Segmentation by minimizing functionals and the merging methods.

G. Koepfler†, J.M. Morel† and S. Solimini‡

†CEREMADE, Université de Paris IX-Dauphine 75775 Paris cedex 16 (France)
‡SISSA, Strada Costiere 1, Trieste (Italy)

Abstract. Most segmentation algorithms are composed of several procedures: split and merge, small region elimination, boundary smoothing,..., each depending on a lot of parameters. The introduction of an energy to minimize leads to a drastic reduction of these parameters. We prove that the most primitive 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 show experimental results.

Résumé. La plupart des algorithmes de segmentation procèdent par étapes: croissance de régions, élimination de petites régions,..., ce qui introduit un grand nombre de paramètres. Nous proposons une approche par minimisation d'énergie ce qui réduit considérablement le nombre de "réglages" à faire. Même le plus primitif des algorithmes de croissance possède des propriétés comme compacité de l'ensemble des solutions, élimination des petites régions,.... Nous donnons des exemples.

1 Introduction.

General principles of segmentation devices. This note does not pretend to propose a new method for image segmentation, but rather to justify mathematically, classify, and hopefully simplify, segmentation devices. We start giving some basic principles for segmentation devices.
1) We admit the possibility of a universal boundary detection device, definable and analysable independently from the kind of channels to be used in the texture discrimination problem (see [Ju,PeMa]). This allows us to begin to get a mathematical understanding of the segmentation problem by considering grey-level segmentation.
2) An algorithm for boundary detection must be scale and space invariant. This means that multiscale segmentation algorithms,invariant by rotation and translation, should be considered.
3) The last point is what we shall call the comparison principle. Given two different segmentations of a datum we shall be able to decide which of them is better than the other. This implies the existence of some ordering on the segmentations which is reflected by some real functional E such that if $E(K_1) < E(K_2)$, then the segmentation K_1 is "better" then the segmentation K_2.

Formalization and examples. 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, if these channels are good indicators of the similarity (or difference) one can caracterize textures, histograms, colors,.... That does not change anything in the theorems and proofs we state later.

Denote by K the whole boundary set of the segmentation. According to the preceding principles, a natural energy functional for segmentation will contain at least two terms, a 2-dimensional one which measures the autosimilarity of g in the connected components of $\Omega \setminus K$ and a 1-dimensional one for controlling the length. Such a generic justification is developped in [MuShI,MuShII]. These authors propose the following functional (see also [GeGe] for an ad hoc formulation):

$$E(u,K) = \int_{\Omega \setminus K} |\nabla u|^2 dx + \int_{\Omega \setminus K} |u-g|^2 dx + \mathcal{H}^1(K),$$

this energy means that in a good segmentation (u,K), the curves K should be the boundaries of homogeneous areas and u a regularized version of g inside such areas. The third term gives control over the length of K, using the Hausdorff measure $\mathcal{H}^1$. As noticed in Zucker [Zu] and Haralick and Shapiro [HaSh], a pure region growing would generate very nonsmooth boundaries.
With a model like the above one hopes to have put all criteria together in the same functional. Many early functionals in image analysis had only 2-dimensional energy terms [PavI,Zu]. The so called "snakes" method [KaWiTe] uses a more sophisticated 1-D term for controlling location and smoothness of the boundary. The functional proposed by Blake and Zisserman [BlZi] has no term imposing the boundary to be close to high gradient pixels. Other work in this direction has been done by Terzopoulos [Te], Leclerc [Le].

Segmentation devices and their relation to energy minimizing. In a recent work Pavlidis and Liow [PavLi] propose an algorithmic integration of all the energy terms mentioned above. The question which rises is wether (and

how) energy functionals and segmentation devices match together. Our purpose is to classify the properties wich are sought by those devices and to decide which are basic and which can be deduced. To make this classification we use the tool of mathematical analysis.
Let us take the simplest Mumford and Shah energy:

$$E(K) = \int_{\Omega \setminus K} |u-g|^2 dx + \lambda \ell(K), \qquad (1)$$

where K is the union of boundaries in Ω, with Hausdorff length $\ell(K)$, and u is piecewise constant on $\Omega \setminus K$, λ is the scaling factor of the model: if it is low a lot of boundary is allowed. We note the energy $E(K)$ because in is easy to see that u is the mean value of g on every connected component defined by K.
Similar functionals have been studied by DeGiorgi's school [DGiCaLe] for liquid crystals. Let us here just quote the results obtained in [MuShII] on the minimum of (1). They showed that either the points of K are regular or they can be classified in two types: ternary crossings or perpendicular to $\partial\Omega$, moreover the boundaries of the segmentation verify inequations for curvature like

