

# Bayesian Object Detection through Level Curves Selection

Charles Kervrann

INRA - Biométrie, Domaine de Vilvert  
78352 Jouy-en-Josas, France  
ck@jouy.inra.fr

**Abstract.** Bayesian statistical theory is a convenient way of taking *a priori* information into consideration when inference is made from images. In Bayesian image detection, the *a priori* distribution should capture the knowledge about objects. Taking inspiration from [1], we design a prior density that penalizes the area of homogeneous parts in images. The detection problem is further formulated as the estimation of the set of curves that maximizes the posterior distribution. In this paper, we explore a posterior distribution model for which its maximal mode is given by a subset of level curves, that is the boundaries of image level sets. For the completeness of the paper, we present a stepwise greedy algorithm for computing partitions with connected components.

## 1 Introduction

In most problems of image analysis, incorporation of prior knowledge is important for making inference based on the images. Bayesian object detection is the problem of how to estimate the number of simply connected objects and their location in a non-ideal environment. Bayesian approaches specify ways for segmenting the entire image using global energy criteria. Indeed, it is usually straightforward to transfer a Bayesian criterion into an energy minimization criterion. In addition, additivity is desirable in models which must be analyzing by Markov Chains Monte Carlo sampling. Thereby, the discrete [8,2] or continuous [18,17] energy functional is traditionally designed as a combination of several terms, each of them corresponding to a precise property which must be satisfied. While this modeling offers a powerful theoretical framework and minimizers exist [17,25], they have several disadvantages. First, these models lead to very difficult optimization problems that are notoriously slow to converge [2]. Second, the weight parameters which are key ingredients of a wide range of segmentation energies, are usually not correctly estimated, yielding to supervised segmentation methods. Third, sampling from a Markov Random Fields distribution does not always produce patterns that look like images.

Object detection belongs to the field of high-level imaging, in which the image modeling is on a more global scale compared to low-level imaging which deals with (smoothing) prior models on a pixel level. In particular, more global prior

models for the simply connected objects can be applied. There has been a growing interest in this field, particularly along the guidelines of Grenander’s general pattern theory using deformable templates [9]. Moreover, Zhu *et al.* attempted to unify snakes [12], balloons [5] and region growing methods within a general energy/Bayes framework [25]. These prior models are generally realistic and incorporate prior information about the outline of objects in a Bayesian image analysis framework. In other respects, both approaches estimate the curves that maximally separate unknown statistics inside and outside the curves [25,9]. The *maximum a posteriori* (MAP) estimate is generally determined by prohibitive stochastic search procedures [9] or other variants of steepest ascent algorithms [25]. Thereby, additional *a priori* knowledge may be specified to ease the segmentation task: statistics inside region boundaries are assumed to be known [9,4] or estimated using *ad-hoc* methods [20,24]. The global energy functional may be then optimized, for instance, within a level set framework [19,22] which provides the advantages of numerical stability and topological flexibility [4,20,24]. In practical imaging, these methods may suffer from the problem of initialization of curves [25], off-line estimation of the mixture model of Gaussians approximating the probability density function of the image [20], or selection of hyperparameters weighting the contribution of energy terms [25,20,4].

In this paper, we address these problems and follow the Bayesian approach for recovering simply connected objects in the plane. The prior model focuses on how the area and number of objects can varied in images (Section 2). It allows to partition the image into few regions, though in a more restrictive manner than previous approaches [25,20] since it can generate irregular boundaries. Unlike other approaches [25,9], we shall see that maximizing the posterior distribution is herein equivalent to select a subset of connected components of image bilevel sets (Section 3). Section 4 presents the numerical implementation of our model and the computation of the image segmentation. In Section 5, we illustrate this approach with some experiments on satellite images. Conclusions and perspectives are presented in Section 6.

## 2 The Bayesian Framework

Let  $S$  be an open subset of  $\mathbb{R}^2$  and  $f$  a grey-scale image treated as a function defined on  $S$ . In practical imaging,  $S$  is a collection of pixels within a discretized rectangle, and possible values of  $f$  are given by integers  $[0, 256[\cap\mathbb{N}$ . Below we will work in the continuous setup, where  $S$  is a subset of a Euclidian space and  $f : S \rightarrow \mathbb{R}^+$  represents the observed data function. The continuous setup allows us to refer to analytic tools, while leaving always a possibility to “discretize” the problem. We use the terminology “site” or “pixel” to denote a point of the image, even in the continuous case. Each point  $x \in S$  is assigned a grey value  $f(x)$ . According to Matheron [15], we interpret the image  $f$  as a family of sets defined by  $L_\gamma(f) = \{x \in S : f(x) \geq \gamma\}$ ,  $\gamma \in \mathbb{R}^+$ . Each level sets  $L_\gamma(f)$  is assumed to be of finite perimeter. Therefore,  $f$  will belong to the bounded variation (noted  $BV$ ) space [1].

Let  $\{\Omega_i \subset S\}$  be a set of disjoint and non-empty image domains or objects, and  $\{\partial\Omega_i\}$  their boundaries. A partition of the space  $S$  consists in finding a set  $\{\Omega_i\}_{i=1}^P$  and a background  $\overline{\Omega}$  defined as the complementary subset of the union of objects  $\overline{\Omega} = S \setminus \cup_{i=1}^P \Omega_i$ ,  $\Omega_i \cap_{i \neq j} \Omega_j = \emptyset$  and  $\Omega_i \cap_i \overline{\Omega} = \emptyset$ .

