very important feature, to the feature vector. Furthermore, the channel based on $I_x I_y$ appears only once and not twice, like in the structure tensor. However, note that, although it seems as if its information was already present in the other two channels of the structure tensor, it could not be neglected, because the sign was lost by squaring the derivatives. So, this channel is very important to ensure rotation invariance.
Finally, our features are computed by applying Eq.3 with initial conditions $u_1 = I$, $u_2 = I_x^2$, $u_3 = I_y^2$, $u_4 = I_x I_y$ and the diffusivity function $g(s) = 1/s$.
For implementation we apply the AOS scheme, proposed in [27], that allows efficient computation of TV flow also for small ϵ. For ϵ in the area of 0.001, where the approximation of TV flow is much better than for larger ϵ causing less blurring effects, the AOS scheme is around 4 orders of magnitude faster than a simple explicit scheme.

3 Adaptive segmentation based on these features

3.1 Variational Formulation

In this article, we restrict our study to the case of the segmentation into two textured regions. Basically, it includes all the images where only one object is to be segmented. Following [16], the image segmentation can be found by maximizing the *a posteriori* partitioning probability $p(\mathcal{P}(\Omega)|I)$ where $\mathcal{P}(\Omega) = \{\Omega_1, \Omega_2\}$ is a partition of the image domain Ω. Instead of the original image I, we use the vector-valued image $u = (u_1, ..., u_4)$ obtained by smoothing $(I, I_x^2, I_y^2, I_x I_y)$ with Eq.3. It has been shown in [16] that such an optimization is equivalent to an energy minimization. Two hypotheses must be made: all the partitions are equally possible and the pixels within each region are independent. Let $p_1(u(x))$ and $p_2(u(x))$ be the probability density functions for the value $u(x)$ to be in Ω_1 and Ω_2 respectively. Let $\partial\Omega$ be the boundary between Ω_1 and Ω_2, the segmentation can be found by minimizing the following energy:

$$E(\Omega_1, \Omega_2) = -\int_{\Omega_1} \log p_1(u(x))\, dx - \int_{\Omega_2} \log p_2(u(x))\, dx \tag{5}$$

The main challenges are to define a family of pdf to modelize each region that can be general enough to capture a wide range of images and to find the minimum of the energy. As we will see, too many parameters in the pdf approximation can trap our variational approach in local minima. This will push us to use a simplified modelization.

3.2 Gaussian Approximation for all the channels

First, a general Gaussian approximation is used to model the information for each region for all the four channels. Since the image u is vector-valued, we have to deal with covariance matrices. Let $\{\mu_1, \Sigma_1\}$ and $\{\mu_2, \Sigma_2\}$ be the vectors'means and the covariance matrices of the gaussian approximation in Ω_1 and Ω_2. The probability of $u(x)$ to be in Ω_i is:

$$p_i(u(x)) = \frac{1}{(2\pi)^2 |\Sigma_i|^{1/2}} e^{-\frac{1}{2}(u(x)-\mu_i)^T \Sigma_i^{-1} (u(x)-\mu_i)} \tag{6}$$

Therefore, the vectors'mean and convariance matrix of each region are additional unkown parameters that must be introduced in the energy (5).
Before minimizing this energy, we introduce the level set representation. The level set function ϕ is defined as:

$$\begin{cases} \phi(x) = \mathcal{D}(x, \partial\Omega), & \text{if } x \in \Omega_1 \\ \phi(x) = -\mathcal{D}(x, \partial\Omega), & \text{if } x \in \Omega_2 \end{cases} \tag{7}$$

Furthermore, let $H_\epsilon(z)$ and $\delta_\epsilon(z)$ be regularized versions of the Heaviside and Dirac functions. Then, the energy (5) can be minimized with respect to the whole set of parameters $\{\partial\Omega, \mu_1, \mu_2, \Sigma_1, \Sigma_2\}$ using the following evolution equation (see [18] for details):

$$\phi_t(x) = \delta_\epsilon(\phi(x)) \left(\log p_1(u(x)) - \log p_2(u(x))\right) \tag{8}$$

while the Gaussian parameters are updated at each iteration following:

$$\begin{cases} \mu_i(\phi) = \int_\Omega u(x)\chi_i dx / \int_\Omega \chi_i dx \\ \Sigma_i(\phi) = \int_\Omega (\mu_i - u(x))(\mu_i - u(x))^T \chi_i dx / \int_\Omega \chi_i dx \end{cases} \tag{9}$$

$$\text{with} \begin{cases} \chi_1(z) = H_\epsilon(z) \\ \chi_2(z) = 1 - H_\epsilon(z) \end{cases}$$

