A Multichannel Variational Model for Robust Image Segmentation under Noise

R. Romano

D. Vitulano [1]

Istituto per le Applicazioni del Calcolo "M. Picone"
C. N. R.
Viale del Policlinico 137
00161 Rome, Italy
Tel. +39–6–88470210, Fax +39–6–4404306
E-mail: {romano,vitulano}@iac.rm.cnr.it

ABSTRACT

This paper [2] presents a novel model for image segmentation under noise: *EVRIST* (Efficient Variational Representation for Image Segmentation Technique). In order to segment general images containing both quite smooth regions and textures, *EVRIST*, based on a hierarchical representation obtained by the classical *weak membrane*, utilizes two channels: one relative to the *mean* and another for the *edges density*. The results achieved on both manual compositions and on real images, show the high robustness to the noise along with a low computational effort.

Keywords: Image Segmentation, Variational Models, Textures

1 INTRODUCTION

One of the more interesting topics in image processing is represented by image segmentation because of its importance as preliminary phase in a lot of practical applications in various fields. Its task consists of splitting a given image in regions, in agreement with the human perception [Lovel92, Berge93].

That's why there are a lot of proposed approaches in literature , and, among them, an interesting class is based on a variational formulation (see for instance [Blake87]). Main features are multiscale detection and selective smoothing, i.e. elimination of the noise with a preservation of the discontinuities representing the information — that is not possible employing classical linear approaches. Contemporaneously to the techniques oriented to improve the performances on quite smooth images, the interest of the scientific community addressed to possible extensions of this formulation to images containing textures (see for instance [Lee92], [Koepf94] and Zhu96).

In [Tronc99] we proposed *WCRM* (Weak Continuity Representation based Model), which utilizes the line process of a classical *Weak Continuity* (*WC*) process for building a measure able to segment textures. In other words the edge points obtained by a classical WC process are utilized for an estimate of the textures coarseness. Thus, based on an earlier Rosenfeld's idea (see [Rosen75] and [Jain86]), where different textures are characterized by a different coarseness, *WCRM* is able to segment images containing only textures, with a high robustness to the noise [Tronc99].

In this paper we want to generalize $WCRM$ so that more than one feature, all robust under noise, will be used and make it able to discriminate among different regions lacking of a significant coarseness — i.e. the classical (black and white) chessboard. In particular, as regards Fig. 3 Top image, $WCRM$ would be able to segment

[1] Corresponding author

[2] This paper has been partially supported by the CNR Finalized Project: "Cultural Heritage" and "5% MURST".

the whole image bar the rightmost and topmost blocks — uniformly black and white ones.

Thus, these two aspects constitute the topic of this paper and, as regards the first one, we introduced a vectorial model, where each channel represents a feature. In order to not excessively complicate the model, two only channels have been used: one for the *edges density* and the other for the *mean* of the pixels' grey levels. The problem of discriminating among flat regions inside an image, can be solved with a suitable "filter" which discriminates between flat regions: it preserves the edges obtained by the first *WC* process, while textured regions are processed by the whole model. These features have been taken into account for designing the evolution of *WCRM*: *EVRIST* (Efficient Variational Representation based Image Segmentation Technique), which achieves very good results under noise from both an objective and a subjective point of view.

The organization of the paper is as follows. Section 2 presents a short review about the variational formulation based approaches as well as *WCRM*. This allows the reader to better understand Section 3 where *EVRIST* is presented. In Section 4, *EVRIST* has been performed on both manual compositions of Brodatz textures and real images showing textures of buildings of historical importance while some concluding remarks constitutes the topic of Section 5.

2 VARIATIONAL MODELS FOR IMAGE SEGMENTATION

Image restoration so as image segmentation under noise, considered as a solution of (see [Jain86] p. 268 and [Tebou98]) $f = gu + \eta$ (where f is an image under noise, u the same image without noise and g and η are two kinds of distortion, $g = Id$ i.e. Id is the identity operator so that we have only additive noise) is an ill-posed problem [Hadam23].