We assume that the observed image  $f$  has been produced by the model  $f = f_{\text{true}} + \varepsilon$ , where  $\varepsilon$  is a zero-mean Gaussian white noise:  $\varepsilon(x) \stackrel{iid}{\sim} \mathcal{N}(0, \sigma^2)$ ,  $x \in S$ . The true image  $f_{\text{true}}(x) = \sum_{i=1}^P \overline{f}_{\Omega_i} \mathbf{1}_{x \in \Omega_i} + \overline{f}_{\overline{\Omega}} \mathbf{1}_{x \in \overline{\Omega}}$  is supposed piecewise constant, where  $\overline{f}_{\Omega_i}$  and  $\overline{f}_{\overline{\Omega}}$  denote respectively the unknown average values of  $f$  over  $\Omega_i$  and  $\overline{\Omega}$ , and  $\mathbf{1}_{x \in E}$  is the set indicator function of the set  $E$ . The variance  $\sigma^2$  is assumed to be known and constant over the entire image [25]. So, the likelihood for the data  $f$  given  $\{\Omega_1, \dots, \Omega_P\}$  is specified by

$$p(f | \Omega_1, \dots, \Omega_P) \propto \exp -\frac{1}{2\sigma^2} \left\{ \sum_{i=1}^P \int_{\Omega_i} (f(x) - \overline{f}_{\Omega_i})^2 dx + \int_{\overline{\Omega}} (f(x) - \overline{f}_{\overline{\Omega}})^2 dx \right\}. \quad (1)$$

We seek a partition of the rectangle  $S$  into a finite set of objects  $\Omega_i$ , each of which corresponding to a part of the image where  $f$  is constant. Therefore, we define the following collection  $\mathcal{C}_P$  of  $P \geq 0$  admissible, closed and connected objects

$$\mathcal{C}_P = \{ \{ \Omega_1, \dots, \Omega_P \} \subset S ; S \setminus \overline{\Omega} = \bigcup_{i=1}^P \Omega_i ; \Omega_i \cap_{1 \leq i \neq j \leq P} \Omega_j = \emptyset \}.$$

When  $P = 0$ , there is no object in the image. Following the Bayesian approach, we use some functional of the posterior distribution of  $\{\Omega_1, \dots, \Omega_P\}$ :

$$p(\Omega_1, \dots, \Omega_P | f) \propto p(f | \Omega_1, \dots, \Omega_P) \pi(\Omega_1, \dots, \Omega_P) \quad (2)$$

where  $p(f | \Omega_1, \dots, \Omega_P)$  is the likelihood given by (1) and  $\pi(\Omega_1, \dots, \Omega_P)$  is the prior distribution of objects. The posterior distribution is used in a further inferential issue concerning the objects within the Bayesian paradigm. The *a priori* distribution should capture the knowledge about  $\{\Omega_1, \dots, \Omega_P\}$ . We define a density that penalizes the area  $|\Omega_i|$  of objects. Additionally, the variables  $\{|\Omega_i|\}$  may be considered as independent random variables with density  $g(|\Omega_i|)$ . Hence, the prior distribution is of the form  $\pi(\Omega_1, \dots, \Omega_P) = Z_P^{-1} \prod_{i=1}^P g(|\Omega_i|)^a$  where  $Z_P$  is a normalization constant and  $a$  a real positive value. The density  $g(|\Omega_i|)$  is chosen to be a non-negative monotonically decreasing function of the object area  $|\Omega_i|$ . For instance, Alvarez *et al.* [1] have realized experimentally that the area distribution of homogeneous parts in images follows a power law  $\beta |\Omega_i|^{-\gamma}$ . The parameters  $\beta$  and  $\gamma$  (close to 1,2 for values of  $|\Omega_i|$  in a certain range) give the intensity of the model. In what follows, we shall consider this model for the density  $g(|\Omega_i|)$ . There are other possible choices of  $g(|\Omega_i|)$ ; the case of  $g(|\Omega_i|) \propto \exp -\beta |\Omega_i|^\gamma$  has been already discussed in [13,1]. This model is related to the Markov connected components fields [16].

### 3 Bayesian Inference

All kinds of inference are made from  $p(\Omega_1, \dots, \Omega_P | f)$ . Finding the *maximum a posteriori* (MAP) estimate is herein our choice of inference. As a consequence, the MAP estimation of objects comes to the minimization of a global energy function  $E_\lambda(f, \Omega_1, \dots, \Omega_P)$  defined as

$$\underbrace{\sum_{i=1}^P \int_{\Omega_i} (f(x) - \bar{f}_{\Omega_i})^2 dx + \int_{\bar{\Omega}} (f(x) - \bar{f}_{\bar{\Omega}})^2 dx}_{E_d(f, \Omega_1, \dots, \Omega_P)} + \lambda \underbrace{\sum_{i=1}^P (\gamma \log(|\Omega_i|) - A)}_{E_p(\Omega_1, \dots, \Omega_P)} \quad (3)$$

where  $E_p(\Omega_1, \dots, \Omega_P)$  is the penalty functional,  $E_d(f, \Omega_1, \dots, \Omega_P)$  the data model,  $\lambda = 2a\sigma^2 > 0$  the regularization parameter and  $A = \log(\beta)$ . The penalty functional tends to regulate the emergence of objects  $\Omega_i$  in the image and gives no control on the smoothness of boundaries. The regularization parameter  $\lambda$  can be then interpreted as a scale parameter that only tunes the number of regions [17,14,13]. If  $\lambda = 0$ , each point is potentially a region and  $\bar{\Omega} = \emptyset$ ; the global minimum coincides with zero and this segmentation is called the “*trivial segmentation*” [17,14].

