A MULTISCALE ALGORITHM FOR IMAGE SEGMENTATION BY VARIATIONAL METHOD.

G. KOEPFLER*, C. LOPEZ†, AND J.M. MOREL†

Abstract. Most segmentation algorithms are composed of several procedures: split and merge, small region elimination, boundary smoothing, ..., each depending on several parameters. The introduction of an energy to minimize leads to a drastic reduction of these parameters. We prove that the most simple segmentation tool, the "region merging" algorithm, made according to the simplest energy, is enough to compute a local energy minimum belonging to a compact class and to achieve the job of most of the tools mentioned above. We explain why "merging" in a variational framework leads to a fast multiscale, multichannel algorithm, with a pyramidal structure. The obtained algorithm is $O(n \ln n)$, where n is the number of pixels of the picture. We apply this fast algorithm to make grey level and texture segmentation and we show experimental results.

Key words. variational methods, nonnumerical algorithm, image processing, texture discrimination

AMS(MOS) subject classifications. 68Q20,68U10,

1. Introduction. The aim of this paper is to describe a fast and universal image segmentation algorithm. Properties of this algorithm, are :

It is multiscale and pyramidal. In other terms, it will not only compute one segmentation, but a hierarchy of segmentations from fine to coarse scales. Moreover the coarser segmentation will be deduced from the finer by "merging" operations, with a pyramidal structure for the computation. This corresponds to Marr's remark [Marr] that textures "live" at several scales. Therefore, a discrimination algorithm should give different kinds of segmentations which depend on the scale.

As a consequence of this pyramidal structure, the computation time is in practice proportional to the size of the datum. Moreover, this algorithmic structure will make it accessible to efficient hardware implementation.

The algorithm is universal, that is, does not depend on any a priori knowledge on the statistics of the image. Texture discrimination is achieved according to a universal criterion depending on very few parameters. More precisely, a picture is defined by a certain number of "channels" (grey level, colour levels, integrodifferential channels obtained by fast wavelet transform). In our segmentation algorithm, the only parameters are the weights attached to each channel, that is, the importance given to each channel as a segmentation criterion. However these parameters can be fixed once for all with reliable results.

The algorithm has been constructed by making a synthesis of several theories : the textons theory of Julesz [Ju], the energy methods in image segmentation introduced by Geman and Geman [GemG], Blake and Zisserman [BlakZ], Mumford and Shah [MumS1] , the Wavelet transform theory of Meyer [Me] , Mallat [Mall] and Cohen [Coh] , which unifies the theory of recursive filtering and pyramidal schemes.

The segmentations provided by the algorithm will be proved to have a large range of good topological and numerical properties, including compactness of the set of approximate solutions, convergence of minimizing sequences made of finer and finer

* COGNITECH, Inc., Santa Monica (CA), now at: U.F.R. de Mathématiques et Informatique, University Paris V-René Descartes 75270 Paris Cedex 06, France

† CEREMADE, University Paris IX-Dauphine 75775 Paris Cedex 16, France

solutions, smoothness of the locally optimal solutions, completeness of the multiscale representation, a priori estimates on the size of the regions of the segmentations.

2. General principles of segmentation devices.

2.1. Formalization. We define an image g as a scalar function, defined on the image domain Ω (generally a rectangle). The function g may also be vectorial, in the case where it has several channels for characterizing textures, histograms, colors, ... That does not change anything in the theorems and proofs that we later state. The only hypothesis is that these channels have been defined in order to be good indicators of the similarity – or difference – of points of the picture, and therefore good indicators for the autosimilarity of regions. We seek for a segmentation, that is, a partition of this rectangle into a finite set of regions, each of which corresponds to a part of the image where g is as constant as possible. Moreover, we wish to compute explicitly the region boundaries and of course control their regularity and location.
More precisely we will adopt the following principles.

The first principle is that the boundary detection problem must follow some universal rules. In other terms, we admit the possibility of a universal boundary detection device, definable and analyzable independently from the kind of channels (grey level, colour, texture, ...) to be used as input for the segmentation ([BecSI, MaliP2, Ju]). This first principle allows us to try to get a complete mathematical understanding of the segmentation problem, by considering the grey level segmentation. It is the simplest case of boundary detection, and it eliminates every early discussion about the concept of texture. (By "boundary in an image", we mean the boundary in the topological sense : it is the boundary of an homogeneous region of the image. Thus boundaries are different from edges obtained by some local filtering.)

A second principle which we adopt in the following is that an algorithm for boundary detection must be scale and space invariant. By space invariance, we mean that all points of the analyzed picture will be treated on the same way, and therefore a segmentation device must be translation and rotation invariant. Since boundaries may be present at every "scale", the scale invariance only means that we shall consider multiscale segmentation algorithms, depending on a scale parameter λ which we do not try to estimate and leave to the user's choice.

Finally we shall adopt a principle without which no discussion about segmentation can even start, and which we call comparison principle. It states that given two different segmentations of a datum, we are always able to decide which of them is considered as better than (or equivalent to) the other. Thus we assume the existence of some total ordering over all possible segmentations, and this can be simply achieved only if this ordering is reflected by some real functional E such that if $E(K_1) < E(K_2)$, then the segmentation K_1 has to be considered "better" than the segmentation K_2. For instance, this principle is verified by segmentation devices based on some Gibbs energy functional like in [GemG]. It is not verified by the region growing methods based on thresholds [Pav1, Pav2, PavL, Zu], nor by the edge detection devices [Marr, MaliP1].

Since, by the comparison principle, all of these criteria must be taken into account in the functional E, we see that this functional necessarily contains terms which control:

- the autosimilarity of each region with respect to the chosen channels. Those channels must be as constant as possible on each region.
- the size, location and regularity of the boundaries.

Denote by K the whole boundary set of the segmentation. According to these principles, a natural energy functional for segmentation will contain two terms, a two dimensional term for the autosimilarity of the regions which will roughly speaking measure the variance of g on each connected component of $\Omega \setminus K$ and a one dimensional one for controlling the length, and eventually the adequacy of location of the boundaries. Such a generic justification is developed e.g. in [MumS1, MumS2]. By the space invariance principle, those terms will be integral terms, the two dimensional one being an integral with respect to the Lebesgue measure and the one dimensional an integral with respect to the Hausdorff measure or "length" along the boundaries. Clearly, the weights given to the terms of the functional will be left to our choice. Consider for example the energy used by Mumford and Shah :

$$E(u,K) = \int_{\Omega\setminus K} |\nabla u|^2 dx + \int_{\Omega} (u-g)^2 dx + \int_K d\sigma \ ,$$

Both first integrals are bidimensional and the third one is with respect to the uniform measure $d\sigma$ supported by K. This energy means that in a "good" segmentation (u,K), the curves of K should be the boundaries of homogeneous regions in the image and u a sort of mean or, more generally, a regularized version of g in the interior of such areas (as illustration see the cartographical application in figure 4.1, u is choosen to be piecewise constant). The third term gives a control on the length, and therefore on the regularity of the boundaries of K (see below). This kind of functional represents of course a compromise between accuracy of the regions and parsimony of the boundaries: in case of grey level segmentation, as noticed in Zucker [Zu] or Haralick and Shapiro [HarS], a pure region growing would simply put together the pixels with similar grey levels. But this generates very nonsmooth boundaries and therefore also "small" or "thin" regions.
So if no control is made on the boundaries, one needs additional criteria and thresholding to achieve a decent segmentation. A functional like the above functional is designed to avoid this kind of mixed methods : one hopes having all the criteria put together in the same functional. For instance, the functional used for the so called "snakes" [KasWT, AmiTW] uses a more sophisticated third term for controlling both location (the boundary, or "snake", is forced to be close to edges) and smoothness of the boundary. Many early functionals in image analysis had only two dimensional energy terms (which were coupled with some thresholding criteria [Pav1, Zu]).