Following the variational formulation, in order to estimate the function u over a domain Ω (open and bounded set in R^2), starting from the data f, we have got to minimize the following functional:

$$E(u) = \underbrace{\int_\Omega (u-f)^2}_{\text{faithfulness to the data}} + \underbrace{\lambda^2 \int_\Omega \varphi(|\nabla(u)|)}_{\text{regularizing term}} \qquad (1)$$

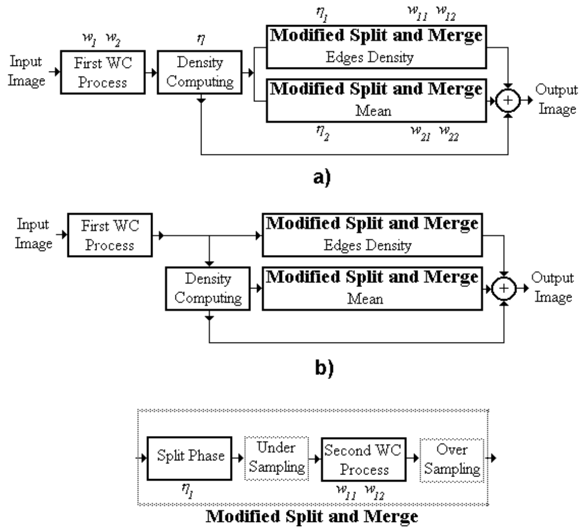

Figure 1: a) *EVRIST*'s block-scheme with the involved thresholds, b) the same after the link between η and η_1 (see text for details).

where $|\nabla(u)|$ represents the modulus of the gradient of u and λ is a weight factor balancing these two components (for a taxonomy of the choices of the function φ see [Tebou98]).

More in general, it can be shown [Tebou98, Lazar96, Chamb95] that the choice of a regularizing function convex and linear at infinity guarantees a unique solution in a Bounded Variation (BV) functions space. Nonetheless, if on one hand using a non convex φ allows us to get a selective smoothing, preserving the information contained in the "good" discontinuities, on the other represents an ill-posed problem, with a minimum only in the discrete version of the functional, where stability and uniqueness are still an open problem.

In this paper, we focus on the functional of Mumford-Shah [Mumfo89], using the minimization proposed by Blake and Zissermann [Blake87]. Other functionals are also suitable (es: Geman and Geman [Geman84]) for our model, but only under the constraint of an edge-preserving regularization.

The functional we consider (for $f \in R^2$ it is known as *weak membrane*) can be written as follows:

$$E(u,K) = \int_\Omega (u-f)^2 + w_1 \int_{\Omega/K} \nabla u \cdot \nabla u + w_2 \int_K d\sigma, \qquad (2)$$

where w_1 and w_2 are again weight factors providing a good compromise among these different components.

Thus, in order to minimize the functional, as non convex, *GNC* (Graduated Non Convex) algorithm can be performed: it consists in a gradual modification (more and more non convex) of a starting and convex functional. Though the procedure guarantees a local minimum, the results are generally good, employing a low computational time.

On a such formulation is based *WCRM*, whose description is contained indetail within the next section.

2.1 *WCRM*'S REVIEW

In this section we want to give a short review about *WCRM* in order to explain some (theoretical and implementation) details so that the following will be more clear for the reader.

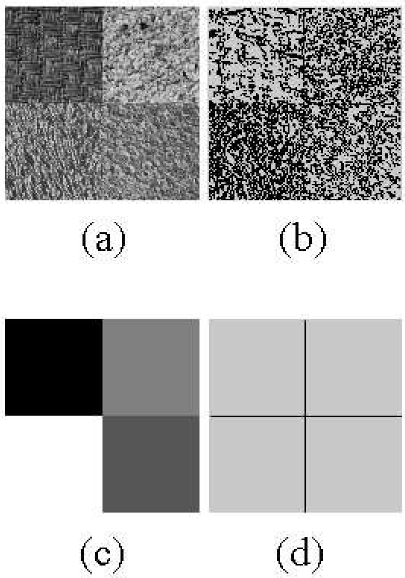

Figure 2: A Brodatz textures composition: a) Original, b) *WC*'s output, c) *Split phase* output (edges density case), d) final segmentation.