Our MAP estimator is defined by (when exists)

$$(\hat{\Omega}_1, \dots, \hat{\Omega}_{\hat{P}}) = \operatorname{argmin}_{0 \leq P \leq T} \operatorname{argmin}_{\{\Omega_1, \dots, \Omega_P\} \in \mathcal{C}_P} E_\lambda(f, \Omega_1, \dots, \Omega_P) \quad (4)$$

where  $\mathcal{C}_P \subseteq \mathcal{C}_T, \forall P \leq T$ , and  $T$  is the maximum number of admissible objects registered in a bank  $\mathcal{C}_T$ . We recall that  $\bar{\Omega} = S \setminus \cup_{i=1}^{\hat{P}} \hat{\Omega}_i$  is the complementary subset of estimated objects  $\{\hat{\Omega}_1, \dots, \hat{\Omega}_{\hat{P}}\}$ . By using classical arguments on lower semi-continuous functionals on the  $BV$  space, we assume here existence of minimizers of  $E_\lambda(f, \Omega_1, \dots, \Omega_P)$  among functions of sets finite perimeter (or of bounded variation) [17,25]. However, a direct minimization with respect to all unknown domains  $\Omega_i$  and parameters  $\bar{f}_{\Omega_i}$  is a very intricate problem, even if  $T$  is low since objects are not designed. In what follows (Lemma 1), we prove that the object boundaries that minimize  $E_\lambda(f, \Omega_1, \dots, \Omega_P)$  are level lines of the function  $f$ , which makes the problem tractable.

**LEMMA 1.** *If there exists minimizers and no pathological minimum exists, then the energy minimizing set of curves is a subset of level lines of  $f$ :*

$$f|_{\partial \hat{\Omega}_i} \equiv \mu_i, \quad i = 1, \dots, \hat{P}.$$

*i.e. the border  $\partial \hat{\Omega}_i$  of each  $\hat{\Omega}_i$  is a boundary of a connected component of a level set of  $f$ .*

*Proof of Lemma 1* Let  $\Omega_\delta$  be a variation of a set  $\Omega$ , i.e. the Hausdorff distance  $d_\infty(\Omega_\delta, \Omega) \leq \delta$ . To prove Lemma 1, we assume that, for any connected perturbation of  $\Omega$  such  $d_\infty(\Omega_\delta, \Omega) \leq \delta$ , two neighboring sets  $\Omega$  and  $\Omega'$  do not merge into one single set  $\Omega \cup \Omega'$  and, for any connected perturbation of  $\Omega$  such  $d_\infty(\Omega_\delta, \Omega) \leq \delta$ ,  $\Omega$  does not split into two new sets. This corresponds to

prohibited topological changes. Without loss of generality, we prove Lemma 1 for one object  $\Omega$  and a background  $\overline{\Omega}$ , that is the closure of the complementary set of  $\Omega$ . For two sets  $A$  and  $B$ , denote  $\int_{A \setminus B} f \stackrel{\text{def}}{=} \int_A f - \int_B f$ . Then, we have

$$\int_{\Omega_\delta \setminus \Omega} 1 \stackrel{\text{def}}{=} |\Omega_\delta| - |\Omega| \quad \text{and} \quad \left( \int_{\Omega_\delta} f \right)^2 - \left( \int_{\Omega} f \right)^2 = 2 \int_{\Omega} f \int_{\Omega_\delta \setminus \Omega} f + \left( \int_{\Omega_\delta \setminus \Omega} f \right)^2. \quad (5)$$

The difference between the involved energies is defined as  $\Delta E_\lambda(f, \Omega) = E_\lambda(f, \Omega_\delta) - E_\lambda(f, \Omega) = T_1 + T_2 + T_3 + T_4 + T_5$  where

$$\begin{aligned} T_1 &= \int_{\Omega_\delta} f^2 - \int_{\Omega} f^2, \quad T_2 = \frac{1}{|\Omega|} \int_{\Omega_\delta \setminus \Omega} \lambda \gamma, \quad T_3 = -\frac{1}{|\Omega_\delta|} \left( \int_{\Omega_\delta} f \right)^2 + \frac{1}{|\Omega|} \left( \int_{\Omega} f \right)^2, \\ T_4 &= \int_{S \setminus \Omega_\delta} f^2 - \int_{S \setminus \Omega} f^2, \quad T_5 = -\frac{1}{|S| - |\Omega_\delta|} \left( \int_{S \setminus \Omega_\delta} f \right)^2 + \frac{1}{|S| - |\Omega|} \left( \int_{S \setminus \Omega} f \right)^2. \end{aligned} \quad (6)$$

Denote  $\Delta|\Omega| = |\Omega_\delta| - |\Omega|$ . Using (5), and passing to the limit  $\Delta|\Omega| \rightarrow 0$ , i.e.  $|\Omega_\delta| \simeq |\Omega|$ , we obtain (higher order terms are neglected)

$$\begin{aligned} T_1 = -T_4 &= \int_{\Omega_\delta \setminus \Omega} f^2, \quad T_2 = \frac{\lambda \gamma}{|\Omega|} \int_{\Omega_\delta \setminus \Omega} 1, \\ T_3 &= -\frac{2}{|\Omega|} \int_{\Omega_\delta \setminus \Omega} f \int_{\Omega} f - \frac{1}{|\Omega|} \left( \int_{\Omega_\delta \setminus \Omega} f \right)^2 + \frac{1}{|\Omega|^2} \int_{\Omega_\delta \setminus \Omega} 1 \left( \int_{\Omega} f \right)^2, \\ T_5 &= \frac{1}{|S| - |\Omega|} \left\{ 2 \int_{\Omega_\delta \setminus \Omega} f \int_{S \setminus \Omega} f - \left( \int_{\Omega_\delta \setminus \Omega} f \right)^2 - \frac{1}{|S| - |\Omega|} \int_{\Omega_\delta \setminus \Omega} 1 \left( \int_{S \setminus \Omega} f \right)^2 \right\}. \end{aligned} \quad (7)$$