2.2. Choice of properties and algorithms for segmentation. The algorithmic description given by Pavlidis and Liow [PavL] is an excellent illustration of the level of sophistication to which the segmentation methods have arrived in Computer Vision, by superposing several tricky devices arisen in the two last decades. Now, the question which is naturally raised by the coexistence of energy functionals on one side, and of segmentation devices on the other is whether they match in any aspect, and how.

Our purpose is to classify the properties which are sought by those devices and to decide which of them are basic, and which can be deduced. Beyond being a way to attain certainty, mathematical proofs are a powerful tool for obtaining such a classification of properties. Once this hierarchy among tools has been established, one may hope to simplify the algorithms and to make them completely transmissible and reproducible. Our aim in this paper is to apply this method and to classify terms

like "small or thin region elimination", "splitting", "merging", "energy minimizing", "region growing", which have proved to be so necessary in most discussions about segmentation, all this following the ideas developed above. In the following we shall focus on the simplest model of Mumford and Shah [MumS1]. According to this model the segmentation (u, K) should be obtained by minimizing the functional

$$E(u,K) = \int_{\Omega \setminus K} (u-g)^2 dxdy + \lambda \ell(K), \tag{2.1}$$

where K is a union of boundaries in Ω with Hausdorff length $\ell(K)$, and u is piecewise constant on $\Omega \setminus K$. The constant λ represents the scale parameter of the functional and measures the amount of boundary : if λ is low, a lot of boundaries are allowed and we get a "fine" segmentation. As λ increases, the segmentation gets coarser and coarser. Similar functionals have been recently introduced by several authors in order to modelize physical phenomena like phase transition [FoT] and liquid crystals [DeG, AmbDeG, CarDeGL] .

In what follows, we shall focus on functional (2.1) because it is the simplest one containing most of the algorithmic difficulties of such functionals. Moreover, it is the only one for which a complete mathematical analysis is available. Indeed, Mumford and Shah have proved in [MumS1] that the boundary set obtained by this functional has the following geometric property : either the points of K are regular (at least C^1) or the singular points are of two types, namely, triple points where three branches meet with 120° angles and boundary points where K meets the boundary of Ω at a 90° angle. Moreover, the boundaries of the segmentation verify variational inequalities of the kind

$$(u^+ - g^-)^2 - (u^- - g^-)^2 \leq \mathrm{curv}(x) \leq (u^- - g^+)^2 - (u^+ - g^+)^2, \tag{2.2}$$

where u^+ and u^- denote the values of $u(x)$ on both sides of the boundary.

In [MorS1, MorS2] Morel and Solimini have given an elementary and constructive proof of the above theorem.
Of course, the fact that the crossing points must be ternary proves that such a functional gives no hope of accurate segmentations. But this is, in our view, secondary, because one can improve the functional by adding or modifying boundary terms (of the kind used for snakes) or surface terms (for instance by imposing that u is a linear, quadratic or cubic spline, ...). This would not change the nature of the proofs and methods which we discuss here and which apply to most energy functionals mentioned above. Note however, that the above mentioned general functional of Mumford and Shah has a different structure. As proved in [BlatM, CarLPP, MumS1], it allows open boundaries and therefore the merging methods which will be discussed here are not adequate. Moreover, the proof of the "Mumford and Shah conjecture", which states that this functional would lead to a finite set of smooth boundaries is not achieved. The best available result [DalMMS1, DalMMS2] is that the minimizing boundaries (which in some weak sense exist [Amb]) can be approximated by a finite number of curves. No regularity result is available, and we see it difficult to comment now mathematically the first and interesting numerical simulations [BlakZ] of this problem. An attempt to do so is however presented in [DibK]. Let us therefore return to our "simplest functional" (2.1). It is well known that functionals of this kind may have many local minimizers. In the case of the functional (2.1), it is easy to give explicit examples.

Moreover, a reasonable conjecture is that to find a minimizer, or even to prove a posteriori that some computed local minimum is global leads to a NP-complete problem.

One has thus to choose between two strategies :

- The global minimization by simulated annealing methods, which leads to huge computations, but ensures that in some asymptotic sense, the global minimum is attained [GemG].
- Another way is then to define some concept of local minimum which should be more accessible to fast computations and verify the same properties which were seeked for the global minimum.

For instance, in the case of simulated annealing, a recent tendency is to define some faster, parallel processes, but which do not pretend to find a global minimum anymore [Az1, Az2]. The homotopy method of Blake and Zisserman [BlakZ] also seeks for "good" local minima. (It seems to be close to a Γ-convergence device suggested by the mathematical school of De Giorgi [AmbT] and developed by Tom Richardson in his Phd. Dissertation at MIT [Ric]).One needs a theory to do that in the deterministic approach to which the segmentation devices mentioned above belong.
In the next section we shall give some basic definitions and notations concerning our "simplest" energy functional. In sections 3.1 and 3.2, we give a precise account of topological properties and the proofs of the announced results concerning segmentations obtained by "merging". The definition and properties of what we could call the "simplest possible" recursive merging algorithm which will be described in section 4.1, there will also be shown results obtained by running this algorithm on grey level images and images containing "textures".

3. Compactness properties of segmentations obtained by "merging".
Before presenting the main result, let us fix a functional and topological glossary.

- The function g. It is a bounded measurable real function on a rectangle Ω (more generally for $g : \Omega \mapsto \mathbb{R}^N$, $N \geq 2$ the same result holds). Without loss of generality we assume that $|g(x, y)| \leq 1$.
- The *energy* $E(K)$: note that given the boundaries K, the corresponding minimal u is completely defined by the fact that its value on each connected component of $\Omega \setminus K$ is equal to the mean value of u on this connected component. Thus we shall always assume in the following that to each K is associated this unique u. Therefore we shall write $E(K)$ instead of $E(u, K)$.
- *Regions*, or connected components of $\Omega \setminus K$: we shall denote them by $(O_i)_i$.
- *Common boundary* of two regions O_i and O_j : we denote it by $\partial(O_i, O_j)$. It is contained in K. If $i = j$, ∂O_i denotes the boundary of O_i.
- *Two dimensional measure* of O_i : denoted by $|O_i|$.
- *Isoperimetric inequality* in $\mathbb{R}^2$ and Ω : denote by O a region in $\mathbb{R}^2$. Then one has

$$\ell(\partial O) \geq 2\sqrt{\pi} \cdot \sqrt{|O|} \, .$$

  In the case of a domain Ω with smooth boundary, the same kind of inequality holds for the *relative boundary* of O in Ω, $\partial O \cap \Omega$, with a smaller constant C :

$$\ell(\partial O \cap \Omega) \geq C \cdot \sqrt{|O|} \, .$$

(Since there is no ambiguity, Ω being fixed, we still denote by $\partial O = \partial O \cap O$ the relative boundary of O in Ω.) By $\ell(\partial O)$ we mean the length of the boundary of O. For a general definition of the length of the boundary ("perimeter") of a region in the plane, see [Fed, §4.5.1] or [Si-L, §14].