In order to better explain the effect of each phase of *WCRM* we consider the textures composition in Fig. 2.a. This latter is composed by four Brodatz textures. On this image we perform the first *WC* process. In practice, for an easy reproducibility we utilized the algorithm contained in [Blake87] pg. 157. The employed thresholds are w_1 and w_2 corresponding to λ and h_0 in [Blake87], i.e. scale factor and noise sensitivity. After this process, we have as output a binary image containing the edges computed (shown in Fig. 1.b): K set in (2). We have the value 1 in correspondence of edge points and zero elsewhere. On this image we can compute the classical Split phase. As criterion for discriminating among different regions (different textures), we utilized the edges density. In particular, we compute 1's number per unit area. This phase utilizes a threshold that we call η_1. It takes into account the difference of edges density between the father and his sons, using a quadtree partitioning. At the end of this phase we obtain uniform areas where for each pixel of a given uniform square block we have the value of the edge density of the whole uniform block. The results is shown in Fig. 1.c, where the different gray color of each one of the four blocks belonging to the composition is proportional to the edges density of the considered block. At this point we perform the last phase, i.e. the Merge phase. Here, instead of performing a classical Merge we utilize again a *WC* process so that we have a Modified Split and Merge. The task of this phase is to find the edges of Fig. 1.c. The thresholds w_1 and w_2 have the same meaning of the first *WC* process (the algorithm is the same). The final output is shown in Fig. 1.d.

WCRM corresponds to the only first channel (relative to the edges density) in Fig. 1.a), without the *Density Computing* phase and the *OR* final phase that will be explained later.

The achieved results have been very good, especially for segmentations under noise so that a generalization of this model, *EVRIST*, has been developed and will be presented in the next section.

3 ABOUT EVRIST

After a short review about some variational approaches for image segmentation along with *WCRM*, in this section we will show *EVRIST* (whose block scheme is shown in Fig. 1). It has been designed:

- to build a *WCRM* generalization able to segment generic images, i.e. containing textures and not;
- utilizing more features so that an improved segmentation under noise is possible — *EVRIST* is a vectorial version of *WCRM.*

More in detail, for a generic input image f on a domain $\Omega \in R^2$, the first step is the minimization of a *WC* process

$$E(u, K) = \int_\Omega (u - f)^2 +$$

$$w_1 \int_{\Omega/K} \nabla u \cdot \nabla u + w_2 \int_K d\sigma \qquad (3)$$

providing the set K and the recovered image u, on which we can define a generalized (in *WCRM* formulation this function was binary, i.e. $u(x,y)$ was replaced by 1) colored characteristic function $\tilde{\chi} : \Omega \to R$ as:

$$\tilde{\chi}(x,y) = \begin{cases} u(x,y) & \text{if } x,y \in K_{w_1,w_2} \\ 0 & \text{otherwise.} \end{cases} \qquad (4)$$

So, the image can be represented by their hierarchically most important points, i.e. the discontinuities, at a given scale level.

Using the symbol $\mathbf{P}(\tilde{\chi})$ for indicating a vectorial statistic — each component represents a statistic — defined on the matrix produced by $\tilde{\chi}$, we can define the *Split* phase as follows:

$$S(\mathbf{P}(\tilde{\chi}(x,y)), \mathbf{t}) = \mathbf{g}, \qquad (5)$$

where $\mathbf{t}$ is an array containing the used thresholds, while $\mathbf{g}$ represents the output images.

Thus, if we have N channels:

$$E(v, M_i) = \int_\Omega (v - \mathbf{g}_i)^2 + w_{i1} \int_{\Omega/M_i} \nabla v \cdot \nabla v + w_{i2} \int_{M_i} d\sigma \qquad 1 \le i \le N. \qquad (6)$$

On the three images coming from each channel, an *or* operation is performed in order to obtain the final segmentation.