A regularization constraint on the length of $\partial\Omega$ can be added, yielding the following final evolution equation for ϕ:

$$\begin{aligned} \phi_t(x) = \delta_\epsilon(\phi(x)) \quad & (\nu \operatorname{div}(\nabla\phi/|\nabla\phi|) \\ & + \log p_1(u(x)) - \log p_2(u(x))) \end{aligned} \tag{10}$$

Considering full covariance matrices leads to lots of unknown parameters. This can result in multiple local minima and makes the energy minimization quite difficult. Such a situation is shown in Figure 1a where the curve stops before capturing one leg of the zebra. If we further analyse the channels depicted in Figure 2, we can see that the information included in each channel is not that much correlated. So, making the hypothesis that the channels are not correlated, the pdf $p_i(u(x))$ for $i = \{1, 2\}$ can be estimated using the joint density probability of each component:

$$p_i(u(x)) = \Pi_{k=1}^{4} p_{k,i}(u_k(x)) \tag{11}$$

This is equivalent to consider a diagonal covariance matrix. With this new approximation the energy has only 8 unknown statistical parameters for each region (4 for the vectors'mean and 4 for the covariance matrix) instead of 14 when a full covariance matrix is considered. Hence, we obtain the favored result as shown in Figure 1b.

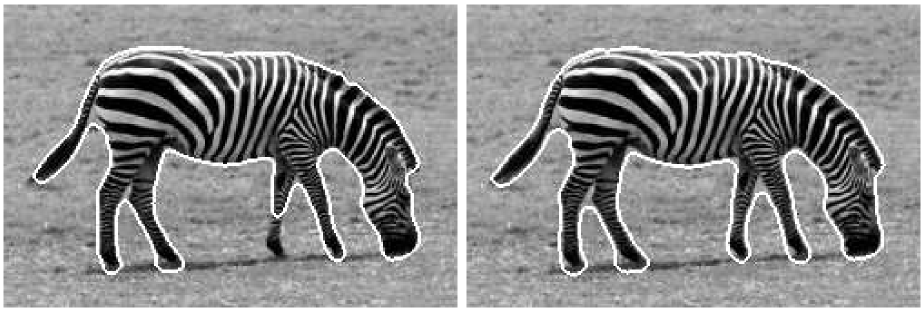

Figure 1: (a) LEFT: Result with full Σ. (b) RIGHT: Result with diagonal Σ.

Figure 2: Feature channels $u_1, .., u_4$.

3.3 Non-Parametric Approximation for $\mathbf{u_1}$

The Gaussian approximation for the first channel, which corresponds to smoothed version of the gray value image I (see the first feature in Figure 2) does not seem to suit very well. So for this channel, an estimation of the pdf based on the histogram should be preferred. Since we make the assumption that different channels are not correlated, the pdf in each channel can be estimated using different approaches. As in [11], we propose to use a continous version of the Parzen density for the first channel. The probability of $u_1(x)$ to be in Ω_i is given by:

$$p_{1,i}(u_1(x)) = \frac{1}{|\Omega_i|} \int_{\Omega_i} g_\sigma(u_1(x) - u_1(\hat{x}))\, d\hat{x} \tag{12}$$

where the Gaussian kernel is $g_\sigma(z) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{z^2}{2\sigma^2}}$

Using the shape derivative tool in [3], we now differenciate the following functional:

$$F(\Omega_i) = \int_{\Omega_i} \log\left(\frac{1}{|\Omega_i|} \int_{\Omega_i} g_\sigma(u_1(x) - u_1(\hat{x}))\, d\hat{x} \right) dx \tag{13}$$

The shape derivative of this functional (see annex and [3] for details) is:

$$
\begin{aligned}
< F'(\Omega_i), V >= \quad & -\int_{\Omega_i}\Big(\log p_{1,i}(u(x)) \\
& + \tfrac{1}{|\Omega_i|}\int_{\Omega_i}\tfrac{g_\sigma(u_1(x)-u_1(\hat{x}))}{p_{1,i}(\hat{x})}\,d\hat{x}\Big)(V.N)da(x)
\end{aligned}
\tag{14}
$$