We define the following image moments  $m_0 = \int_{\Omega} 1$ ,  $m_1 = \int_{\Omega} f$ ,  $K_0 = \int_S 1$ ,  $K_1 = \int_S f$ . Using the *mean value theorem for double integral*, which states that if  $f$  is continuous and a connected subset  $E$  is bounded by a simple curve, then for some point  $x_0$  in  $E$  we have  $\int_E f(x) dE = f(x_0) \cdot |E|$  where  $|E|$  denotes the area of  $E$ , it follows that

$$\begin{aligned} \Delta E_\lambda(f, \Omega) &= \left\{ \overbrace{\left[ \frac{m_1^2}{m_0^2} - \frac{(K_1 - m_1)^2}{(K_0 - m_0)^2} + \frac{\lambda \gamma}{m_0} \right]}^{M_0} + \overbrace{\left[ \frac{2(K_1 - m_1)}{K_0 - m_0} - \frac{2m_1}{m_0} \right]}^{M_1} f(x_0) \right. \\ &\quad \left. - \left[ \frac{1}{m_0} + \frac{1}{K_0 - m_0} \right] f(x_0)^2 \int_{\Omega_\delta \setminus \Omega} 1 \right\} \int_{\Omega_\delta \setminus \Omega} 1. \end{aligned} \quad (8)$$

Let  $x_b$  be a fixed point of the border  $\partial\Omega$ . Choose  $\Omega_\delta$  such that  $\partial\Omega_\delta = \partial\Omega$  except on a small neighborhood of  $x_b$ . The energy having a minimum for  $\Omega$ ,  $f(x_b)$  needs to be solution of the following equation

$$\lim_{\Delta|\Omega| \rightarrow 0} \frac{\Delta E_\lambda(f, \Omega)}{\Delta|\Omega|} = [M_0 + M_1 f(x_b)] + O(\Delta|\Omega|) = 0. \quad (9)$$

By passing to the limit  $\Delta|\Omega| \rightarrow 0$ , we obtain  $M_0 + M_1 f(x_b) = 0$ . This equation has one single solution. The coefficients  $M_0$  and  $M_1$  do depend on neither  $x_b$  nor  $f(x_b)$ , and  $M_0 \neq 0$ . The function  $f$  is continuous and  $\partial\Omega$  is a connected curve. Therefore  $f(x_b)$  is constant when  $x_b$  covers  $\partial\Omega$ . This completes the proof.  $\square$

We have proved Lemma 1 with a connected perturbation including the situation when  $|\Omega_\delta| - |\Omega| = |\partial\Omega|$  where  $|\partial\Omega|$  is the boundary length of  $\Omega$ . Equation 9 states a *necessary condition* which is essential to prove that a subset of level lines globally minimizes the energy.

If  $f$  is of bounded variation, the connected components of level sets can be characterized by their boundaries, that is the so-called *level lines* of  $f$  [3]. In consequence of Lemma 1, those curves constitute the borders  $\{\partial\Omega_i\}$  of objects  $\{\Omega_i\}$ .

## 4 A Stepwise Greedy Algorithm for Image Segmentation

This section describes our algorithmic procedure for object boundaries estimation. Our recommendations for the concrete choice of the input parameters are collected in this section. The algorithm we propose does not require neither the number of regions nor any initial mean gray values for regions and background.

### 4.1 Level Sets and Object Boundaries

The key ingredient of the procedure is the construction of objects whose boundaries are image level lines [3]. In practical imaging, we can associate with an image 255 level sets  $\{L_\gamma(f)\}$ ,  $0 \leq \gamma \leq 255$ . We consider the scenario where a point  $x$  belongs to one single connected component at once within the image level sets. We take into account this fact and define the bilevel sets of  $f$  as the set of pixels  $x \in S$  such that  $v \leq f(x) \leq w$ ,  $0 \leq v \leq w$ . Instead of computing all the 255 level sets, we restrict only this computation to a small number of  $K (< 255)$  level sets and adaptively quantize the image histogram using an entropy method [11]. For  $l \in \mathbb{N}$  varying from 1 to  $K$ , let  $b_l$  be the binary image with  $b_l(x) = 1$  if  $f(x) \in [t_{l-1}, t_l)$  and  $b_l(x) = 0$  otherwise, where  $t_l$  is a threshold. We call those images the  $K$ -bilevel sets of  $f \in [f_{\min}, f_{\max}]$  [1]. In general, each bilevel set is made up of  $n(t_l)$  disjoint connected components, where  $n(t_l)$  is a function of the threshold  $t_l$  and  $S = \cup_{l=1}^{l=K} [\Omega_{t_l,1} \cup \Omega_{t_l,2} \cup \dots \cup \Omega_{t_l,n(t_l)}]$ . A crude way to build pixels sets corresponding to objects would be to proceed to a connected components labeling of binary images  $\{b_l\}$ ,  $1 \leq l \leq K$ , and to associate each label with an object  $\Omega_i$ .

