HAL open science

Smooth Curve Extraction by Mean Field Annealing

Laurent Herault, Radu Horaud

To cite this version:

Laurent Herault, Radu Horaud. Smooth Curve Extraction by Mean Field Annealing. Annals of Mathematics and Artificial Intelligence, 1995, 13 (3-4), pp.281–300. 10.1007/BF01530832 . inria-00590040

HAL Id: inria-00590040

https://inria.hal.science/inria-00590040v1

Submitted on 10 Aug 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Annals of Mathematics and Artificial Intelligence 13(1995) 281–300

Smooth Curve Extraction by Mean Field Annealing*

Laurent Hérault
LETI (CEA-Technologies Avancées)
Département Systèmes
CENG, 17 avenue des Martyrs
38054 Grenoble FRANCE

Radu Horaud
LIFIA–IMAG
46, avenue Félix Viallet
38031 Grenoble FRANCE

Abstract

In this paper we attack the figure-ground discrimination problem from a combinatorial optimization perspective. In general the solutions proposed in the past solved this problem only partially: either the mathematical model encoding the figure-ground problem was too simple or the optimization methods that were used were not efficient enough or they could not guarantee to find the global minimum of the cost function describing the figure-ground model. The method that we devised and which is described in this paper is tailored around the following contributions. First, we suggest a mathematical model encoding the figure-ground discrimination problem that makes explicit a definition of shape (or figure) based on cocircularity, smoothness, proximity, and contrast. This model consists of building a cost function on the basis of image element interactions. Moreover, this cost function fits the constraints of a interacting spin system which in turn is a well suited physical model to solve hard combinatorial optimization problems. Second, we suggest a combinatorial optimization method for solving the figure-ground problem, namely mean field annealing which combines mean field approximation and annealing. Mean field annealing may well be viewed as a deterministic approximation of stochastic methods such as simulated annealing. We describe in detail the theoretical bases of this method, derive a computational model, and provide a practical algorithm. Finally, some experimental results are shown both with synthetic and real images.

I Introduction

The problem of separating figure from ground is a central one in computer vision. One aspect of this problem is the problem of separating shape from noise. Two-dimensional shapes are the input data of high-level visual processes such as recognition. In order to maintain the complexity of recognition as low as possible it is important to determine at an early level what is shape and

*This research has been sponsored in part by "Commissariat à l'Energie Atomique," and in part by the ORASIS project (PRC Communications Homme/Machine).

what is noise. Therefore one needs a definition of shape, a definition of noise, and a process that takes as input *image elements* and separates them into shape and noise.

In this paper we suggest an approach whose goal is twofold: (i) it groups image elements that are likely to belong to the same (locally circular) shape while (ii) noisy image elements are eliminated. More precisely, the method that we devised builds a cost function over the entire image. This cost function sums up image element interactions and it has two terms, i.e., the first enforces the grouping of image elements into shapes and the second enforces noise elimination. Therefore the shape/noise discrimination problem becomes a combinatorial optimization problem, namely the problem of finding the global minimum for the cost function just described. In theory, the problem can be solved by any combinatorial optimization algorithm that is guaranteed to converge towards the global minimum of the cost function.

In practice, we implemented three combinatorial optimization methods: simulated annealing (SA), mean field annealing (MFA), and microcanonical annealing (MCA) [8]. Here we concentrate on mean field annealing.

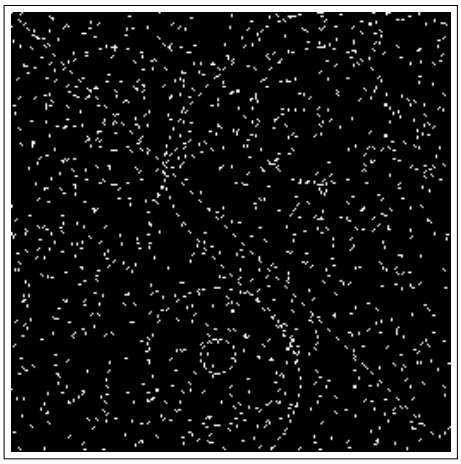

Figure 1: A synthetic image with 1250 elements. Circles, a straight line, and a sinusoid are plunged into randomly generated elements.