Using the level set representation, it gives us a new evolution equation for ϕ:

$$
\begin{aligned}
\phi_t(x) = \delta_\epsilon(\phi)\underbrace{\left(\nu\,\mathrm{div}(\nabla\phi/|\nabla\phi|) + \sum_{k=1}^{4}\log\frac{p_{k,1}(u(x))}{p_{k,2}(u(x))}\right)}_{\text{usual term}} \\
\left.\begin{aligned}
&+\delta_\epsilon(\phi)\Big(\tfrac{1}{|\Omega_1|}\int_\Omega H_\epsilon(\phi)\tfrac{g_\sigma(u_1(x)-u_1(\hat{x}))}{p_{1,1}(u(\hat{x}))}\,d\hat{x} \\
&-\tfrac{1}{|\Omega_2|}\int_\Omega(1-H_\epsilon(\phi))\tfrac{g_\sigma(u_1(x)-u_1(\hat{x}))}{p_{1,2}(u(\hat{x}))}\,d\hat{x}\Big)
\end{aligned}\right\}\text{additional term}
\end{aligned}
\tag{15}
$$

while Gaussian parameters for the channels u_2, u_3 and u_4 are updated at each iteration according to (9). In this new evolution equation an additional term appears due to the first channel. Regarding the other terms of the equation, they are the same as the ones obtained with gaussian approximation for all channels when a diagonal covariance matrix is considered.
By using this non-parametric representation to approximate the first channel information, experimental results are improved a lot, as can be seen in Figure 3. Actually, the parametric estimation for the structure channels is very robust and when it is combined with gray level information, the method can deal with a much larger range of images, in particular with low-textured images.

3.4 Implementation remarks

To implement the evolution equation (15), we used an explicit scheme in time and the curvature was estimated using the classical centered finite difference. The level set was updated only in a small narrow band around its zero crossing. A very tiny narrow band is sufficient (we used a width of 12) since only local information is important. Therefore, the size of the active area was relatively small, and it was possible to reinitialize the level set function to the distance function at each iteration in reasonable time. We used the reinitialization method described in [1] which limits the displacements of the zero level. For all the examples shown in the next section, the same initialization and the same parameters were used. Small circles were used as initialization because even if the active region covers almost the whole image at

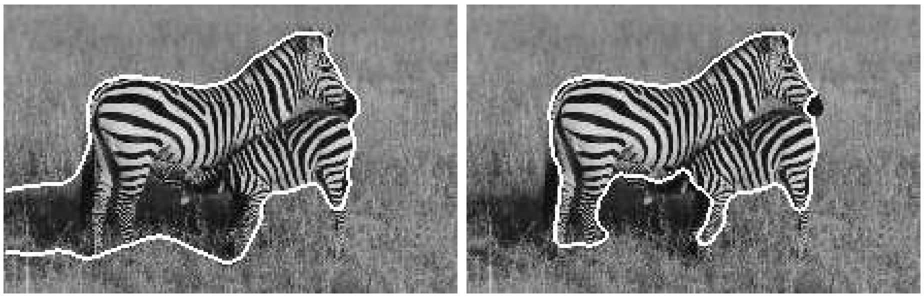

Figure 3: (a) LEFT: Result with Gaussian approximation for each channel. (b) RIGHT: Result with non-parametric approximation for the 1st channel.

the beginning, the final solution is reached after only few iterations (less than 30 iterations are mostly enough). But other initializations have also been tested and the method seems very robust by giving the same result with different initializations. Results from different initializations are shown in Figure 10.
Two parameters must be set, the regularization weight ν and the parzen window size σ, we set it once and for all: $\nu = 0.5$ and $\sigma = 5$. We were able to get good results on a wide range of textured images with these parameters.
Since we use an explicit scheme and so a small timestep, the second order terms in (15) are negligibles and can be omitted. With this approximation the curve evolution takes around ten seconds on 250x200 images and standard hardware. Also the feature extraction is quite efficient since we use the AOS scheme. The computation time of this part is also around ten seconds.