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

where u^+ and u^- denote the values of $u(x)$ on both sides of K. Two of the authors have proposed in [MoSoI] a constructive approach.
Now it is well known that functionals of the kind of (1) have many local minimizers. One has thus to choose between two strategies:
α) The global minimization is achieved by simulated annealing, this method leeds to huge and long computations but in some asymptotic sense the global minimum is attained.
β) Define some concept of local minimum which should be more accessible to fast computations and verify the same properties which are seeked for the global minimum. For instance a tendency in simulated annealing is to define some faster algorithms which do not pretend to find a global minimum [Az]. The homotopy method of [BlZi] also seeks for "good" local minima.

2 Definitions and notations.

Regions, or connected components of $\Omega \setminus K$: we shall denote them by $(O_i)_i$.
Common boundary of two areas 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: we denote it by $|O_i|$.
Isoperimetric inequality in $\mathbb{R}^2$ *and* Ω: denote by O a region in $\mathbb{R}^2$. Then one has $\ell(\partial O) \geq 2\pi^{\frac{1}{2}}|O|^{\frac{1}{2}}$. In the case of a region Ω, the same kind of inequality holds for the relative boundary of O in Ω with a smaller constant C: $\ell(\partial O) \geq C|O|^{\frac{1}{2}}$. Since there is no ambiguity, Ω being fixed, we still denote by ∂O the relative boundary of O in Ω (instead of $\partial O \cap \partial\Omega$). By $\ell(\partial O)$ we mean the length of the boundary of O. For a general definition of the length see [Fa].
Normal segmentations: a segmentation K will be called normal if every subsegmentation K' of K verifies $E(K') > E(K)$. By subsegmentation of K we mean a segmentation obtained by merging an arbitrary number of regions. The following definition is a particular case of normality (see [PavI,Zu]):

Definition 1 *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 two regions verifies*

$$E(K') > E(K).$$

3 Properties of segmentations obtained by merging.

We shall only consider segmentations having the following properties, wich are easy to check for computationnally defined segmentations, generally made of affine curves.

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

b) Every region has no internal boundary. This means that 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 is only the 1-normality of K.

We will now state a compactness result for the set of 2-normal segmentations.

Theorem 1 *For every sequence (K_n) of 2-normal 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. Thus every minimizing sequence of 2-normal segmentations provides a minimizing segmentation.

Comments.
This result may seem to be weak, since one generally hopes to get a unique solution for a given problem, or at least a finite number of solutions. However, we already know that our problem may have infinitely many local minimizers and therefore a restriction which we can hope to obtain for a given set of "solutions" to our problem is compactness. It indicates that the set of 2-normal segmentations is much smaller than the set of all possible segmentations (which is not compact). One can proof [MoSoII] that these properties are specific to the dimension 2. In dimension 3 the 2-normal segmentations are no compact set and one needs the full normality assumption to restore this compactness. Thus region growing algorithms based on recursive merging seem to be inconsistent in dimensions greater than 2.
Proof.
We will not give the complete proof of theorem 1 here, we will just state two lemmas which are important for the proof and, to our aims, give interesting properties for 2-normal segmentations (for more details see [MoSoI]).

Lemma 1 *For every region O of a 2-normal segmentation, denote by $N(O)$ the number of neighbouring regions. Then $N(O) \geq C\lambda|O|^{-\frac{1}{2}}\mathrm{osc}(g)^{-2}$, where C is the isoperimetric constant in Ω and $\mathrm{osc} = \max(g) - \min(g)$ is the oscillation of g.*

Lemma 2 *Let K be a segmentation with no internal boundaries. Then the number of regions N, and the number of edges E verify $E \leq 3(N-1)$.*

By edge we mean any connected component of the boundary common to two different regions. Thus an edge is a curve, which is either circular, ends at points where at least three regions and three edges meet or ends on $\partial\Omega$ (see [MuShII]). We impose the fact that the segmentation is 1-normal in order that K is exactly the union of all edges.