Similarly to [Tronc99], for each channel, the decimation of the image g, ("Under Sampling" in Fig. 1.b.) whenever possible, allows a saving of computational effort in the latter *WC* process. Obviously, the output of the second *WC* process will have to be over sampled, in order to obtain the same size of the input image.

The problem of discriminating among "flat" regions characterized by different gray scale levels can be solved introducing a filter before the input of the channels as shown in Fig. 1.a: *Density Computing.* Its task consists, utilizing the first *WC* process output edges density, in distinguishing among textures and flat regions. A generic texture will have a lot of edges, that is a consistent coarseness, while, as regards flat regions, the edges density will be little or absent.

The implementation of *EVRIST* is very simple as it results in a vectorial version of *WCRM* where besides the edges density the mean of the pixels' grey levels has been employed. In addition, we have the third channel which is relative to those regions (non textured) that go directly to the *or* operator.

3.1 THE THRESHOLDS PROBLEM

EVRIST utilizes two only channels so that it involves $N_{par} = 9$ parameters to be tuned. Nonetheless, as we will see in the following, only four thresholds will be free. In Fig. 1.a) have been written the thresholds employed for each phase.

w_1 and w_2 of the first *WC*, since they represent respectively the scale level at which we want to study the input image and the threshold oriented to discard the noise (see [Blake87] p. 52), have to be free.

While η_i relative to the *Split* processes have to be free, representing a measure of uniformity of the textures, η is tied to η_1 by the following constraint: $\eta = \eta_1 \times 1$, where 1 is the highest density. In other words, a given region is considered flat if its density is not greater than η, that is, homogeneous with regions with density $= 0$ and then perfectly flat. It is evident that such a constraint leads the edges density split phase to not detect flat regions, in the sense that, η_1's working contains η. Hence the new model becomes as in Fig. 1.b.

As regards the other two *WC* processes, all their parameters can be fixed. In fact, w_{11} and w_{21} (the first index represents the number of the channel) can be set at 1 whenever the decimation/interpolation is performed [Tronc99], and also w_{12} and w_{22} can be set a priori as follows. Let us consider only one channel, for instance the *edges density* one, for the sake of simplicity. Generally, the output we obtain from the Split phase is an image containing uniform regions, each one containing pixels' set at the value of its density — see the foregoing Section. So w_{22} can be fixed by the following constraint: $w_{22} < \eta_2 \times mindens$ where $mindens$ represents the value of the lowermost density value computed inside the image. Summing up, after these considerations, the free thresholds are w_1 (scale level), w_2 (elimination of noise), η_1 (uniformity relative to edges density predicate) and η_2 (uniformity relative to mean predicate).

4 SOME EXPERIMENTAL RESULTS

EVRIST has shown very good performances on various images. For the sake of brevity, we present here two only examples along with some comparisons. These have been evaluated as good for understanding the potentialities of the proposed approach.

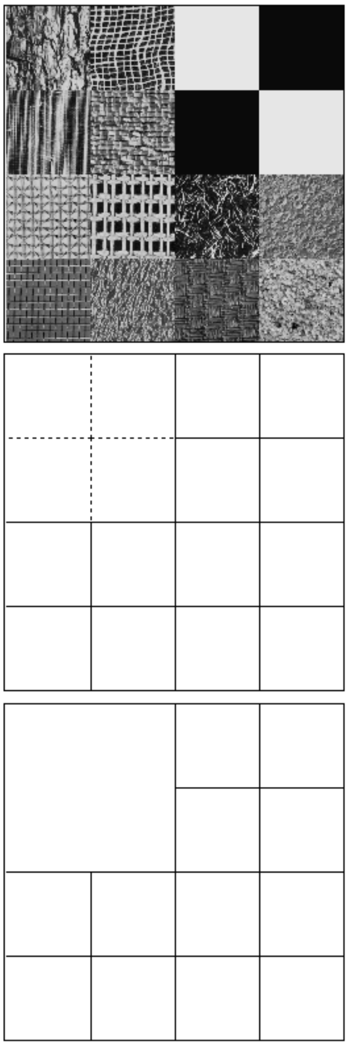

Figure 3: The textured chessboard where the black and white blocks are set at 10 and 240 (Top); $EVRIST$ segmentation (see text) at 90% (middle); $EVRIST$ segmentation at 91% (bottom).