4 Results

We tested the performance of our method with synthetic as well as real test images. The synthetic test images are composed of textures from the Brodatz texture database. The real images are from papers on texture segmentation that were published earlier, so they allow direct comparison.
Figure 4a reveals our method to work quite fine with common textured images. Note the average gray value between the two textured regions not to be too much different, and there is also no dominant orientation for the background texture. Also note that the correct way of smoothing the structure tensor (using Eq.3) is very important to get such good results. Figure 4b shows that the segmentation fails completely if the

smoothing is omitted. Actually the smoothing is responsible for the fact that the three feature channels based on the nonlinear structure tensor are able to outperform a whole set of Gabor features. However, Figure 4c also shows the limitation that was already mentioned in Section 2: If two textures can only be distinguished due to their scale, there is no possibility for the structure tensor of only one scale to separate the regions. Note that this problem could be fixed very easily by adding the features of a structure tensor of different scale. However, it is also interesting that one has to construct very special cases to make the structure tensor of only one scale to fail completely.

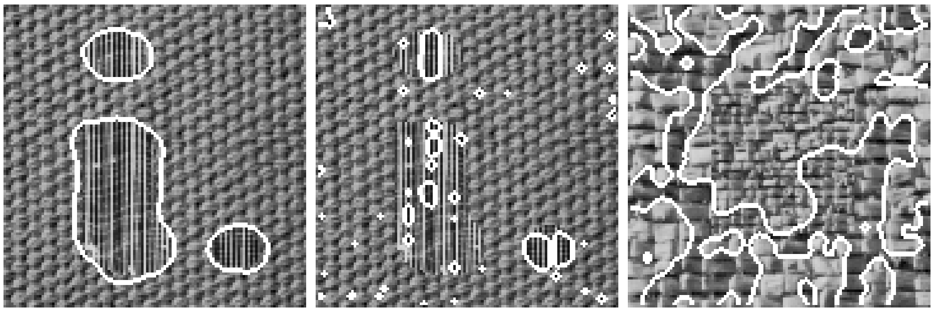

Figure 4: (a) LEFT: Result for a synthetic test image with smoothed features using Eq.3. (b) CENTER: Segmentation fails if the feature channels are not smoothed. (c) RIGHT: Segmentation fails if texture only differs in scale.

More interesting than the synthetic images are the real images. Figures 5 and 6 prove our method to be fully competitive to recent approaches published in [11] and [22] (these results as well as the results extracted from [16] and [21] are the ones shown in the original articles). Considering the results from the method in [21], where they used smoothed features based on Gabor filters, our method uses less feature channels and compares favourably. Figures 8 and 9 show that we can even improve the good results obtained in [16] where a supervised scheme was used. Finally, Figure 11 illustrates the capabilities of our approach on an other set of synthetic and natural images.

So obviously, our features based on the nonlinear structure tensor are very powerful in discriminating different textured regions, and our dynamic modelling of the active regions is able to incorporate the features in a way that enables the algorithm to cope with real textured images. We also want to stress that our method is almost

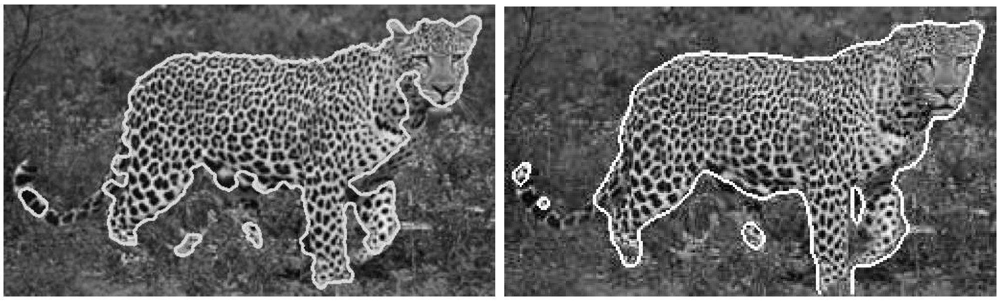

Figure 5: (a) TOP: Result from [11]. (b) BOTTOM: Our result.

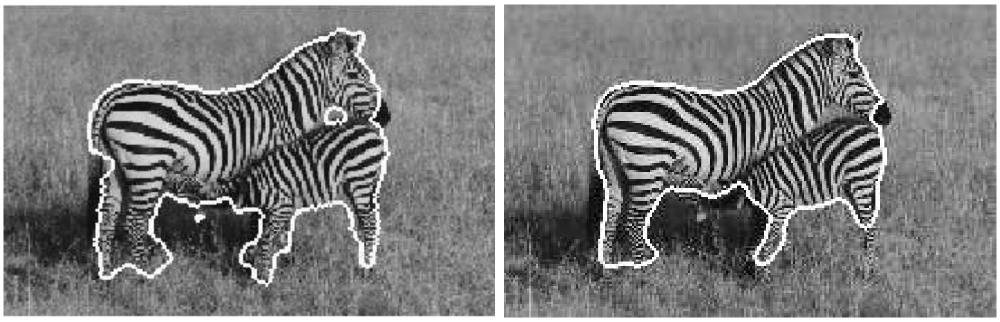

Figure 6: (a) TOP: Result from [22]. (b) BOTTOM: Our result.