DEFINITION 3.1. *A segmentation K will be called normal if every subsegmentation K' of K verifies $E(K') > E(K)$.*

By a subsegmentation of K, we mean a segmentation obtained by merging an arbitrary number of adjacent regions. The normality of a segmentation means that no merging operation of a set of regions of the segmentation K can decrease the energy. This definition implies in particular that if K is normal, no boundary between two regions of the segmentation may be removed without increasing the energy E.

DEFINITION 3.2. *A segmentation K will be called 2-normal if for every pair of regions O_i and O_j, the new segmentation K' obtained by merging these regions verifies $E(K') > E(K)$.*

2-normality, with the same definition, but under the name of "optimality" has been introduced by Pavlidis [Pav1].

We shall consider only segmentations having the following properties, which are easy to check for computationally defined segmentations.

a) The number of regions is finite. In other terms, $\Omega \setminus K$ has a finite number of connected components.

b) No region has internal boundaries. In other terms, the interior of the closure of each region O is equal to the interior of O. Indeed, if this is not the case, one can remove the internal boundaries without increasing the energy. This property is called the "1-normality" (compare to the definitions above).

c) The segmentations are made of piecewise affine curves. (This assumption can be changed into piecewise C^1, or piecewise Lipschitz without changing the proofs.)

Let us introduce the framework in which we will study the topological properties of affine 1-normal segmentations.

- A *curve* c from $[0,1]$ into Ω has *tips* $c(0)$ and $c(1)$, all other points of the range of c are *interior points*.
- *Geometrical support* of a curve c : denoted by $S(c)$, it is the range of the curve c. For several curves it is the union of the ranges.
  For instance $K = \bigcup_i S(c_i)$.
- A segmentation is said to be *piecewise affine* – resp. C^1 – if the corresponding c_i are *piecewise affine* – resp. C^1 –. A *polygon* is a connected region of an affine segmentation.
- *Geometrical crossings* : all the points of K where either a curve meets the interior of another, either three curves at least have a common tip, or a curve meets $\partial\Omega$.
- *Geometrical curve* : subset of K which is in the geometrical support of a curve c such that its boundary is contained in the set of geometrical crossings and that no one of its interior points is a geometrical crossing.
- A *Jordan curve* is a continuous curve, such that for all $\sigma, \sigma' \in]0,1[$: $c(\sigma) \neq c(\sigma')$ if $\sigma \neq \sigma'$; if $c(0) = c(1)$ the Jordan curve is said to be *closed*.

3.1. Topological properties of affine 1-normal segmentations. Let us recall a classical result on Jordan curves (for a proof see for example [Ale, Th])

LEMMA 3.1. *Every closed Jordan curve c divides the plane in exactly two connected components, one is bounded "enclosed by c" and one unbounded.* We will say that a closed Jordan curve c *encloses* a part of the plane if this part is included in the bounded connected component delimited by c.

By using the preceding lemma one easily proofs the following (see[MorS1, MorS2]).

LEMMA 3.2. *Take an affine 1-normal segmentation K having more than one region, then there are two regions enclosed, O and O', such that $\partial(O, O')$ is a Jordan curve.*

LEMMA 3.3. *Let K be an affine 1-normal segmentation with α regions. Then K is the union of $\alpha - 1$ affine Jordan curves having no common segments.*

Proof. Let c_1 be an affine Jordan curve as given by lemma 3.2 . Let us define a new segmentation K_1 obtained by removing the segments of c_1. We merge exactly two regions of K. The segmentation K_1 is also affine and 1-normal, so we can iterate the process and define a series of segmentations K_i such that every segmentation has a region less than the preceding one and has been obtained by removing an affine Jordan curve. The process stops if there is only one region left, Ω. □

LEMMA 3.4. *Let α be the number of regions of an affine 1-normal segmentation, β the number of geometrical curves and γ the number of geometrical crossings. Then*

$$\gamma \leq 2 \cdot (\alpha - 1) \quad \textit{and} \quad \beta \leq 3 \cdot (\alpha - 1) - 2 \ .$$

Proof. Consider the proof of the preceding lemma (3.3). The curve c_1 contains at most two geometrical crossings of K_1 : its tips if it is an open Jordan curve. So passing from $K = K_0$ to K_1 eliminates at most two geometrical crossings. By iterating the process the first inequality follows. Let us now analyze how the number of geometrical curves decreases when passing from K_0 to K_1. If c_1 is a closed Jordan curve it is the only geometrical curve to disappear.
If c_1 is an open Jordan curve there are two geometrical crossings at its tips. If one of these crossings meets more than four segments of K the other geometrical curves are not modified. But if there are only three segments at this crossing it will vanish when c_1 is removed and the other two geometrical curves will merge into one.
Removing c_1 decreases β of at most 3 and iterating the process yields the second inequality. □

3.2. Estimates on segmentations obtained by merging. LEMMA 3.5. *Let (u, K) be a 2-normal segmentation. Then every pair of regions O and O' verifies*

$$\ell(\partial(O, O')) \leq \frac{\mathrm{osc}(g)^2}{\lambda} \cdot \min(|O|, |O'|) \ ,$$

where $\mathrm{osc}(g) = \sup(g) - \inf(g)$ *is the oscillation of g.*

Proof. Let $K' = K \setminus \partial(O, O')$. By 2-normality we have

$$0 \leq E(K') - E(K) = \int_{O \cup O'} (u_{ij} - g)^2 - \int_O (u_i - g)^2 - \int_{O'} (u_j - g)^2 - \lambda \ell(\partial(O, O')),$$

where u_i, u_j and u_{ij} are the mean values of g on O, O' and $O \cup O'$. We get, supposing for example that $|O| \leq |O'|$,

$$\begin{aligned} \lambda \cdot \ell(\partial(O, O')) &\leq \int_O \left[(u_{ij} - g)^2 - (u_i - g)^2 \right] \\ &\leq \min(|O|, |O'|) \operatorname{osc}(g)^2 \quad \square \end{aligned}$$

LEMMA 3.6. *For every region O of a 2-normal segmentation, denote by $N(O)$ the number of neighbouring regions. Then*

$$N(O) \geq \frac{C\lambda}{\sqrt{|O|}\, \operatorname{osc}(g)^2} ,$$

where C is the isoperimetric constant in Ω.