The first image we deal with is the chessboard shown in Fig. 3 (Top). This case is very interesting, since it allows us, by means of the Haralick's metrics [Haral84], to get objective measures under noise on a relatively complicate image containing Brodatz textures [Broda66] along with black and white blocks.

$EVRIST$ achieves a correct segmentation both on the original chessboard ($w_1 = 2, w_2 = 25, \eta_1 = 40, \eta_2 = 22$) and on the same image with additive gaussian noise (zero mean) up to 90% (SNR=6.98 db) with $w_1 = 2, w_2 = 37, \eta_1 = 14, \eta_2 = 22$, as

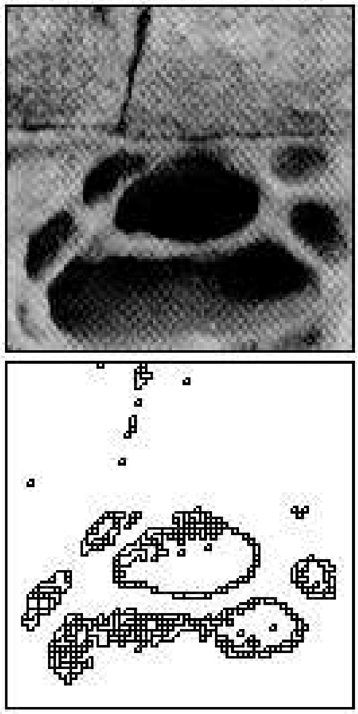

Figure 4: An example of segmentation of a pudding stone with its degradation: Top) original, bottom) a possible segmentation

shown in Fig 3 (Middle).

In this latter, noisy case, we can see that both channels are fundamental for getting good performance. In fact, the solid line shows the edges obtained using the mean channel while the dotted line is relative to the edge density one. The combination of both leads to a correct result. Moreover, it is worth pointing out that the filter discriminating between flat regions and textured ones is required only when the noise amplitude is low, otherwise the two channels above are enough. In fact, when the noise amplitude is high, some edges due to the noise are not discarded producing an artificial coarseness. Since we have a bias due to the gray level of the underlying information (without noise) we are still able to perform the segmentation. In other words, in $EVRIST$ there isn't the hypotesis that high amplitudes constitute the information, while low gradients are the noise.

From 91% onwards, $EVRIST$ becomes unstable, in the sense that is able to correctly segment the image only for some percentage of noise. In particular, at 91% of noise, the Haralick's values are $P(DTE|TE) = .83$ and $P(DTE|DE) = 1$, as shown in Fig. 3 (Bottom): some contours have been missing, but no spurious edge is detected, as $WCRM$ did [Tronc99].

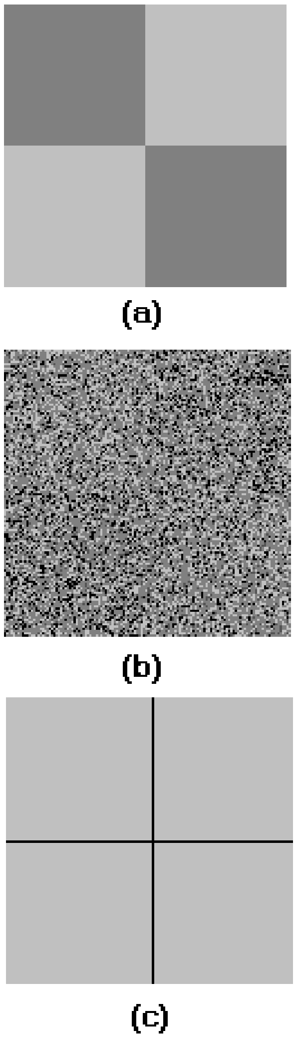

Figure 5: Black and white chessboard: a) Original, b) with 60% of gaussian noise, c) $EVRIST$ segmentation.