free of parameters. The parameters that appear for the nonlinear diffusion and the active contour are very robust and can be set to fixed values. All our results have been computed with the same parameters. This property is very important for an unsupervised approach, because we think an approach can actually not be called unsupervised, if for each image someone has to figure out the right parameters first. We are also in the process to apply our method to much more images.

5 Conclusions

In this paper two novel ideas were proposed and applied in order to cope with unsupervised texture segmentation. First, the idea of using the structure tensor for feature extraction was emphasized and a nonlinear version was shown to be compet-

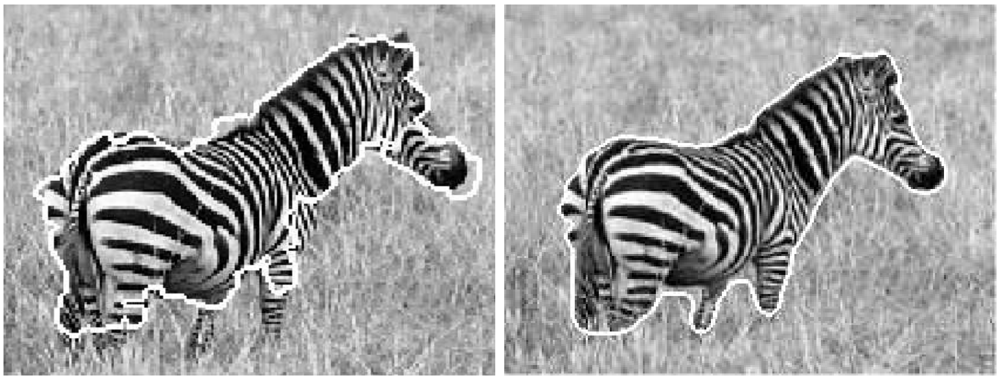

Figure 7: (a) LEFT: Result from [21]. (b) RIGHT: Our result.

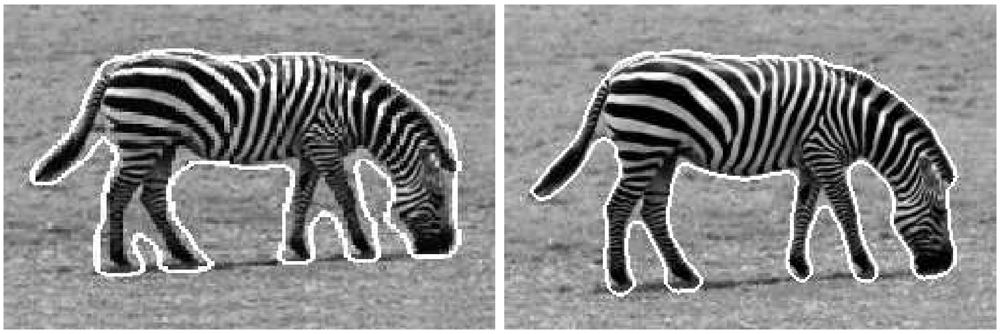

Figure 8: (a) TOP: Result from [16]. (b) BOTTOM: Our result.

itive or even superior to a whole set of Gabor features. Second, we have proposed a new segmentation model based on a variational formulation and on the level representation. Starting from a general multi-dimensional Gaussian to approximate region information of the feature channels, we have simplified and specialized the model so as to get a robust segmentation process. The robustness of this approach allows to fix all appearing parameters, so our technique is completely parameter free. The comparison of our method to some recent approaches was very convincing. It was possible to reproduce or even improve the former results. Of course, there still exist several examples where our method is not appropriate. One problem is caused by textures that only differ in their scale. Another limitation is given by the fact

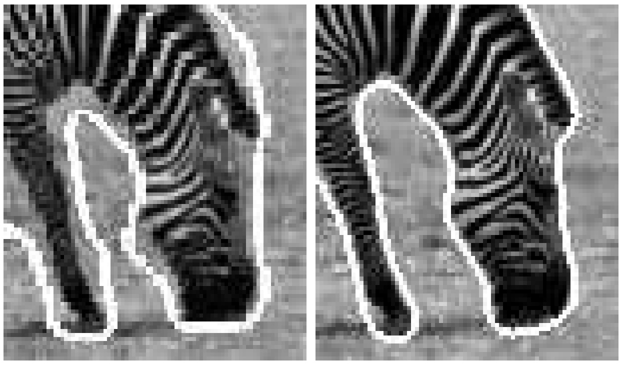

Figure 9: Zoom on result (a) LEFT: Result from [16]. (b) RIGHT: Our result.