Remark 1. Elimination of small regions.
For any region O of the segmentation the proof of theorem 1 and lemma 1 give a lower bound for $|O|$ depending only on g, λ and $|\Omega|$. Therefore a merging method based on minimizing $E(K)$ should spontaneously eliminate the small regions.

Remark 2. Elimination of thin regions.
One can also deduce that the regions of a 2-normal segmentation are not too "thin". Indeed they verify an inverse isoperimetric inequality: there is a constant C depending only on g and Ω such that for every region O : $|O|^{\frac{1}{2}} \geq C|\partial O|$ (see [MoSol]). Thus the devices based on the elimination of thin regions (see [PapJa] for instance, and many clustering algorithms [PavII]) can be considered as implicit in the search of a 2-normal segmentation.

Remark 3. Smoothing of boundaries.
The 2-normal segmentations have no chance of having boundaries smooth everywhere. However, due to the length term in (1), one can show that they are almost everywhere C^1 (see [Fe,Si]). The boundaries can not increase indefinetly as it is the case for some methods. Moreover with (2) one can impose that the boundary is smoothed according to the criterion imposed by the energy. This equation shows that the length term of our simple energy model, coupled to its bidimensional contrast measuring term, is enough to ensure that the boundaries are analogous to snakes. The idea to treat region boundaries as snakes is implicit in [PavLi].

4 Pyramidal algorithms constructing 2-normal segmentations

We now consider the problem of defining and computing a 2-normal segmentation. If we follow the main ideas of region growing methods [Zu], we see that all they do is precisely compute a 2-normal subsegmentation of the initial segmentation, obtained by recursive merging.
Assume that the datum g is defined on a rectangle which is divided in small squares where g is constant (trivial segmentation). We require the following properties for the segmentations calculated by a region growing algorithm, defined as application associating to g and λ a segmentation (u, K):

- Correctedness (fixed point property): Assume g piecewise constant, the there exists λ_0 such that for all $\lambda < \lambda_0$ the segmentation (u, K), obtained for λ, verifies $u = g$ and K is union of the boundaries of the areas where g is constant (see also [Ri]).

- 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 areas of the segmentation associated to λ' are the unions of some of the areas obtained for λ.

We will explain now an algorithm computing a 2-normal subsegmentation from the trivial segmentation verifying the preceding requests.
The basic step, the merging, is done relatively to the energy (1). Let O and O' be two connected components of $\Omega \backslash K$, the decision of merging will be made according to the sign of $E(K) - E(K \backslash \partial(O, O'))$: if this quantity is positive than the merging operation decreases the global energy.
Consider an increasing sequence of parameters $\lambda_1, \lambda_2, \ldots, \lambda_k$, for example $\lambda_i = 2^i$. Let (u_0, K_0) be the trivial segmentation, define recursively a sequence (u_n, K_n) as follows: (u_{n+1}, K_{n+1}) is obtained by sweeping over the areas of $\Omega \backslash K_n$ and trying to merge every region with its neighbour giving the most energy decrease. Repeat this until there is no more merging possible using parameter λ_{n+1}, i.e. we have a 2-normal segmentation for $\lambda = \lambda_{n+1}$, then pass to parameter λ_{n+2}. We stop when we have a 2-normal segmentation for λ_k.
To make the method quite independend from machine representation of data the sequence should start with small values and little increase (in order to avoid CPU-time expensive global best-merging lists). Indeed having small weight for the length term only allows "evident" mergings and postphones less evident, but unreverseble, mergings to a bigger value of the parameter.

5 Experiments.

We tried the preceding algorithm on several pictures on a SPARC1+ (40MB). An example is given in figure 1, we used an trivial segmentation composed of 1 pixel per region. The segmentation is achieved in 100s. We used the piecewise constant model with grey-channels outlined in the preceding section.

6 Conclusion.

First we pointed out basic principles for segmentation and proposed a mathematical formalization in order to group the several tasks of a good segmentation device. Then we showed, using a simple merging algorithm, the relations between a simple energy and region growing devices.
We conclude that it is possible to unify most segmentation tools under a generalized merging algorithm based on minimization of an energy having the minimal complexity of the energy which we considered here. However we do not pretend that this is enough to characterize good segmentations, we only tried to justify that it is a necessary step. The following would be to define evolution equations generating boundaries and regions. In our opinion the correct formalization of such an equation is a good challenge.