If  $f$  is bounded, the connected components of level sets can be characterized by their surrounding curves, that is the level lines [3,1]. If we map these level lines for a given set of  $K$  levels, we get a segmentation of the image also called *topographic map* [3,7]. More generally, one can consider a segmentation achieved using only some connected components of level sets, which is the philosophy of our approach. The most perceptible level lines can be determined by an isoperimetric criterion [7] or the detection of T-junctions of level lines [3]. Both criteria

are strong indicators of region boundaries. Instead, we use herein a simpler criterion where perceptually significant level lines are the level sets boundaries of an quantized image by using  $K$  quantizers and an entropy method [11]. Entropy methods seek to maximize the information content between objects and background pixels of an image. The method due to Kapur *et al.* chooses the thresholds  $\{t_l\}$  to be the values at which the information is maximum. As a consequence, the detection of meaningful level lines will depend on the quantization parameter  $K$ . Unlike previous criteria [3,7], this quantization operation is not invariant to contrast changes. Nevertheless, we shall see that, in practice,  $K = \{4, \dots, 8\}$  seems sufficient to detect physically meaningful objects in the image.

## 4.2 The Segmentation Procedure

The proposed algorithm is not a region growing algorithm as described in [14,17] since all objects are built once and for all. Although our work is related to morphological approaches based on connected operators [21,6,10], it is an independent approach since we seek minimizers of a global objective functional. In addition, it differs from the watershed approach since regions that emerge from the watershed segmentation are not necessarily connected components within the image level sets [23].

We post-process the connected components to remove any components whose surface area  $|\Omega_i|$  is less than some threshold  $|\Omega_{\min}|$  (a parameter of the method) to eliminate regions corresponding to noise and artifacts in the original image [21,6,10]. To implement our level set image segmentation based on energy minimization, a four step method is used. Let  $K, \lambda, |\Omega_{\min}|$  be the input parameters set by the user.

*1. Bilevel Set Construction.* The first step completes a crude mapping of each image pixel on a given bilevel set. At present, we quantize the function  $f \in [f_{\min}, f_{\max}]$  in  $K = \{4, \dots, 8\}$  non-equal-sized and non-overlapping intervals  $[t_{l-1}, t_l)$ ,  $l = \{1, \dots, K\}$ . Given this set of intervals estimated using the maximum entropy sum method [11], let  $b_l$  be the bilevel set image with  $b_l(x) = 1$  if  $f(x) \in [t_{l-1}, t_l)$  and  $b_l(x) = 0$  otherwise.

*2. Object Extraction.* A crude way to build pixels sets corresponding to objects is to proceed to a connected components labeling of images  $\{b_l\}$  and to associate each label with an object  $\Omega_i$ . Though this process may work in the noise-free case, in general we would also need some smoothing effect of the connected components labeling. So we consider a size-oriented morphological operator acting on sets that consists in keeping all connected components of the output of area larger than a limit  $|\Omega_{\min}|$ . This *connected operator* in mathematical morphology will never introduce new features or edges and boundaries of remained connected components are preserved [21,6,10]. The list of connected components then forms the bank  $\mathcal{C}_T$  of admissible objects  $\{\Omega_1, \dots, \Omega_T\}$  with  $|\Omega_i| \geq |\Omega_{\min}|$ .

*3. Configuration Determination.* The connected components are then combined during the third step to form object configurations. For instance, these configurations can be built by enumeration of all possible object combinations, i.e.  $2^T$  configurations. Each configuration is made of a subset of objects taken in the bank  $\{\Omega_1, \dots, \Omega_T\}$ . The background  $\overline{\Omega}$  corresponds to the complementary set of objects selected for each configuration.

*4. Energy Computation and Object Configuration Selection.* Energy calculations take the image intensities of the original (not quantized) image to establish piecewise-constant approximation errors. Energies of the form  $\{\int_{\Omega_i} (f(x) - \overline{f}_{\Omega_i})^2 dx\}$  are computed once and stored on a RAM memory. The energy term  $\int_{\overline{\Omega}} (f(x) - \overline{f}_{\overline{\Omega}})^2 dx$  is efficiently updated for each configuration since  $\overline{\Omega}$  is the complementary subset of the union of objects  $\{\Omega_i\}_{i=1}^P$ . The configuration that globally minimizes the energy functional corresponds to the MAP segmentation. The time necessary to perform image segmentation essentially depends on the size of the object bank  $\mathcal{C}_T$ .

### 4.3 Computational Issues

Now we discuss how some parameters of the procedure can be selected and indicate one possible choice used in our experimental results. On the discrete domain  $S$ , the neighborhoods of a pixel  $x$  are typically defined via 4-connectivity or 8-connectivity.

Number of Bilevel Sets. The value of  $K$  is mainly determined by the number of meaningful objects that one wishes to extract and the computational effort one is able to spend. Decreasing  $K$  allows to reduce the number of connected component. In our approach, we determine the optimal configuration of objects by supervising a small set of levels. In practice, our approach successfully segmented various images into only 4 or 8 levels.

Minimal Area of Objects. The area-oriented operator affects the image by remaining connected components within the image level sets that do not satisfy the minimum criterion [21,6,10]. Boundaries of connected components are not distorted by this operator as occurs with other types of image filters (such as openings and closings using structuring elements). Our default choice is  $|\Omega_{\min}| \in [0.0001, 0.001] \times |S|$ .

Prior Parameters  $A$  and  $\gamma$ . For fixed  $K$ , we consider the sets of observations  $\{\log(|\Omega_i|), \log(g(|\Omega_i|)), 1 \leq i \leq T\}$ . We perform a linear regression on this set so as to find the straight line (in the log-log coordinates)  $\log(g(|\Omega_i|)) = A - \gamma \log(|\Omega_i|)$  the closest to the data in the least squares sense [1].

Hyperparameter  $\lambda$ . The choice of this parameter determines mostly the properties of the segmentation result. Increasing this parameter reduces the final number of objects to be extracted. If  $f$  is a function from  $S$  to  $[0, 255]$ , a default choice for the hyperparameter is  $\lambda \in [0.1, 1.] \times 255^2$ . Of course larger values of  $\lambda$  lead to even extraction of only one object. In practice, it's possible for us to tune this parameter according to image contents.