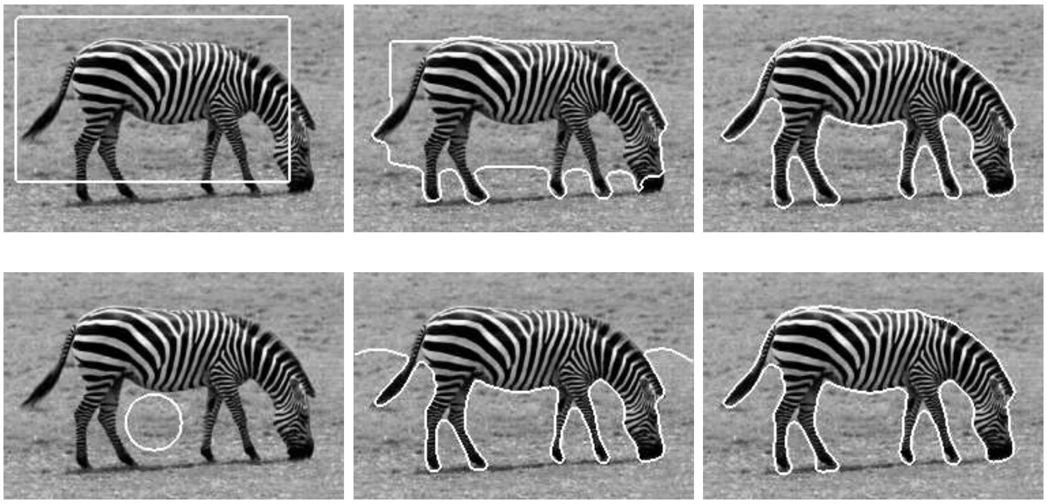

Figure 10: (a) TOP: Curve evolution with first initialisation (wide recangle). (b) BOTTOM: Curve evolution with second initialisation (one small circle).

that we only consider two regions. Our future research will focus on resolving these problems.

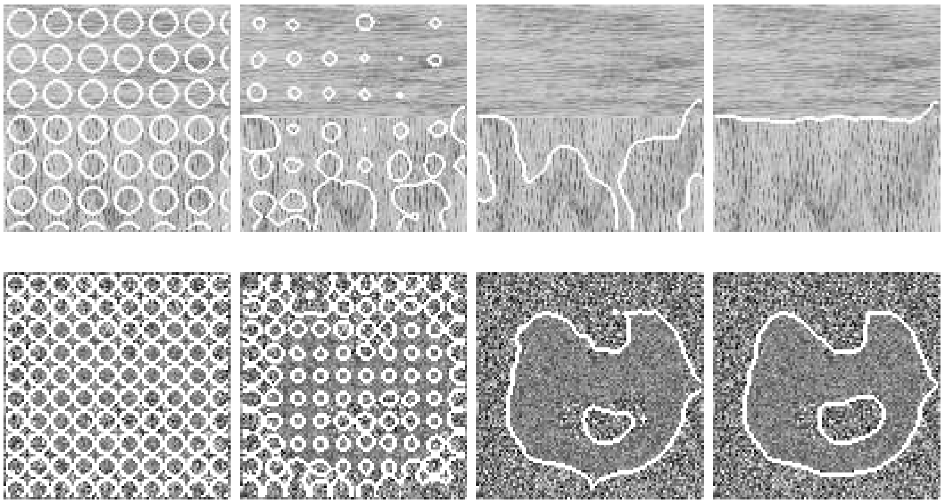

Figure 11: (a) TOP: Curve evolution for a synthetic image composed where the regions have same means but different variances. (b) BOTTOM: Curve evolution for the *wood* image.