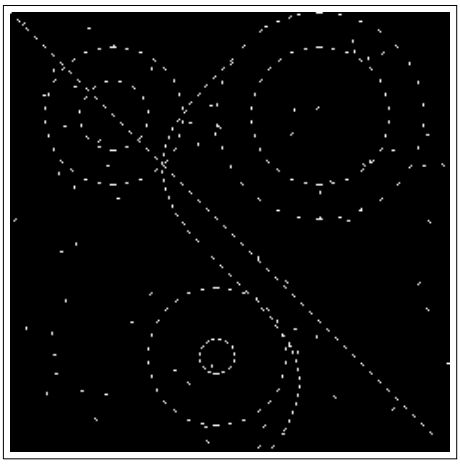

Figure 2: The result of applying mean field annealing to the synthetic image. The elements shown on this figure were labelled "shape".

The figure-ground or shape/noise separation problem is best illustrated by an example. Figure 1 shows a synthetic image. In this image some elements belong to such shapes as circles, a straight line, and a sinusoid while some other elements are noise. Two independent sequences of random numbers were used to generate this noise: One sequence for the position in the image and another sequence for the orientation. This image has a total of 1250 elements. Figure 2 shows the result obtained with mean field annealing. This image contains 309 elements that were labelled "shape" by this algorithm. Notice that edges belonging to such generic shapes as circles, straight lines, and sinusoids were correctly labelled as shapes while most of the noise was thrown out. Also, there are a number of isolated noise edgels that were incorrectly labelled shape: these isolated edgels can be easily eliminated by some straightforward post-processing.

II Background and contribution

The interest for shape/noise separation stems from Gestalt psychologists' figure-ground demonstrations [12]: certain image elements are organized to produce an emergent figure. Ever since the figure-ground discrimination problem has been seen as a side effect of feature grouping. The prerequisite of any method is the extraction from the raw data of the image elements that are either to be grouped or to be thrown out. Edge detection is in general first performed. Grouping edges is done on the basis of their connectivity or by using a clustering technique. Noise is eliminated by thresholding.

The connectivity analysis produces edge chains which are further fitted with piecewise analytic curves (lines, conics, splines, etc.). The clustering technique maps image edges into a parameter space and such a parameter space is associated with each curve type: this is the well known Hough transform. The problem of finding curves in the image is equivalent to a clustering problem in parameter space. With this technique, curves can be detected even if their constituting edges are not connected in the image. There are two major problems with the techniques just described:

1. One has to specify analytically the type of curve or curves that are to be sought in the image and this is done at a low-level of the visual process. At an early stage of an image interpretation system one may just want to know whether a set of edges form a shape or not, without any specific knowledge about the analytic curve that may fit this shape. For example, recognition of object prototypes needs such a vague description.

2. The notion of *noise* is not clearly specified with respect to the notion of shape. Hence noise elimination takes the form of adhoc methods. For example, edges associated with low contrast are usually thrown out. There are many examples that tend to prove that this is a bad heuristic. Indeed, a low contrasted shape may be present in an image in the same time as high-contrasted noise. Or, a shape may be composed of low-contrasted edges and high-contrasted ones. Simply eliminating the low-contrasted edges will have as an effect the truncation and/or elimination of shapes.

Parent & Zucker [14] define the notion of shape locally on the basis of curvature computed on a discrete grid and map the curve inference problem into a global optimization problem. The optimization itself is carried out by relaxation labelling. Finally, curve points are labelled "1" and noise points are labelled "0". We did not compare relaxation labelling with stochastic optimization and therefore we cannot assert which method is the best. However, relaxation labelling is a local optimization process which needs good initialization. The analysis of discrete curvature provided by Parent & Zucker is very interesting but the structure of the associated cost function is not well suited for such methods as mean-field annealing.

A similar approach was followed by Shashua & Ullman [18]. Curve inference takes the form of searching for the best sequence of edge elements in order to form the longest and the smoothest curves that pass through each image point. The search itself is carried out by dynamic programming which is not a global optimization technique. Therefore a global minimum cannot be guaranteed unless a good initialization is provided.

Curve/noise separation is also the concern of Gutfinger & Sklansky [7]. The coding of the problem is inspired from [14]. Curve/noise separation is then viewed as a classification problem. The classification of dots into curve-dots and noise-dots is carried out by "mixed adaptation", a

method combining supervised and unsupervised training. The training stage computes statistics on noise images and on curve images. Unfortunately this a priori separation is possible only with simulated data and the method becomes impracticable when applied to real images.