References

[Az] R. Azencott. *Synchronous Boltzmann machines and Gibbs fields; Learning algorithms.* To appear, Springer Verlag 1989.

[BlZi] A. Blake & A. Zisserman. *Visual Reconstruction.* MIT Press, 1987.

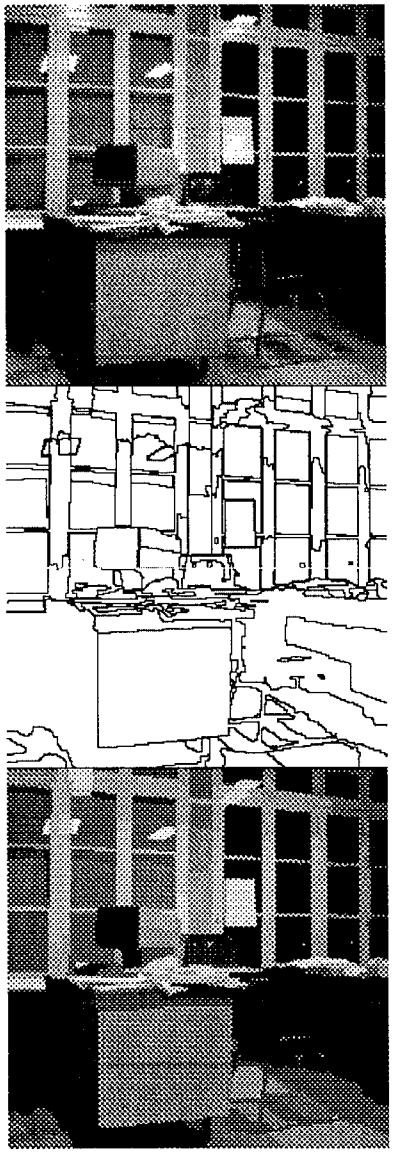

Figure 1: Original g, boundaries K, reconstrution u.

[CaDGiLe] M. Carriero, E. DeGiorgi & A. Leaci. *Existence theorem for a minimum problem with free discontinuity set.* Preprint Univ. Lecce, 1988.

[DMaMoSo] G. Dal Maso, J.M. Morel & S. Solimini. *A variational method in image segmentation: existence and approximation results.* S.I.S.S.A. 48 M, 1989 (to appear in Acta Matematica 90).

[DGiCaLe] E. DeGiorgi, M. Carriero & A. Leaci. *Existence theorem for a minimum problem with free discontinuity set.* Preprint Univ. Lecce, 1988.

[Fa] K.J. Falconer. *The Geometry of Fractal Sets.* Cambridge University Press,1985.

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

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

[HaSh] R.M. Haralick & L.G. Shapiro. *Image segmentation techniques.* Comp.Vi.Gr. and Im.Pr. 29, 100-132, 1985

[Ju] B. Julesz *Texton gradients: the texton theory revisited.* Biol. Cyb.,54,254-251, 1986.

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

[Le] Y. Leclerc. *Constructing stable descriptions for image partitioning.* I.J.C.Vision 3, 73-102, 1989.

[MoSoI] J.M. Morel & S. Solimini. *Segmentation of images by variational methods: a constructive approach.* Rev. Matematica de la Universidad Complutense de Madrid, Vol.1, no.1,2,3, 169-182, 1988.

[MoSoII] J.M. Morel & S. Solimini. *Segmentation of multidimensional data: on the number of regions obtained by "split and merge" methods.* To appear.

[MuShI] D. Mumford & J. Shah. *Boundary detection by minimizing funcionals.* Proc. IEEE Conf. C.V.P.R., San Francisco 1985.

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

[PapJa] N.T. Pappas & N.S. Jayant *An adaptive clustering algorithm for image segmentation.* Proc. of ICCV 88, IEEE n° 883, 310-315, 1988.

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

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

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

[PeMa] P. Perona & J. Malik. *A computational model of texture perception* Preprint.

[Ri] T.J. Richardson. *Scale independant piecewise smooth segmentation of images via variational methods.* PhD. MIT 1991.

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

[Te] D. Terzopoulos. *Regularization of inverse visual problems involving discontinuities.* IEEE PAMI-8,4, July 1986.

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