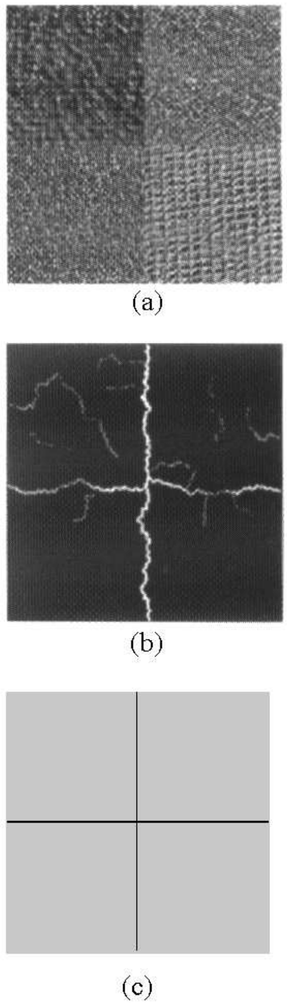

Figure 6: The textures composition proposed in [Liu00]: a) Original, b) their segmentation, c) our segmentation at 99% of noise.

Before presenting the second example, there are some considerations to be made. First, it is trivial to prove that $EVRIST$ is invariant to any bias of the adding noise so that the hypothesis of zero mean is not necessary. Second, though there are four thresholds to be tuned, their behavior is very simple. The scale level w_1 is fixed, w_2 obviously grows as the noise increases, while both η_1 and η_2 have a decreasing behavior, since the elimination of the noise edges leads to eliminate also the good edges [Tronc99].

The second example is a "natural" image showing a pudding stone medium grain with an evident degradation: the alveolization (Fig. 4, Top) along with its segmentation (Fig. 4, Bottom). The result shows a possible $EVRIST$'s segmentation on a given image without noise.

Summing up, the combination of the *weak* representation along with the couple of the two WC processes seems to constitute a very promising approach for image segmentation. Moreover, its robustness to the noise along with a reasonably low computational effort (five seconds are required to segment the chessboard at 90% of noise on a workstation Octane/SI R10000 175 MHz/1Mb cache) leads to consider $EVRIST$ as a good candidate for an extension of classical *weak membrane* approach to the textures.

Before concluding this section, it is possible to compute the upper-bound of $EVRIST$ and to do some comparisons. As regards "flat regions" (i.e. non textured ones) we built a chessboard composed by only four blocks (black=120 and white=130). On this image, it is possible to obtain a perfect segmentation up to 60% of noise, as in Fig. 5. As regards textures, we considered the composition in Fig. 6 (Top), presented in [Liu00].

In the same Figure, we find the segmentation proposed in [Liu00] without noise(Medium), and our result (Bottom) at 99% of noise. These results are very promising and lead us to consider *EVRIST* as a candidate for the *Weak Membrane* extension, able to segment also textures.

5 CONCLUDING REMARKS

In this paper, we have presented *EVRIST*, a model able to segment images using *mean* of the pixels' grey levels and *edges density* as features of a suitable representation of the input image. It represents an evolution of *WCRM* which represented an efficacious technique for segmenting images containing only textures under noise. As *WCRM* did, *EVRIST* presents a high robustness to the noise, a low computational time along with an easy implementation. In conclusion, we outline that the weak continuity representation seems to open a promising way to be better investigated. Nonetheless, the future research involves a more efficacious criterion for looking for the edges. In other words, the modified Split and Merge, for segmentation under noise, is not efficacious for complicate shapes. This is due to the fact that, because of the quadtree partition, in corrispondence of some shapes the Split phase produces very little regions, where optimal detection under noise is not possible (see [Zhu96] for similar considerations using the uncertainty principle). So, the future work will consist of producing a more efficacious model for this aim preserving the same hierarchical representation.