Energy Minimization. For a fixed bank  $\mathcal{C}_T = \{\Omega_1, \dots, \Omega_T\}$  of  $T$  objects, one way to choose the optimal set of objects  $\{\widehat{\Omega}_1, \dots, \widehat{\Omega}_{\widehat{P}}\}$ ,  $\widehat{P} \leq T$ , is to search for all possible combinations of  $P$  objects and compute the corresponding energy  $E_\lambda(f, \Omega_1, \dots, \Omega_P)$ . Enumerating all possible sets of objects in the object bank and comparing their energies is computationally too expensive if  $T$  is large (typically, it is infeasible if  $T > 32$ ). Instead of a such brute force search, we propose the following stepwise greedy algorithm for minimizing  $E_\lambda(f, \Omega_1, \dots, \Omega_P)$ .

We start from  $P = 0$  and introduce one object  $\Omega_j$  at a time. Energies of all objects are assumed to be already stored in a RAM memory. At the first step, we compute the  $T$  energies with one single object  $\Omega_j$  at once against the complementary subset  $\overline{\Omega} = S \setminus \bigcup_{j \neq i=1}^T \Omega_i$ . Let  $\widehat{\Omega}_1$  be the estimated object that best lowers  $E_\lambda$ . This object is stored on a RAM memory as an object of the optimal configuration. It is removed from the initial bank  $\mathcal{C}_T$ . At any steps of the algorithm, a new object is chosen to maximally decrease the energy  $E_\lambda$ .

Suppose that at the  $P$ -th step,  $\widehat{P}$  and  $\widehat{\Omega}$  are not known but we have estimated  $P$  objects  $\{\widehat{\Omega}_1, \dots, \widehat{\Omega}_P\}$  and a current background  $\overline{\Omega} = S \setminus \{\widehat{\Omega}_1, \dots, \widehat{\Omega}_P\}$ . Let  $E_\lambda(f, \widehat{\Omega}_1, \dots, \widehat{\Omega}_P)$  be the current computed energy. Then at the  $(P+1)$ -th step, we choose the object  $\Omega_j \in \mathcal{C}_T \setminus \{\widehat{\Omega}_1, \dots, \widehat{\Omega}_P\}$  which has the maximal difference, i.e.

$$\widehat{\Omega}_{P+1} = \arg \max_{\Omega_j \in \mathcal{C}_T \setminus \{\widehat{\Omega}_1, \dots, \widehat{\Omega}_P\}} E_\lambda(f, \widehat{\Omega}_1, \dots, \widehat{\Omega}_P, \Omega_j) - E_\lambda(f, \widehat{\Omega}_1, \dots, \widehat{\Omega}_P) \quad (10)$$

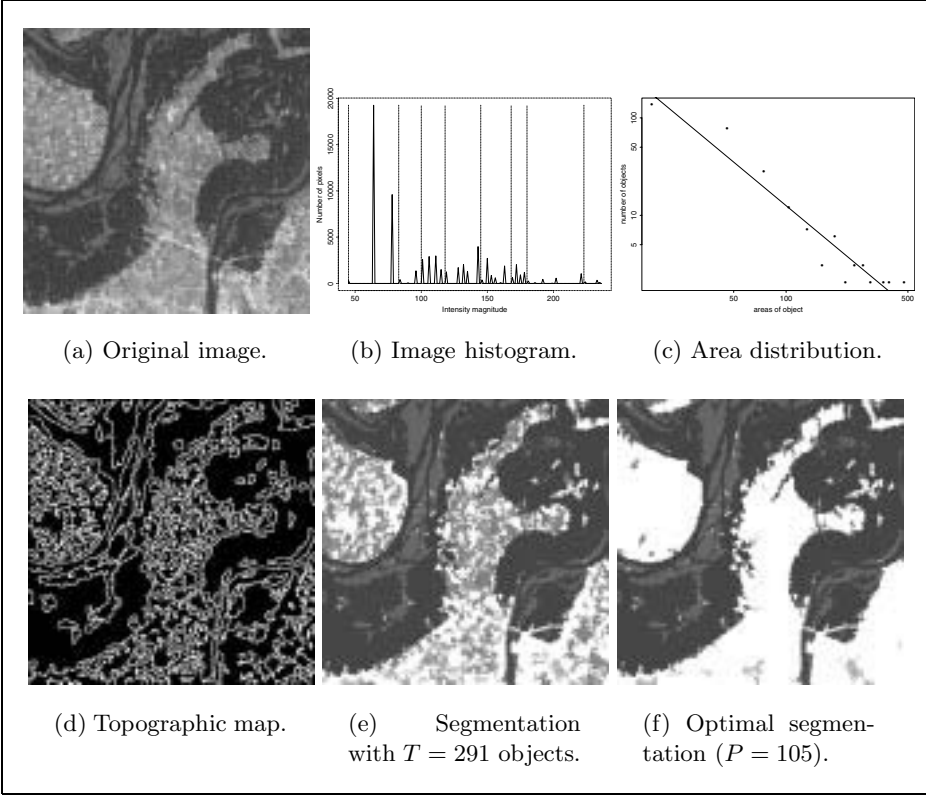
The algorithm stops at  $P$ -th step when the adding of any object does not decrease  $E_\lambda$ . This means that the optimal number of objects is  $\widehat{P} = P$  and the remained objects of the bank are a part of the estimated background, i.e.  $\widehat{\Omega} = S \setminus \{\widehat{\Omega}_1, \dots, \widehat{\Omega}_{\widehat{P}}\}$ .