Proof. Call O_j a neighbouring region of O. The fact that O and O_j cannot be merged without increasing the energy E implies that

$$\lambda\, \ell\left(\partial\left(O, O_j\right)\right) \leq |O|\, \operatorname{osc}(g)^2.$$

Thus by adding these inequalities for all neighbours of O, we obtain

$$\lambda\, \ell(\partial O) \leq N(O) |O|\, \operatorname{osc}(g)^2.$$

We conclude by applying to O the isoperimetric inequality in Ω.$\square$

LEMMA 3.7. *Let α be the number of regions of a 2-normal affine segmentation. Then*

$$\alpha \leq \frac{288\, |\Omega|\, \operatorname{osc}(g)^4}{C^2 \lambda^2} .$$

Proof. The union of all regions O_i is equal to Ω and therefore $\sum_i |O_i| = |\Omega|$. Thus the number of O_i verifying $|O_i| \leq \dfrac{2}{\alpha}\, |\Omega|$ is greater than $\dfrac{\alpha}{2}$.
Let us now apply lemma 3.6 to all of these O_is. Each one of them has at least

$$\frac{C\lambda}{\sqrt{|O_i|}\, \operatorname{osc}(g)^2} \geq C\lambda \cdot \sqrt{\frac{\alpha}{2\, |\Omega|}} \cdot \frac{1}{osc(g)^2}$$

neighbouring regions. Consequently the number β of common boundaries, and therefore of geometrical curves, verifies

$$\beta \geq \frac{\alpha}{4} \cdot C\lambda \sqrt{\frac{\alpha}{2\, |\Omega|}} \cdot \frac{1}{\operatorname{osc}(g)^2} = C\lambda 2^{-\frac{5}{2}} \cdot \frac{\alpha^{\frac{3}{2}}}{\sqrt{|\Omega|}} \cdot \frac{1}{\operatorname{osc}(g)^2} .$$

By lemma 3.4, we have

$$\beta \leq 3\alpha,$$

this implies that

$$C\lambda 2^{-\frac{5}{2}}\, \frac{\alpha^{\frac{3}{2}}}{\sqrt{|\Omega|}\, \operatorname{osc}(g)^2} \leq 3\alpha$$

and thus we obtain

$$\alpha \leq \frac{288\,|\Omega|\,\mathrm{osc}(g)^4}{C^2\lambda^2} \qquad \square$$

REMARK 3.1. *This is the same estimate (with a bigger constant) than in Mumford and Shah [MumS1, Th.5.2]. However, in the mentioned paper, this estimate is obtained for a global minimum. The fact that we get this estimate in the case of 2-normal segmentations indicates an analogy of structure between global minima and this kind of local minimum.*

COROLLARY 3.8. *The set of 2-normal affine segmentations has the following compactness property : for every sequence K_n of such segmentations, there exists a subsequence converging to a segmentation K such that*

$$E(K) \leq \liminf_n E(K_n).$$

K is not necessarily 2-normal, but has anyway a 2-normal subsegmentation with still less energy.

Proof. The proof of the announced compactness property is based on the fact that the number of edges of any 2-normal segmentation is now bounded from above by the preceding estimates (cf. lemmas 3.4 and 3.7). By the Ascoli-Arzela theorem, each one of the edges can be supposed to converge to a limit edge if we extract an ad hoc subsequence. The limit segmentation is then defined by these limit edges. When passing to the limit, we know by Fatou's lemma that neither the integral part of the energy nor the length of the edges can increase, and therefore this limit segmentation has an energy smaller than the inf limit of the energies of the sequence. For more (technical) details see [MorS1].$\square$

REMARK 3.2. **: Elimination of small regions ([Zu, HarS]).**
It is easy to deduce from the preceding proof a lower bound on the area of each region of the segmentation. Indeed, take any region O of the segmentation. By Lemma 3.6, the number of neighbouring regions O_j is at least

$$N(O) \geq \frac{C\lambda}{\sqrt{|O|}\,\mathrm{osc}(g)^2}\,.$$

Thus the number α of regions of the segmentation is at least

$$\frac{C\lambda}{\sqrt{|O|}\,\mathrm{osc}(g)^2}\,.$$

By using the upper bound for α, given in lemma 3.7 we get

$$\frac{288\,|\Omega|\,\mathrm{osc}(g)^4}{C^2\lambda^2} \geq \frac{C\lambda}{\sqrt{|O|}\,\mathrm{osc}(g)^2}\,.$$

Thus the area of O is bounded from below by a positive constant only depending on g, λ and Ω. Therefore a merging method based on the minimizing of the energy $E(K)$ will spontaneously eliminate the small regions. This process was considered as heuristic and parameter-dependent in [HarS].

REMARK 3.3. **: Elimination of thin regions ([Zu, HarS]).**
It is also easy to deduce from the above estimates that the regions are not too "thin",

that is, verify an "inverse isoperimetric inequality". Indeed, each region O verifies for a constant C depending only on g and Ω,

$$|O|^{\frac{1}{2}} \geq C \cdot \ell(\partial O) \, .$$

This last result can be deduced from remark 3.2, since the surface of O is bounded from below and $\ell(\partial O)$ from above (as the length of the segmentation is under control). The estimate obtained can be proved with a better constant C (see [MorS1]). Thus the devices based on the elimination of thin regions (see [PapJ]) for instance, and many clustering algorithms [Pav2]) can be considered as implicit in the search of an optimal 2-normal segmentation and do not depend anymore on extra threshold parameters.

REMARK 3.4. **: Smoothing of the boundaries ([PavL]).**
The 2-normal segmentations have no chance of having boundaries smooth everywhere. However, since their length is under control, a classical geometric measure theory (see for example [Si-L, th.14.3]) asserts that they are almost everywhere C^1. Thus the presence of noise, for instance, can alter the regularity of these boundaries but not make them increase indefinitely, as it is the case for some region growing devices. What can be done in order to restore this regularity and how can such a regularization device be deduced from the energy to be minimized ?
The answer is in relation (2.2), which asserts that the curvature of an energy minimizing boundary is controlled. This equation shows that the length term of our "simplest energy", coupled to its bidimensional contrast measuring term, is enough to ensure that the boundaries are analogous to snakes. They tend to keep at places where g has a jump and they remain smooth, since from the bound on the curvature follows a bound on the first derivative [MorS1]. The merging of regions should therefore be completed by a merging of the boundaries in the same sense : imposing therefore that this boundary is smoothed according to the criterion on curvature imposed by the energy. This idea is implicit in [PavL]. As noticed by these last authors, that imposes to treat region boundaries like snakes, but the relation (2.2) proves that still this tool can be deduced from the "simplest energy" that we have considered.

4. A pyramidal algorithm constructing 2-normal affine segmentations.

We now consider the problem of defining and computing a 2-normal segmentation. Notice that not all 2-normal segmentations are equally interesting : for instance, the empty segmentation, where Ω is the single region is clearly a 2-normal segmentation. If the scale parameter λ is very large, it is however also a reasonable segmentation since one "pays" a too large energy amount for having any boundary. Now, it is obvious from the definition that the empty segmentation is 2-normal for every λ, which certainly proves that the assertion that a segmentation is 2-normal is not enough to ensure that it is "good". But if we follow the main idea of the region growing methods [Zu], *we shall see that what they compute is precisely a 2-normal subsegmentation of a fine initial segmentation, obtained by recursive merging.*

Assume that the datum g is defined on a rectangle. This rectangle is divided in small squares of constant size (the pixels) and g is assumed to be constant on each pixel. Here are the properties which we require for the segmentations computed by a region growing algorithm, defined as an application associating to g and λ a segmentation (u, K).

a) "Correctedness" (Fixed point property) : Assume that g is piecewise constant on some regions of the rectangle. Then there exists a value λ_0 of the parameter λ such that for every $\lambda < \lambda_0$, the segmentation (u, K) obtained by the algorithm verifies $u = g$ and K is the union of the boundaries of the areas where g is constant. This property has been proved to be asymptotically true for the segmentations which are global minima of the energy E as λ tends to zero [Ric]. But we impose it here as a nonasymptotic property.

b) "Causality" (Pyramidal segmentation property) : If $\lambda > \lambda'$, then the boundaries provided by the algorithm for λ are contained in those obtained for λ' and the regions of the segmentation associated to λ are the unions of some of the regions obtained for λ'.