A Derivation details

To simplify the notations we note Ω the (bounded and open) domain of integration which is moving. Actually Ω is either Ω_1 or Ω_2, and Γ will be the boundary of Ω. Using the shape derivative tool [3], we want to differenciate the functional:

$$F(\Omega) = \int_\Omega \log\left(\frac{1}{|\Omega|}\int_\Omega g_\sigma(u_1(x) - u_1(\hat{x}))\,d\hat{x}\right)\,dx \tag{16}$$

We introduce some notations:

$$\begin{array}{l} F(\Omega) = \int_\Omega f(x,\Omega)\,dx = \int_\Omega \log\frac{G_1(x,\Omega)}{G_2(\Omega)}\,dx \\ \text{with } G_1(x,\Omega) = \int_\Omega g_\sigma(u_1(x) - u_1(\hat{x}))\,d\hat{x} \\ \text{and } G_2(\Omega) = |\Omega| = \int_\Omega d\hat{x} \end{array}$$

The Gâteaux derivative of $F(\Omega)$ in the direction of V (a vector field) is:

$$\begin{array}{rl} < F'(\Omega), V >== & \int_\Omega f_s(x,\Omega,V)dx \\ & - \int_\Gamma f(x,\Omega)(V(x).N(x))da(x) \end{array}$$

where $f_s(x, \Omega, V)$ is the shape derivative of $f(x, \Omega)$. This can be expressed as:

$$f_s(x, \Omega, V) = f_{G_1} < G_1'(x, \Omega), V > + f_{G_2} < G_2'(\Omega), V >$$

Each term can be calculated:

$$\begin{array}{l} f_{G_1} = 1/G_1, \quad f_{G_2} = -1/G_2 \\ < G_1'(x, \Omega), V >= - \int_\Gamma g_\sigma(u_1(x) - u_1(\hat{x}))(V(\hat{x}).N(\hat{x}))\, d\hat{x} \\ < G_2'(\Omega), V >= - \int_\Gamma (V(\hat{x}).N(\hat{x}))\, d\hat{x} \end{array}$$

Putting everything together we get the following shape derivative:

$$\begin{aligned} < F'(\Omega), V > \quad &= - \int_\Gamma \Bigg(\log p_1(u(x)) \\ &\quad + \tfrac{1}{|\Omega|} \int_\Omega \tfrac{g_\sigma(u_1(x) - u_1(\hat{x}))}{p_1(u(\hat{x}))}\, d\hat{x} \Bigg) (V(x).N(x))\, dx \end{aligned} \tag{17}$$

Note that this result could be related to the one obtained in [11] through a different approach.

References

[1] D. Adalsteinsson and J. Sethian. The fast construction of extension velocities in level set methods. *Journal of Computational Physics*, 148:2–22, 1998.

[2] F. Andreu, V. Caselles, J. I. Diaz, and J. M. Mazón. Qualitative properties of the total variation flow. *Journal of Functional Analysis*, 2001. To appear.

[3] G. Aubert, M. Barlaud, O. Faugeras, and S. Jehan-Besson. Image segmentation using active contours: calculus of variations of shape gradients? *Research Report, INRIA*, July 2002.

[4] J. Bigün, G. H. Granlund, and J. Wiklund. Multidimensional orientation estimation with applications to texture analysis and optical flow. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(8):775–790, August 1991.

[5] T. Brox and J. Weickert. Nonlinear matrix diffusion for optic flow estimation. In *Proc. 24th DAGM Symposium*, volume 2449 of *Lecture Notes in Computer Science*, pages 446–453, Zürich, Switzerland, September 2002. Springer.

[6] D. Cremers, T. Kohlberger, and C. Schnörr. Nonlinear shape statistics in mumford-shah based segmentation. In *Proc. 7th European Conf. on Computer Vision*, volume 2351 of *Lecture Notes in Computer Science*, pages 93–108, Copenhagen, June 2002. Springer.

[7] G. R. Cross and A. K. Jain. Markov random field texture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 5:25–39, 1983.

[8] W. Förstner and E. Gülch. A fast operator for detection and precise location of distinct points, corners and centres of circular features. In *Proc. ISPRS Intercommission Conference on Fast Processing of Photogrammetric Data*, pages 281–305, Interlaken, Switzerland, June 1987.

[9] D. Gabor. Theory of communication. *J. IEEE*, 93:429–459, 1946.

[10] G. Gerig, O. Kübler, R. Kikinis, and F. A. Jolesz. Nonlinear anisotropic filtering of MRI data. *IEEE Transactions on Medical Imaging*, 11:221–232, 1992.