This fast algorithm selects a suboptimal configuration of objects corresponding to a local minima of the energy functional. Using this algorithm,  $\frac{T \times (T+1)}{2}$  object configurations are examined at the most, whereas the supervision of all the configurations corresponds to  $2^T$  global energy computations.

## 5 Experimental Results in Image Segmentation

Experiments were conducted on satellite and meteorological images to evaluate the performance of the algorithm. Recall that obtaining the most meaningful objects is the goal of this work. For this reason,  $K$  was set fairly low in the experiments ( $K = 4$  or  $K = 8$ ) to obtain large regions and to improve robustness to noise and artifacts in the image. Regions which areas  $|\Omega_i| < [0.0001, 0.001] \times |S|$  are discarded. For our method,  $\lambda$  varies across the images depending on the image contents. It is set empirically and values that gave visually better results were chosen. Most segmentations took approximately about 1-15 seconds on a 296MHz workstation.

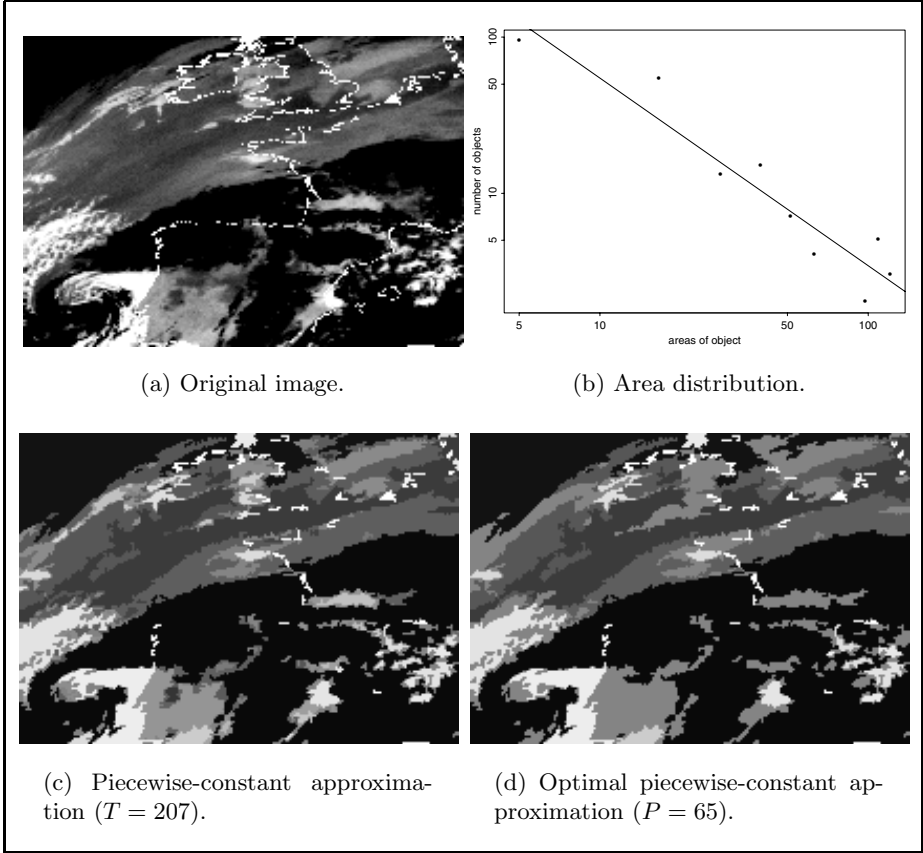
Figure 1a shows an aerial  $256 \times 256$  image (in the visual spectrum) depicting the region of Saint-Louis during the rising of the Mississippi and Missouri rivers in July 1993. We are interested in extracting the rivers and a background



**Fig. 1.** Satellite image ( $K = 8$ ,  $|\Omega_{min}| = 0.00025 \times |S|$ ,  $\lambda = 0.25 \times 255^2$ ).

corresponding to textured urban areas. Figure 1 shows the segmentation results when  $K = 8$ ,  $|\Omega_{min}| = 0.00025 \times |S|$  and  $\lambda = 0.25 \times 255^2$ . In this experiment, the maximum number of significant components is  $T = 291$  (Fig. 1d-e). The connected components that do not satisfy the minimum area criterion are labeled “white” in Fig. 1e. The image histogram has been quantized with  $K = 8$  quantizers and an entropic method (Fig. 1b). We estimated the values of  $A = 3.727$  and  $\gamma = 1.486$  by fitting a straight line  $\log(g(|\Omega_i|)) = A - \gamma \log(|\Omega_i|)$  to the observed data by linear regression. In that case, the least squares error is 2.007 and  $17 \leq |\Omega_i| \leq 2.886 \cdot 10^4$  pixels. Figure 1f displays the crudely piecewise-constant approximation results by setting  $\lambda = 0.25 \times 255^2$ . It takes about 15 seconds ( $25095 < \frac{T \times (T+1)}{2} = 42486$  iterations) of computing time for building  $\mathcal{C}_T$  and selecting the best configuration ( $P = 105$  objects) using the stepwise greedy algorithm. Enumerating all the configurations is infeasible since  $2^T = 3.9810^{87}$  iterations ! The non-connected background is labeled in “white” in Fig. 1f and the objects are filled with their mean gray values  $\{\bar{f}_{\Omega_i}\}$ .