The advantage of stochastic optimization over more classical approaches for solving the figure-ground discrimination problem was stressed by Sejnowski & Hinton [17]. They introduce a figure-ground model based on two possible labels for the image elements: region and edge. Starting with a random labelling, a gradient-descent procedure gets trapped in one of the many local minima of the energy landscape while simulated annealing converges to a solution where the region elements are "inside" the edge elements. The experimental results shown by the authors deal only with synthetic data.

The method proposed by Carnevali, Coletti, & Patarnello [3] uses simulated annealing and a simple pixel interaction model in order to classify the pixels of a binary image into object and noise. The pixel interaction model they propose uses pixel proximity on the premise that objects are dense sets of pixels while noise is formed of a sparsely distributed set of pixels.

The method proposed by Peterson [16] for tracking particles in high energy physics may be applied to the grouping problem. Peterson suggests a combinatorial optimization formulation based on mean field theory. However the model proposed by Peterson for track finding is not well suited when noise is present in the data. Moreover, Peterson's model maps n points onto n^2 variables and hence there are n^4 connections. Such a model is well suited when the number of variables is relatively small. Our model maps n points onto n variables and therefore only n^2 connections are necessary. More interesting, a VLSI implementation of the mean field theory is suggested in Peterson's paper.

Recently, Kass, Witkin, and Terzopoulos [11] proposed the snake model: it is an energy-minimizing spline guided by external forces and images forces that pull the snake towards image edges. Nevertheless, snakes do not try to solve the entire problem of finding salient image contours. They rely on other mechanisms to place them nearby the desired contours.

The work described in this paper has the following contributions:

- We suggest a mathematical encoding of the figure-ground discrimination problem that consists of separating shape from noise using combinatorial optimization methods. The particular cost (or energy) function that we devised fits the constraints of a interacting spin system – a physical model well suited for solving hard combinatorial optimization problems.

- We suggest a combinatorial optimization method for solving the figure-ground problem, namely mean field annealing which combines mean field approximation theory and annealing. Mean field annealing may well be viewed as a deterministic approximation of stochastic methods such as simulated annealing.

In the past, the theoretical bases of mean field annealing were studied by Orland [13] and Peterson [15]. Mean field annealing has been used by Hérault & Niez to solve NP-complete graph problems [10], by Geiger & Girosi to solve the reconstruction problem [5], by Geiger & Yuille to solve the image segmentation problem [6], and by Zerubia & Chellappa to solve the edge detection problem [21].

The remainder of this paper is organized as follows. Section III briefly describes the mathematical structure of a class of cost functions that are well suited for stochastic optimization methods.

This structure places strong constraints on the mathematical coding of the figure-ground discrimination problem. Next we formulate our problem in terms of such a cost function. This function involves interactions between the image elements that are considered. These interactions are made clear in Section IV. Section V describes in detail the physical basis of the mean field approximation theory and introduces the mean field annealing algorithm (MFA). Section VI shows the experimental results obtained with both synthetic and real images. Section VII contains a general discussion on figure-ground discrimination and gives some directions for future work.

III A combinatorial optimization formulation

We consider a particular class of combinatorial optimization problems for which the cost function has a mathematical structure that is analog to the global energy of a complex physical system, that is a *interacting spin system*. First, we briefly describe the state of such a physical system and give the mathematical expression of its energy. We also show the analogy with the energy of a recursive neural network. Second we suggest that the figure-ground discrimination problem can be cast into a global optimization problem of the type mentioned above.

The state of an interacting spin system is defined by: (i) A spin state-vector of N elements $\vec{\sigma} = [\sigma_1, \cdots, \sigma_N]$ whose components are described by discrete labels which correspond to *up* or *down* Ising spins: $\sigma_i \in \{-1, +1\}$. The components σ_i may well be viewed as the outputs of binary neurons. (ii) A symmetric matrix J describing the interactions between the spins. These interactions may well be viewed as the synaptic weights between neurons in a network. (iii) A vector $\vec{\delta} = [\delta_1, \cdots, \delta_N]$ describing an external field in which the spins are plunged.