[11] Junmo Kim, John Fisher, Anthony Yezzi, Mujdat Cetin, and Alan Willsky. Nonparametric methods for image segmentation using information theory and curve evolution. In *IEEE Internation Conference on Image Processing*, Rochester, NY, setpember 2002.

[12] M. Leventon, E. Grimson, and O. Faugeras. Statistical Shape Influence in Geodesic Active Contours. In *Proceedings of the International Conference on Computer Vision and Pattern Recognition*, pages 316–323, Hilton Head Island, South Carolina, June 2000. IEEE Computer Society.

[13] J. Malik, S. Belongie, T. Leung, and J. Shi. Contour and texture analysis for image segmentation. *The International Journal of Computer Vision*, 43(1):7–27, June 2001.

[14] R. Malladi, J.A. Sethian, and B.C. Vemuri. Shape modeling with front propagation: A level set approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17(2):158–175, February 1995.

[15] Stanley Osher, Andres Sole, and Luminita Vese. Image decomposition and restoration using total variation minimization and the h^{-1} norm. *UCLA C.A.M. Report 02-57*, October 2002.

[16] Nikos Paragios and Rachid Deriche. Geodesic active regions: A new paradigm to deal with frame partition problems in computer vision. *Journal of Visual Communication and Image Representation*, March/June 2002.

[17] P. Perona and J. Malik. Scale space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12:629–639, 1990.

[18] M. Rousson and R. Deriche. A variational framework for active and adaptative segmentation of vector valued images. In *Proc. IEEE Workshop on Motion and Video Computing*, Orlando, Florida, December 2002.

[19] Mikael Rousson and Nikos Paragios. Shape priors for level set representations. In *ECCV*, volume 2, pages 78–92, Copenhagen, Denmark, May 2002.

[20] L. I. Rudin, S. Osher, and E. Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D*, 60:259–268, 1992.

[21] C. Sagiv, N. A. Sochen, and Y. Y. Zeevi. Texture segmentation via a diffusion-segmentation scheme in the gabor feature space. In *Proc. Texture 2002, 2nd International Workshop on Texture Analysis and Synthesis*, Copenhagen, June 2002.

[22] B. Sandberg, T. Chan, and L. Vese. A level-set and gabor-based activ contour algorithm for segmenting textured images. Technical Report 39, Math. Dept. UCLA, Los Angeles, USA, July 2002.

[23] D. Tschumperlé and R. Deriche. Diffusion tensor regularization with contraints preservation. In *Proc. 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 1, pages 948–953, Kauai, HI, December 2001. IEEE Computer Society Press.

[24] R. van den Boomgaard and J. van de Weijer. Robust estimation of orientation for texture analysis. In *Proc. Texture 2002, 2nd International Workshop on Texture Analysis and Synthesis*, Copenhagen, June 2002.

[25] J. Weickert. *Anisotropic Diffusion in Image Processing.* Teubner, Stuttgart, 1998.

[26] J. Weickert and T. Brox. Diffusion and regularization of vector- and matrix-valued images. Technical Report 58, Department of Mathematics, Saarland University, Saarbrücken, Germany, March 2002.

[27] J. Weickert, B. M. ter Haar Romeny, and M. A. Viergever. Efficient and reliable schemes for nonlinear diffusion filtering. *IEEE Transactions on Image Processing*, 7(3):398–410, March 1998.

Unité de recherche INRIA Sophia Antipolis
2004, route des Lucioles - BP 93 - 06902 Sophia Antipolis Cedex (France)

Unité de recherche INRIA Lorraine : LORIA, Technopôle de Nancy-Brabois - Campus scientifique
615, rue du Jardin Botanique - BP 101 - 54602 Villers-lès-Nancy Cedex (France)
Unité de recherche INRIA Rennes : IRISA, Campus universitaire de Beaulieu - 35042 Rennes Cedex (France)
Unité de recherche INRIA Rhône-Alpes : 655, avenue de l'Europe - 38330 Montbonnot-St-Martin (France)
Unité de recherche INRIA Rocquencourt : Domaine de Voluceau - Rocquencourt - BP 105 - 78153 Le Chesnay Cedex (France)

Éditeur
INRIA - Domaine de Voluceau - Rocquencourt, BP 105 - 78153 Le Chesnay Cedex (France)
http://www.inria.fr
ISSN 0249-6399