The last property ensures that a fast pyramidal algorithm can be implemented, computing a hierarchy of segmentations from fine to coarse scales. Moreover the coarser segmentation will be deduced from the finer by "merging" operations, with a pyramidal structure for the computation. Note that, as a consequence of the fixed point property, if λ is very small, the computed segmentation is attained with (u_0, K_0) where $u_0 = u$ and K_0 consists of all the boundaries of all the pixels, and therefore coincides with the global minimum as λ is zero. We shall call this segmentation, where each pixel is a region, the "trivial segmentation". The recursive merging algorithm which we use verifies all the above mentioned properties.

4.1. Description of the algorithm.

The criterion. The decision to proceed to a merging of two regions O_i and O_j depends on the sign of $E(\tilde{u}, K \setminus \partial(O_i, O_j)) - E(u, K)$. The algorithm looks for a decrease of global energy by merging these regions. This is the criterion of 2-normal segmentations introduced in the discussion of section 2. The simplified Mumford and Shah model is implemented by choosing the following energy functional :

$$E(u, K) = \int_\Omega \|u - g\|^2 \; + \; \lambda \ell(K).$$

Here g is a vector valued function, whose components are different channels, defined on the rectangle Ω, u is the approximating vector function, and K is the set of boundaries with total length $\ell(K)$. As in the piecewise constant case, u is the mean value of g on each region O. Thus the above functional is just denoted by $E(K)$, i.e. to obtain (u, K) one needs to know only K. Then the merging criterion is

$$E\left(K \setminus \partial\left(O_i, O_j\right)\right) - E(K) = \frac{|O_i| \cdot |O_j|}{|O_i| + |O_j|} \cdot \|u_i - u_j\|^2 - \lambda \cdot \ell\left(\partial\left(O_i, O_j\right)\right),$$

where $|.|$ is the area measure and u_i the approximation of g on O_i. When g is scalar the norm $\|.\|$ is just the absolute value. For multichannel data a weighted norm $\|.\|$ is used. It is specific to each application and to the meaning of the different channels. This will be emphasized in the next section.

To obtain the necessary data for evaluating the criterion the following information has to be used: Suppose $g = {}^t(g^1, \ldots, g^n)$, then to each region O we associate its area $|O|$ and n channels $c_O^l = \displaystyle\int_O g^l$, $(l = 1, \ldots, n)$. These yield the values for u restricted to O : $u_O = {}^t(u_O^1, \ldots, u_O^n)$ by simply computing $u_O^l = \dfrac{c_O^l}{|O|}$ to get the mean value.

The channels of region O_{new} obtained by merging O_i and O_j are given by $|O_{\text{new}}| = |O_i| + |O_j|$ and $c^l_{O_{\text{new}}} = c^l_{O_i} + c^l_{O_j}$, $(l = 1, \ldots, n)$.

Thus a merging of two regions only implies adding the corresponding channels and updating the data structure.

The algorithm.

(i) Take the segmentation (u_0, K_0) resulting from the initialization and $\lambda_i = \lambda_1$ as a scale parameter.

(ii) Scan the list of regions and for every candidate region look for the adjacent region which yields the best merging score (i.e. the maximal energy decrease). If such a region exists proceed to merge and update the data. The next region in the list becomes a candidate for merging.
For $\lambda = \lambda_i$ fixed, repeat the scanning of the picture until no merging is possible. After this step, a 2-normal subsegmentation (u_i, K_i) of the initial segmentation for scale parameter λ_i is achieved.

(iii) For every λ_i, i=1,...,L, calculate a 2-normal segmentation by iterating step (ii). The algorithm stops if there is just one region left or after computing a 2-normal segmentation (u_L, K_L) for the last scale parameter λ_L.

Experimental results. In figure 4.1 we show in the top left corner a satellite image which has been segmented with respect to a single channel, the grey level: in the upper right image we show the piecewise constant reconstruction u, in the lower left picture the boundary set K is represented and in the lower right corner we have drawn the boundaries upon the original picture.

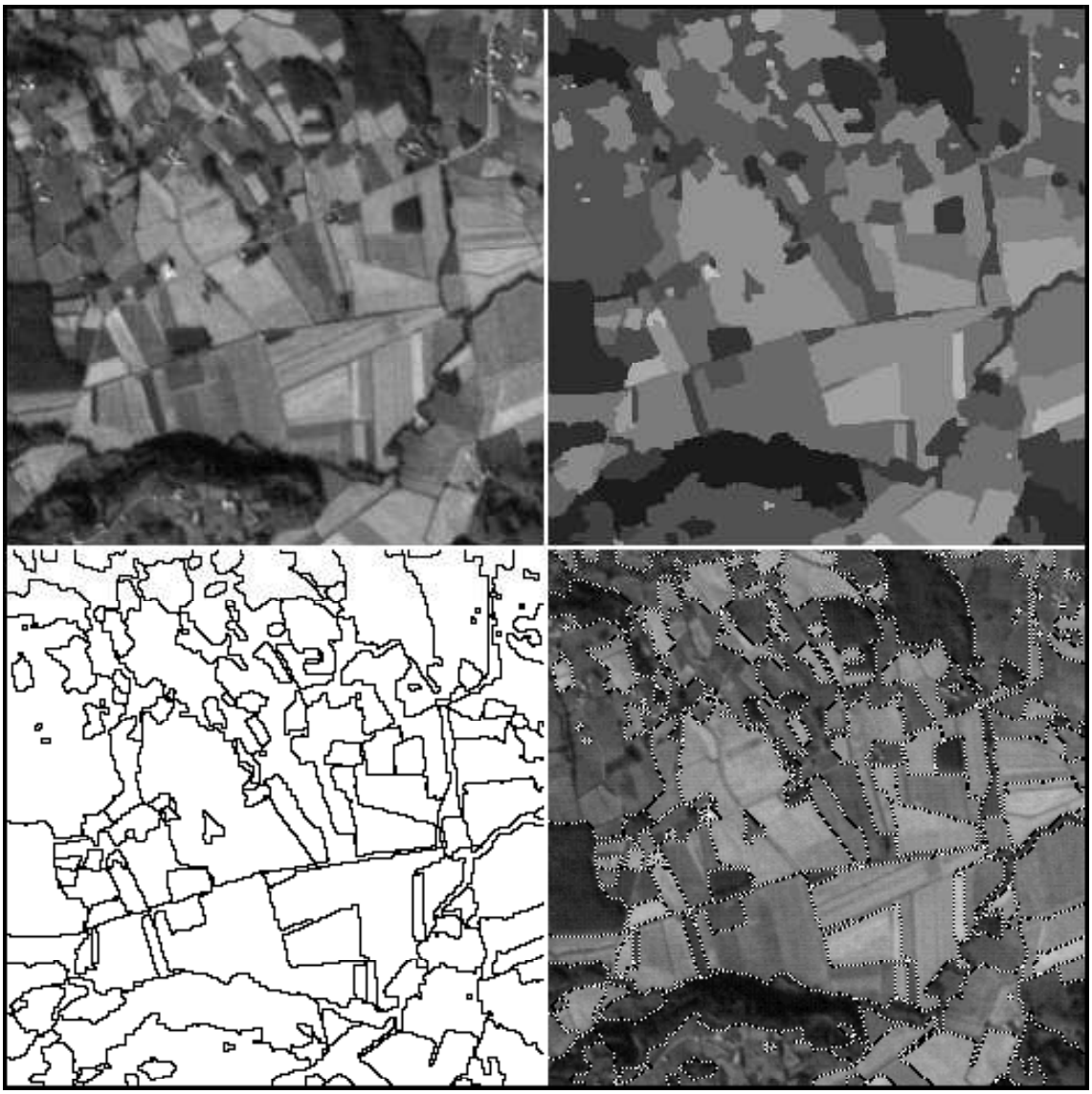

FIG. 4.1. *Scale $\lambda = 2048$, 117 regions (original picture : GRECO TdSI). Representation of the original picture g, the piecewise constant approximation u, the boundaries K and the original picture with boundaries.*

5. Application to texture discrimination. We follow the ideas of David Marr [Marr] concerning the "raw primal sketch". According to this theory the grouping process of the human visual system should be based on the detection of local features which from the mathematical viewpoint are simply differential or semilocal integrodifferential operators like the derivatives or convolution by Gaussians. The preattentive texture discrimination depends on the fact that at least one of the features has a bigger or lower local density (see [MaliP2]).