Therefore, the interacting spin system has a "natural" neural network encoding associated with it which describes the microscopic behaviour of the system. A macroscopic description is given by the energy function which evaluates each spin configuration. This energy is given by:

$$E(\sigma_1, \sigma_2, \ldots, \sigma_N) = -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} J_{ij} \sigma_i \sigma_j - \sum_{i=1}^{N} \delta_i \sigma_i \tag{1}$$

The main property of interacting spin systems is that at low temperatures the number of local minima of the energy function grows exponentially with the number of spins. Hence the adequation between the mathematical model of interacting spin systems and combinatorial optimization problems with many local minima is natural.

We consider now N image elements. Each such element has a label associated with it, p_i, which can take two values: 0 or 1. The set of N labels form the state vector $\vec{p} = [p_1, \cdots, p_N]$. We seek a state vector such that the "shape" elements have a label equal to 1 and the "noise" elements have a label equal to 0. If c_{ij} designates an interaction between elements i and j, one may write by analogy with physics an interaction energy: $E_{saliency}(\vec{p}) = -\sum_{i=1}^{N} \sum_{j=1}^{N} c_{ij} p_i p_j$

Obviously, the expression above is minimized when all the labels are equal to 1. In order to avoid this trivial solution we introduce the constraint that some of the elements in the image are *not* significant and therefore should be labelled "noise": $E_{constraint}(\vec{p}) = \left(\sum_{i=1}^{N} p_i\right)^2$

The function to be minimized could be something like the sum of these energies:

$$E(\vec{p}) = E_{saliency}(\vec{p}) + \lambda \, E_{constraint}(\vec{p}) \tag{2}$$

In this expression λ is a positive real parameter that has to be adjusted and is closely related to the signal-to-noise ratio. With the substitution $p_i = (\sigma_i + 1)/2$ in eq. (2), the function to be minimized is given by eq. (1) where: $J_{ij} = (c_{ij} - \lambda)/2$ and $\delta_i = (\sum_{j=1}^{N} c_{ij} - N\lambda)/2$.

IV Computing image interactions

An image array contains two types of information: changes in intensity and local geometry. Therefore the choice of the *image elements* mentioned so far is crucial. Edge elements, or edgels are the natural candidates for making explicit the two pieces of information just mentioned.

An edgel can be obtained by one of the many edge detectors now available. An edgel i is characterized by its position in the image (x_i, y_i) and by its gradient computed once the image has been low-pass filtered. The x and y components of the gradient vector are:

$$g_x(x_i, y_i) = \frac{\partial I_f(x_i, y_i)}{\partial x}$$

$$g_y(x_i, y_i) = \frac{\partial I_f(x_i, y_i)}{\partial y}$$

I_f is the low-pass filtered image. From the gradient vector one can easily compute the gradient direction and magnitude. The direction θ_i of the edgel is perpendicular to the gradient direction. It is also the direction of the tangent to the curve that may pass through this edgel, e.g., Figure 3. The magnitude of the gradient g_i is proportional to the height of the intensity change at the edgel location:

$$\theta_i = \arctan\left(\frac{g_y(x_i, y_i)}{g_x(x_i, y_i)}\right) + \frac{\pi}{2}$$

$$g_i = \left(g_x^2(x_i, y_i) + g_y^2(x_i, y_i)\right)^{1/2}$$

Let A and B be two edgels. We want that the interaction between these two edgels encapsulates the concept of shape. That is, if A and B belong to the same shape then their *interaction* is high. Otherwise their interaction is low. Notice that a weak interaction between two edgels has several interpretations: (i) A belongs to one shape and B belongs to another one, (ii) A belongs to a shape and B is noise, or (iii) both A and B are noise. The interaction coefficient must therefore be a co-shapeness measure. In our approach, co-shapeness is defined by a combination of *cocircularity, smoothness, proximity*, and *contrast*.

The definition of cocircularity is derived from [14] and it constrains the shapes to be as circular as possible, or as a special case, as linear as possible. Smoothness enforces shapes with low curvature. Proximity restricts the interaction to occur in between nearby edgels. As a consequence, cocircularity and smoothness are constrained to be local shape properties. The combination of cocircularity, smoothness, and proximity will therefore allow a large variety of shapes that are circular and smooth only locally. Contrast enforces edgels with a high gradient module to have a higher interaction coefficient than edgels with a low gradient module.