The performance of the segmentation procedure is demonstrated for a satellite ( $210 \times 148$ ) image shown in Fig. 2. For the set of parameters  $K = 4$ ,



**Fig. 2.** Meteorological image ( $K = 4$ ,  $|\Omega_{\min}| = 0.00015 \times |S|$ ,  $\lambda = 0.75 \times 255^2$ ).

$\lambda = 0.75 \times 255^2$  and  $|\Omega_{\min}| = 0.00015 \times |S|$ , the algorithm selected  $P = 65$  objects from the bank which contains  $T = 213$  objects (Fig. 2d). The piecewise-constant approximation of the image using  $T = 213$  objects is shown in Fig. 2c. The algorithm stopped at the 11765<sup>th</sup> iteration (2s of CPU time), i.e. before the maximal iteration  $\frac{T \times (T+1)}{2} = 22791$ . We performed a linear regression to estimate the values of  $A$  and  $\gamma$  ( $5 \leq |\Omega_i| \leq 1.159 \cdot 10^4$ ):  $A = 1.482$ ,  $\gamma = 1.202$ . This corresponds to a least squares error of 1.34.

## 6 Conclusion and Perspectives

In this paper, we have presented a Bayesian approach for extracting structures in images. The prior model penalizes the area of the homogeneous parts of the image. Morphological approaches based on connected operators have already applied a such criterion but the filtering/segmentation process is not generally based on the optimization of a global objective functional. In addition, we proved

that our MAP estimator can be determined by selecting a subset of image level lines. A total CPU time of a few seconds using a suboptimal stepwise greedy algorithm for partitioning a  $256 \times 256$  image into meaningful regions makes the method attractive for many time-critical applications. In terms of future directions for research, we propose to create a non-linear scale-space by successive applications of an area morphology operator to select most meaningful regions in the image.

## References

1. L. Alvarez, Y. Gousseau, and J.M. Morel. Scales in natural images and a consequence on their bounded variation. In *Int. Conf. on Scale-Space Theories Comp. Vis.*, pages 247–258, Kerkyra, Greece, September 1999.
2. A. Blake and A. Zisserman. *Visual Reconstruction*. MIT Press, Cambridge, Mass, 1987.
3. V. Caselles, B. Coll, and J.M. Morel. Topographic maps and local contrast changes in natural images. *Int J. Computer Vision*, 33(1):5–27, 1999.
4. T. Chan and L. Vese. Active contour model without edges. In *Int. Conf. on Scale-Space Theories Comp. Vis.*, pages 141–151, Kerkyra, Greece, September 1999.
5. L.D. Cohen. On active contour models and balloons. *CVGIP: Image Understanding*, 53(2):211–218, 1991.
6. J. Crespo, R. Schafer, J. Serra, C. Gratin, and F. Meyer. The flat zone approach: a general low-level region merging segmentation method. *Signal Processing*, 62(1):37–60, 1997.
7. J. Froment. Perceptible level lines and isoperimetric ratio. In *Int. Conf. on Image Processing*, Vancouver, Canada, 2000.
8. S. Geman and D. Geman. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Trans. Patt. Anal. and Mach. Int.*, 6(6):721–741, 1984.
9. U. Grenander and M.I. Miller. Representations of knowledge in complex systems. *J. Royal Statistical Society, series B*, 56(4):549–603, April 1994.
10. R. Jones. Connecting filtering and segmentation using component trees. *Computer Vision and Image Understanding*, 75(3):215–228, 1999.
11. J.N. Kapur, P.K. Sahoo, and A.K.C. Wong. A new method for gray-level picture thresholding using the entropy of the histogram. *Comp. Vis. Graphics and Image Proc.*, 29:273–285, 1985.
12. M. Kass, A. Witkin, and D. Terzopoulos. Snakes: active contour models. *Int J. Computer Vision*, 12(1):321–331, 1987.
13. C. Kervrann, M. Hoebeke, and A. Trubuil. Level lines as global minimizers of energy functionals in image segmentation. In *Euro. Conf. on Comp. Vis.*, pages 241–256, Dublin, Ireland, June 2000.
14. G. Koepfler, C. Lopez, and J.M. Morel. A multiscale algorithm for image segmentation by variational method. *SIAM J. Numerical Analysis*, 31(1):282–299, 1994.
15. G. Matheron. *Random Sets and Integral Geometry*. John Wiley, New York, 1975.
16. J. Møller and R.P. Waagepetersen. Markov connected component fields. *Adv. in Applied Probability*, pages 1–35, 1998.
17. J.M. Morel and S. Solimini. *Variational methods in image segmentation*. Birkhauser, 1994.

18. D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and variational problems. *Communication on Pure and applied Mathematics*, 42(5):577–685, 1989.
19. S. Osher and J. Sethian. Fronts propagating with curvature dependent speed: algorithms based on the hamilton-jacobi formulation. *J. Computational Physics*, 79:12–49, 1988.
20. N. Paragios and R. Deriche. Coupled geodesic active regions for image segmentation: a level set approach. In *Euro. Conf. on Comp. Vis.*, pages 224–240, Dublin, Ireland, June 2000.
21. P. Salembier and J. Serra. Flat zones filtering, connected operators, and filters by reconstruction. *IEEE Trans. Image Processing*, 4(8):1153–1160, 1995.
22. J. Sethian. *Level Sets Methods: Evolving Interfaces in Geometry, Fluid Mechanics, Computer Vision, and Material Science*. Cambridges University Press, 1996.
23. L. Vincent and P. Soille. Watershed in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Trans. Patt. Anal. and Mach. Int.*, 13(6):583–598, 1991.
24. A. Yezzi, A. Tsai, and A. Willsky. A statistical approach to snakes for bimodal and trimodal imagery. In *Int. Conf. on Comp. Vis.*, pages 898–903, Kerkyra, Greece, September 1999.
25. S.C Zhu and A. Yuille. Region competition: unifying snakes, region growing, and bayes/mdl for multiband image segmentation. *IEEE Trans. Patt. Anal. and Mach. Int.*, 18(9):884–900, 1996.