It is important to note that most of the channels introduced by the algorithms of [MaliP2], [VoP], [BovCG], ... are linear operators which, according to the Wavelet theory, can be computed at each scale in linear time by a pyramidal scheme (also called in this context quadrature mirror filter [Coh], [EsG]).

The wavelet transform allows to compute a multiscale gradient in every direction in linear time (by a quadrature mirror filter). Since, as recalls David Marr, textures "live independently at different scales", it will be much easier to get discrimination of most textures with a fast multiscale analysis. One can for instance use the Daubechies filters [Daub1, Daub2], which can be computed with quadrature mirror filters having very few coefficients (eight coefficients are enough for a smooth wavelet with two zero moments). In a work in collaboration with Jaffard and Journé, Daubechies has proved a fast algorithm for a variant of Gabor filtering ([DaubJJ]). Fast Gabor filtering may also provide cheap channels for our segmentation algorithm, which prove very useful for almost periodic textures.

Note however that the human eye is by no means perfect in texture discrimination and that it can reasonably be argued that a discrimination algorithm could be more accurate and discriminate textures which are not directly accessible to the human eye. In our opinion, a successful computational model of texture discrimination should do better than the human eye in some cases, and therefore be used as an enhancement operator (see figure 5.3).

Let us summarize the assumptions on which we base our segmentation algorithm :

- We assume, following the theory of Beck and Julesz (see [Bec, Ju]) and some recent experimental and computational confirmation of it [BovCG, MaliP2, VoP] that a reasonable number of channels defined on the image is enough in order that for any pair of preattentively different textures, at least one of the channels will help to discriminate them.
- We are not able now to say which channels are necessary or sufficient in order to get a discrimination similar to that of the human eye. Now, our universal segmentation algorithm is designed for fast specific applications, but also as a fast and robust experimental device allowing progress in the question of which channels are necessary in order to match the human visual performance.

Examples. Let $g = (g_1, \ldots, g_n)$ be the initial data calculated from the picture $\tilde{g}$ and $u = (u_1, \ldots, u_n)$ the piecewise constant estimate (it turns out that u_i is the mean value of g_i on each connected component). Using the functional

$$\int_\Omega \|u - g\|^2_{pond.} + \lambda \cdot \ell(K) \, ,$$

where $\|.\|_{pond.}$ is a weighted norm, we calculate a 2-normal segmentation.

The initial datum g for the experiments related below is given in our experiments by

an oversampled Haar-Wavelet transform of the picture $\tilde{g}$. More precisely the image is convloved with a bank of linear filters F_k, followed by half-wave rectification.

$$R_{2k} = (\tilde{g} \star F_k)^+ (x, y) \ ; \ R_{2k+1} = (\tilde{g} \star F_k)^- (x, y)$$

It should be noticed that all the filters F_k are of zero mean and separable:

$$\begin{aligned} H_a^1(x, y) &= \chi_{[-a,a]}(x) \left(\chi_{[0,a]}(y) - \chi_{[-a,o]}(y) \right) \\ H_a^2(x, y) &= \chi_{[-a,a]}(y) \left(\chi_{[0,a]}(x) - \chi_{[-a.0]}(x) \right) \\ H_a^3(x, y) &= \left(\chi_{[0,a]}(x) - \chi_{[-a,o]}(x) \right) \left(\chi_{[0,a]}(y) - \chi_{[-a,o]}(y) \right) \end{aligned}$$

where χ is the caracteristic function on $\mathbb{R}$ and $a = 2^j$, $1 \leq j \leq J$, J is the decomposition order of the analysis.
The R_is are then filtered by Gaussians in order to obtain texton densities. The size of these Gaussians corresponds to the "Δ-neighbourhoods" of Julesz. Let us recall the definition of Julesz [Ju]: " The Δ-neighbourhood is the area, in which differences in texton densities are determined. Textons are formed only if the adjacent elements lie within the Δ-neighborhood."
Finally we obtain the g_is, the components of our initial datum g, as the filtered versions of the R_i channels.

The aim of our experiences is to get a verification of the Julesz doctrine : we stop our region growing if the desired number of regions is reached (e.g. if there are two textures we proceed until we have a partition in two regions of our image) . If the discrimination is successful (i.e the two regions correspond to the textures' location) one can say that the used channels are able to discriminate the given textures (it is important to notice that we don't use the grey-channel information).

The experimentations used Brodatz pictures (see [Bro]) and a synthetic image which illustrates the need to keep all the channel involved in the discrimination process.

6. Conclusion. We have proved that the minimizing of the "simplest" segmentation energy entails the implicit realization of properties sought by most segmentation devices. Indeed, the most primitive segmentation tool, the "merging", applied to the simplest possible segmentation energy, is enough to ensure a compact and therefore small set of possible segmentations, with no small regions and no thin regions. Uniform a priori estimates for the size and number of the regions can be given for all segmentations obtained by exhaustive "merging". Moreover, the region growing method associated with the recursive merging is enough to retrieve all piecewise constant functions. Such a merging method is not accurate enough to obtain smooth boundaries, but it controls anyway their length. The big advantage of this method is its velocity.
Aknowledgment: We thank Yves Meyer, Pietro Perona and Lenny Rudin for valuable conversations and remarks.

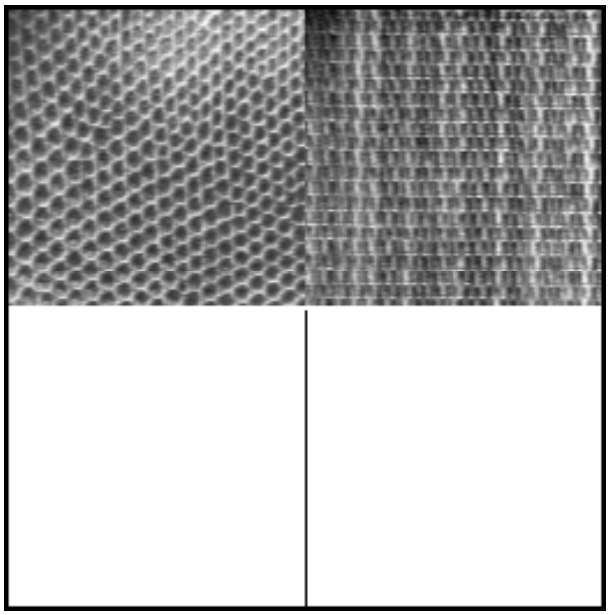

FIG. 5.1. *Above two Brodatz textures, below the boundary calculated (Δ-neighbourhood of 16 pixels, level of wavelet transform $J = 3$) .*