REFERENCES

[Berge93] J. R. Bergen, B. Julesz, *Rapid Discrimination of Visual Patterns*, IEEE Trans. on System Man and Cybernetics, Vol. 13, pp. 857-863, 1993.

[Blake87] A. Blake and A. Zissermann, *Visual Reconstruction*. MIT Press, 1987.

[Broda66] P. Brodatz, Textures: A Photographic Album for Artists and Designers, New York: Dover, 1966.

[Chamb95] A. Chambolle and P. L. Lions, Image Recovery via Total Variation Minimization and Related Problems, Research note95-09, 1995, CEREMADE, Univ. Paris-Dauphine, Paris, France.

[Geman84] S. Geman and D. Geman, Stochastic relaxation, Gibbs distributions and the Bayesian restoration of images, *IEEE Trans. PAMI*, Vol. PAMI-6, pp. 721-741, Nov. 1984.

[Hadam23] J. Hadamard, Lectures on the Cauchy Problem in Linear Partial Differential Equations, Yale University Press, 1923.

[Haral84] R. M. Haralick, *Digital Step Edges from Zero Crossing of Second Directional Derivatives*, IEEE Trans. on PAMI, 6, pp.58-68, 1984.

[Jain86] A. K. Jain, Fundamentals of Digital Image Processing. *Prentice Hall, Englewood Cliffs*, 1986.

[Koepf94] G. Koepfler, C. Lopezand J. M. Morel, A Multi scale Algorithm for Image Segmentation by Variational Method, *SIAM J. Numer. Anal.*, Vol. 31, No. 1, pp. 282-299, February 1994.

[Lazar96] L. Lazaroaía, A Study in a BV Space of a Denoising-Deblurring Variational Problem, prepubl. no. 439, Laboratory J-A Dieudonné, URA 168 du CNRS, University of Nice-Sophia Antipolis, France, 1996.

[Lee92] T. S. Lee, D. Mumford, A. Yuille, Texture Segmentation by Minimizing Vector-Valued Energy Functionals: The Coupled-Membrane Model, *Proc. of Computer Vision ECCV '92*, Vol. 588, LNCS Springer, May 1992.

[Lovel92] R. Lovell, W. R. Uttal, T. Shepherd, S. Dayanand, *A Model of Visual Texture Discrimination using Multiple Weak Operators and Spatial Averaging*, Pattern Recognition, Vol. 25, No. 10, pp. 1157-1170, 1992.

[Mumfo89] D. Mumford and J. Shah, Optimal Approximation by Piecewice Smooth Functions and Associated Variational Problems, *Commun. Pure Appl. Math*, vol. 42, pp. 577-685, 1989.

[Rande99] T. Randen, J. H. Husøy, Filtering for Texture Classification: A Comparative Study, *IEEE Trans. on PAMI*, Vol. 21, No. 4 April 1999.

[Rosen75] A. Rosenfeld, *Visual Texture Analysis: An Overview*, Tech. Rep. TR-406, Computer Science Center, University of Maryland, August 1975.

[Tebou98] S. Teboul, L. Blanc-Féraud, G. Aubert and M. Barlaud, Variational Approach for Edge-Preserving Regularization using Coupled PDE's, *IEEE Trans. on Image Processing*, Vol. 7, No. 3, March 1998.

[Tronc99] A. Troncone, D. Vitulano, A Novel Approach for Texture Segmentation under Noise, *Proc. of VMV '99*, 17-19 November 1999, Erlangen, Germany, pp. 179-185 (B. Girod, H. Niemann, H. P. Seidel Eds.).

[Liu00] X. Liu, D. L. Wang, R. Ramirez, *Boundary Detection by Contextual Nonlinear Smoothing*, Pattern Recognition, 33, pp. 263-280, February 2000.

[Zhu96] S. C. Zhu, A. Yuille, *Region Competition: Unifying Snakes: Region Growing, and Bayes/MDL for Multiband Image Segmentation*, IEEE Trans. on PAMI, Vol. 18, No. 9, pp. 884-900, 1996.