Cocircularity. Following [14] and from Figure 3 it is clear that two edgels belong to the same circle if and only if:

$$\lambda_i + \lambda_j = \pi \tag{3}$$

In this formula, λ_i is the angle made by one edgel with the line joining the two edgels. Notice that a circle is uniquely defined if the relative orientations of the two edgels verify the equation above. This equation is also a local symmetry condition consistent with the definition of local symmetry of Brady & Asada [1]. Moreover, linearity appears as a special case of cocircularity, namely when $\lambda_i = 0$ and $\lambda_j = \pi$ or when $\lambda_i = \pi$ and $\lambda_j = 0$.

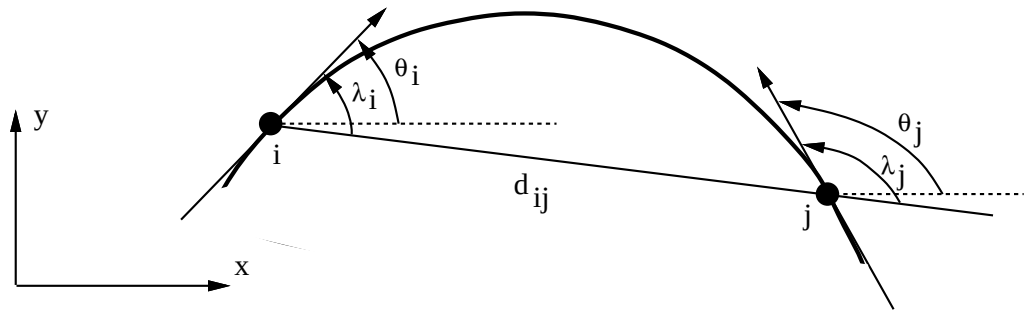

Figure 3: The definition of cocircularity between two edgels i and j.

From this cocircularity constraint we may derive a weaker constraint which will measure the similarity between a two-edgel-configuration and a circular shape:

$$\Delta_{ij} = \mid \lambda_i + \lambda_j - \pi \mid .$$

Δ_{ij} will vary between 0 (a perfect circle) and π (no shape). Finally the cocircularity coefficient is constrained to vary between 1 for a circle and 0 for noise and is defined by the formula:

$$c_{ij}^{cocir} = \left(1 - \frac{\Delta_{ij}^2}{\pi^2}\right) \exp\left(-\frac{\Delta_{ij}^2}{k}\right) \tag{4}$$

The parameter k is chosen such that the cocircularity coefficient vanishes gently for non-circular interactions.

Smoothness. Consider an edgel and two nearby edgels. From Figure 4 it is clear that the cocircularity coefficient between edgels i and j is the same as the cocircularity coefficient between i and k. Indeed there are two circles, one passing through i and j and the other passing trhough i and k. One would like to give some preference to one of these configurations. If smooth curves (rather than rapidly turning curves) are prefered then the $i-j$ interaction should be stronger than the $i-k$ interaction.

With the same definition for λ as above we define a smoothness coefficient between two image edgels. This coefficient varies between 0 (a sharp corner) and 1 (a straight line):

$$c_{ij}^{smooth} = \left(1 - \frac{\lambda_i(\pi - \lambda_i)}{\pi^2}\right)\left(1 - \frac{\lambda_j(\pi - \lambda_j)}{\pi^2}\right) \tag{5}$$

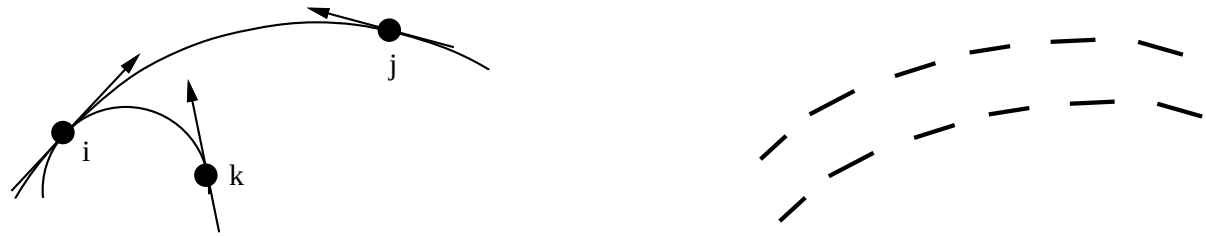

Figure 4: In the absence of the smoothness constraint, the interaction between i and j is the same as the interaction between i and k (left). The smoothness constraint reduces the interaction in between parallel curves (right).