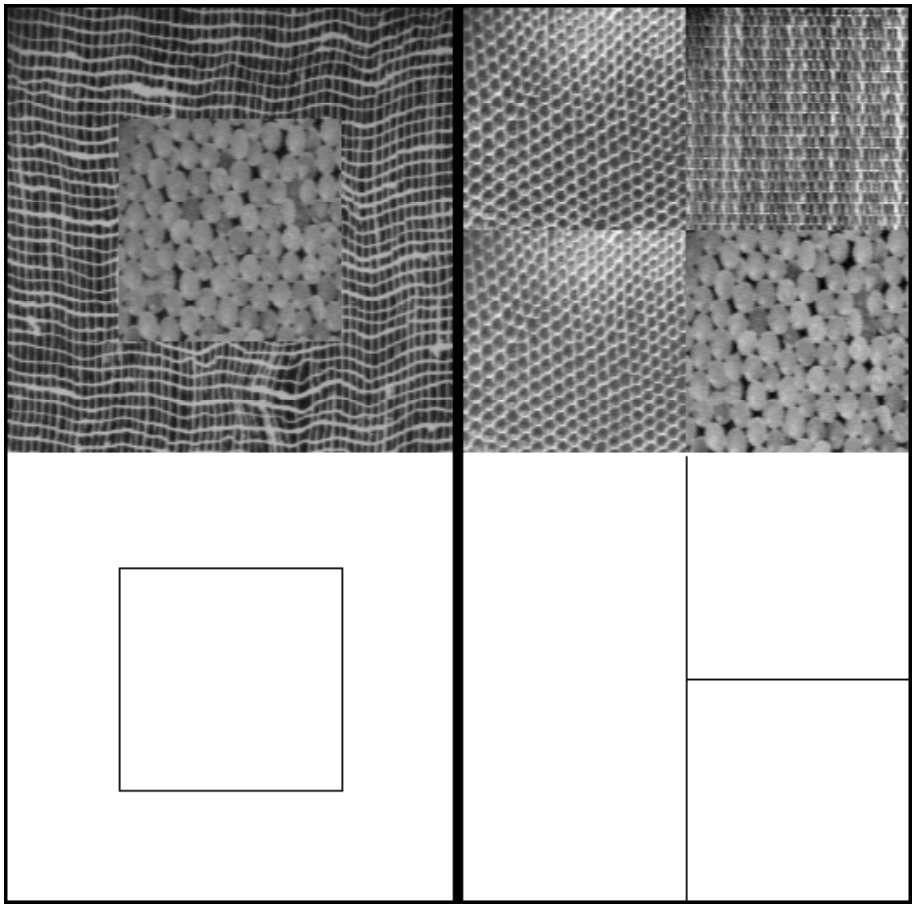

FIG. 5.2. *Left part : two Brodatz textures, right part : three different Brodatz textures. Below each the boundary calculated (same parameters : Δ-neighbourhood of 16 pixels, level of wavelet transform $J = 3$) . Notice that in this case the same texture is repeated twice left. Unless the boundary between them is visible, the algorithm recognizes the identity of the texture and segments correctly.*

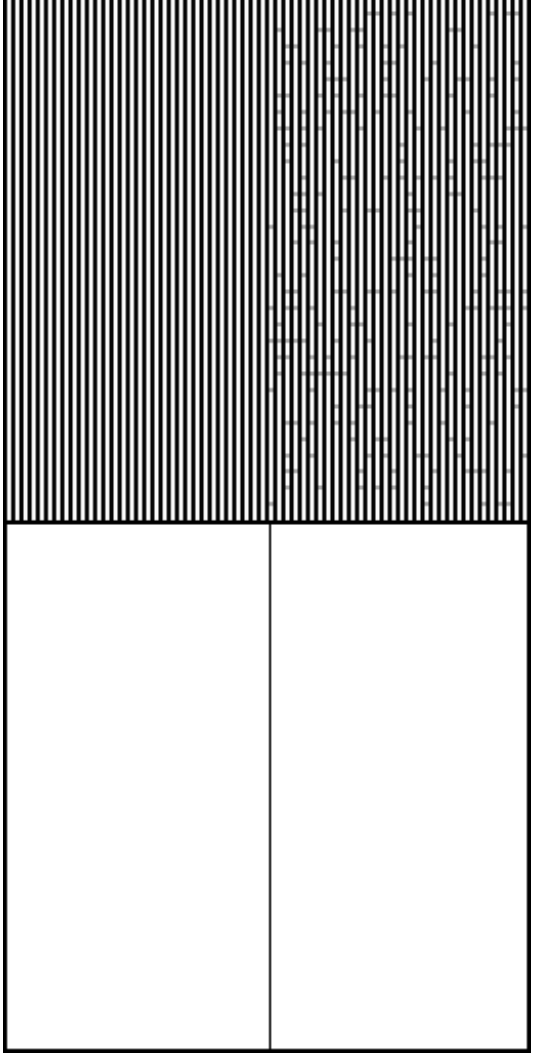

FIG. 5.3. *Above the textures : on the left hand one dominant channel; on the right hand the same dominant channel with a weak second channel. Below the boundary calculated (same parameters:* Δ*-neighbourhood of 16 pixels, level of wavelet transform* $J = 3$*) . This experiment illustrates the power of the algorithm for discrimination based on nondominant channels. Most texture analysis devices tend to base it on the dominant channel. If the dominant channel is the same in the two regions, no discrimination [BovCG,MaliP2,Vop2] is possible. It is important to notice that a variational method yields an easy formalization for segmentation based on non-necessary dominant channel.*

[Ale] P.S. Aleksandrov. *Combinatorial Topology.* Vol. 1, chap. 2, Graylock Press NY, 1956.

[Amb] L. Ambrosio. *A compactness theorem for a special class of functions of bounded variation.* Boll. Un. Mat. Ital. (to appear).

[AmbDeG] E. De Giorgi, and L. Ambrosio. *Un nuovo tipo di funzionale del calcolo delle variazioni.* To appear in Atti Acad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.

[AmbT] L. Ambrosio, and V. M. Tortorelli. *Approximation of functionals depending on jumps by elliptic functionals via G-convergence.* Preprint. Scuola Normale Superiore di Pisa.

[AmiTW] A. Amini, S. Tehrani and T. E. Weymouth. *Using dynamic programming for minimizing the energy of active contours in the presence of hard constraints.* Second International Conference on Computer Vision (Proceedings of), ICCV'88, IEEE n°883.

[Az1] R. Azencott. *Synchronous Boltzmann machines and Gibbs fields; Learning algorithms.* To appear, Springer Verlag 1989, (NATO Series, Les Arcs Congress Proceedings).

[Az2] R. Azencott. *Boltzmann Machines : high order interactions and synchronous learning.* Submitted to IEEE PAMI Trans.

[Bec] J. Beck. *Organization and representation in perception.* Chapter in *Textural se gmentation*, Erlbaum, Hillsdale, NJ, 1982.

[BecSI] J. Beck, A. Sutter and R. Ivry. *Spatial frequency channels and perceptual grouping in texture segregation.* Computer Vision, Graphics and Image Processing, 37, 299-325, 1987.

[BlakZ] A. Blake and A. Zisserman. *Visual Reconstruction.* MIT Press, 1987.

[BlatM] J. Blat, and JM. Morel. *Elliptic problems in image segmentation and their relation to fracture theory.* Proceedings of the Int. Conf. on nonlinear elliptic and parabolic problems, Nancy 88 (Longman 89).

[BovCG] A.C. Bovik, M. Clark and W. S. Geisler. *Computational texture analysis using localized spatial filtering.* IEEE 1987.

[Bro] Ph. Brodatz. *Textures for artists and designers.* Dover Publications Inc. NY 1966.

[CarDeGL] M. Carriero, E. DeGiorgi and A. Leaci. *Existence theorem for a minimum problem with free discontinuity set.* Arch. Rat. Mech. Anal., vol. 108, 1989.

[CarLPP] M. Carriero, D. Leaci, E. Pallara and E. Pascali. *Euler conditions for a minimum problem with discontinuity surfaces.* Preprint Universita degli studi di Lecce (Italy)

[Coh] A. Cohen. *Ondelettes, analyses multirésolutions et filtres miroir en quadrature.* Ann. Inst. Henri Poincaré, Analyse non linéaire, Vol. 7, n°5, p.439-459, 1990.

[CohDF] A. Cohen, I. Daubechies and J.C. Feauveau. *Biorthogonal bases of compactly supported wavelets.* Preprint Bell Labs, 1989.

[CongT] G. Congedo and I. Tamanini. *On the existence of solutions to a problem in image segmentation.* Preprint, Dipartimento di Matematica, Univ. di Trento, 38050 Italy.

[DalMMS1] G. Dal Maso, J.M. Morel and S. Solimini. *Une approche variationnelle en traitement d'images : résultats d'existence et d'approximation.* C.R. Acad. Sci. Paris, t.308, Série I, p. 549-554, 1989.

[DalMMS2] G. Dal Maso, J.M. Morel and S. Solimini. *A variational method in image segmentation: existence and approximation results.* Acta Matematica Vol. 168, 1992.

[Daub1] I. Daubechies. *Orthonormal bases of compactly supported wavelets.* Comm. Pure and Applied Math., vol. 41, p.909-996, 1988.

[Daub2] I. Daubechies. *Orthonormal bases of compactly supported wavelets.* Part II, II, Preprint Bell Labs, 1989.

[DaubJJ] I. Daubechies, S. Jaffard and J.L. Journé. *A simple Wilson orthonormal basis with exponential decay.* Preprint Bell Labs, 1989.

[DeG] E. De Giorgi. *Free discontinuity problems in calculus of variations.* To appear, Proc. of the meeting in J.L. Lions's honour, Paris, 1988.

[DeGCL] E. DeGiorgi, M. Carriero and A. Leaci. *Existence theorem for a minimum problem with free discontinuity set.* Arch.Rat.Mech.Anal., Vol.108, 1988, p.195-218.

[DibK] F. Dibos and G. Koepfler. *Propriété de régularité des contours d'une image segmentée.* C.R. Acad. Sci. Paris, t. 313, Série I, p. 573-578, 1991.

[EsG] D. Esteban and C. Galand. *Application of QMF to split-band voice coding schemes.* Proc. of Int. Conf. Acoust. Speech, Signal Processing, Harford, 1977.

[Fed] H. Federer. *Geometric Measure Theory.* Springer-Verlag,1969.

[FoT] I. Fonseca and L. Tartar. *The gradient theory of phase transitions for systems with two*

potential wells. Carnegie Mellon Univ. To appear.

[FroM] J. Froment and J.M. Morel. *Analyse multiéchelle, vision stéréo et ondelettes.* Exposé n°5 du Séminaire d'Orsay "Les ondelettes en 89", ed. P.G. Lemarié, Springer Verlag 1990.

[GemG] S. Geman and D. Geman. *Stochastic relaxation, Gibbs distributions and the Baysian restoration of images.* IEEE PAMI 6, 1984.

[HarS] R.M. Haralick and L.G. Shapiro. *Image segmentation techniques.* Computer Vision Graphics and Image Processing, 29, 100-132, 1985.

[Ju] B. Julesz. *Texton gradients : the texton theory revisited.* Biological Cybernetics, 54 : 245-251, 1986.

[KasWT] M. Kass, A. Witkin and D. Terzopoulos. *Snakes: active contour models.* 1st Int.Comp.Vis.Conf. 1987, IEEE n°777.

[KoeMS] G. Koepfler, J.M. Morel and S. Solimini. *Segmentation by minimizing a functional and the "merging" methods.* Proceedings '*GRETSI* Colloque' in Juan-les-Pins, September 1991 (France).

[MaliP1] J. Malik and P. Perona *A scale space and edge detection using anisotropic diffusion.* Proc. IEEE Computer Soc. Workshop on Computer Vision, 1987.

[MaliP2] J. Malik and P. Perona. *Preattentive texture discrimination with early vision mechanisms.* Journ. of the Opt. Society. of America. A, 923-932, vol. 7, n°5, 1991.

[Mall] S.Mallat. *A theory for multiresolution signal decomposition: the wavelet representation.* IEEE Trans. on Pattern Analysis and Machine Intelligence, 674-693, July 1989.

[Marr] D. Marr. *Vision.* Freeman and Co. 1982.

[Me] Y. Meyer. *Ondelettes, fonctions splines et analyses graduées.* Rapport CEREMADE n°8703, 1987.

[MorS1] J.M. Morel and S. Solimini. *Segmentation of images by variational methods: a constructive approach.* Rev. Matematica de la Universidad Complutense de Madrid, Vol.1, n°1,2,3, 169-182, 1988.

[MorS2] JM. Morel and S. Solimini. *Segmentation d'images par méthode variationnelle : une preuve constructive d'existence.* C. R. Acad. Sci. Paris Série I, 1988.

[MumS1] D. Mumford and J. Shah. *Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems.* Communications on Pure and Applied Mathematics. vol. XLII No.4, 1989.

[MumS2] D. Mumford and J. Shah. *Boundary detection by minimizing functionals.* Image Understanding, 1988, Ed. S. Ullman and W. Richards.

[PapJ] N.T. Pappas and N.S. Jayant. *An adaptive clustering algorithm for image segmentation.* Proc. of ICCV'88, IEEE n° 883, 310-315, 1988.

[Pav1] T. Pavlidis. *Segmentation of pictures and maps through functional approximation.* Comp.Gr. and Im.Proc. 1, 360-372, 1972.

[Pav2] T. Pavlidis. *Structural Pattern Recognition.* Springer, New York 1977.

[PavL] T. Pavlidis and Y.T. Liow. *Integrating Region Growing and Edge Detection.* Proc. of the IEEE Conf. Comp. Vision and Patt. Recognition 1988.

[Ric] T.J. Richardson. *Scale independant piecewise smooth segmentation of images via variational methods.* PhD. Dissertation, Lab. for Information and Decision Systems, MIT, Cambridge MA 02139, February 1990.

[Si-L] L. Simon. *Lectures on Geometric Measure Theory.* Centre for Math. Analysis. Australian Nat. Univ., vol. 33, 1983.

[Th] C. Thomassen. *The Jordan-Schönflies theorem and the classification of surfaces.* Am.Math.Monthly, 116-130, February 1992.

[VoP] H. Voorhees and T. Poggio. *Computing texture boundaries from images.* Nature, 333, 364-367, 1988.

[Zu] S.W. Zucker. *Region growing: Childhood and Adolescence (Survey).* Comp. Graphics and Image Proc. 5, 382-399, 1976.