Obviously, for the example of Figure 4-left we obtain:

$$c_{ij}^{smooth} > c_{ik}^{smooth}$$

Figure 4-right illustrates another advantage of using a coefficient that enforces smoothness. Indeed, the absence of the smoothness constraint, edgel interactions occur between parallel curves: an edgel equally interacts with its longitudinal neighbours and with its *lateral* neighbours. The smoothness constraint reduces the lateral interactions occuring between such parallel curves.

Proximity. The surrounding world is not constituted only by circular shapes. Cocircularity must therefore be a local property. That is, the class of shapes we are interested to detect at a given scale of resolution are shapes that can be approximated by a sequence of smoothly connected circular arcs and straight lines. The proximity constraint is best described by a coefficient that vanishes smoothly as the two edgels are farther away from each other:

$$c_{ij}^{prox} = \exp\left(-\frac{d_{ij}^2}{2\,\sigma_d^2}\right) \tag{6}$$

where d_{ij} is the distance between the two edgels and σ_d is a fraction of the standard deviation of all these distances over the image. Hence, the edgel interaction will adjust itself to the image distribution of the edgel population.

Contrast. A classical approach to figure-ground discrimination is to compare the gradient value at an edgel against a threshold and to eliminate those edgels that fall under this threshold. An improvement to this simply-minded nonlinear filtering is to consider two thresholds such that edgel connectivity is better preserved [2]. Following the same idea, selection of shapes with high contrast can be enforced by multiplying the interaction coefficient with a term whose value depends on contrast:

$$c_{ij}^{contrast} = \frac{g_i g_j}{g_{max}^2} \tag{7}$$

g_{max} is the highest gradient value over the edgel population.

Finally the interaction coefficient between two edgels becomes:

$$c_{ij} = c_{ij}^{cocir}\; c_{ij}^{smooth}\; c_{ij}^{prox}\; c_{ij}^{contrast} \tag{8}$$

V Mean field annealing (MFA)

The states reachable by the system described by eq. (1) correspond to the vertices of a N-dimensional hypercube. We are looking for the state which corresponds to the absolute minimum of the energy function. Typically, when $N = 1000$, the number of possible configurations is $2^N \approx 10^{301}$. The problem of finding the absolute minimum is complex because of the large number of local minima of the energy function and hence this problem cannot be tackled with local minimization methods (unless a good initialization is available).

We already mentioned that the functional to be minimized has the same structure as the global energy of an interacting spin system. To find a near ground state of such a physical system we will use statistical methods. Two analysis are possible depending on the interaction of the system with its environment: either the system can exchange heat with its environment (case of the *canonical* analysis) or the system is isolated (case of the *microcanonical* analysis) [8]. We will consider here the canonical analysis.

This analysis makes the hypothesis that the physical system can exchange heat with its environment. At the equilibrium, statistical thermodynamics shows that the free energy F is minimized. The free energy is given by: $F = E - TS$, where E is the internal energy (the energy associated with the optimization problem) and S is the entropy (which measures the internal disorder). Hence, there is a competition between E and S. At low temperatures and at equilibrium, F is minimal and TS is close to zero. Therefore, the internal energy E is minimized. However the minimum of E depends on how the temperature parameter decreases towards the absolute zero. It was shown that annealing is a very good way to decrease the temperature.

We are interested in physical systems for which the internal energy is given by eq. (1). The remarks above are expressed in the most fundamental result of statistical physics, the Boltzmann (or Gibbs) distribution:

$$Pr(E(\vec{\sigma}) = E_i) = \frac{\exp\left(-E_i/(kT)\right)}{Z(T)} \tag{9}$$

which gives the probability of finding a system in a state i with the energy E_i, assuming that the system is at equilibrium with a large heat bath at temperature $k\,T$ (k is the Boltzmann's constant). $Z(T)$ is called the *partition function* and is a normalization factor:

$$Z(T) = \sum_n \exp\left(-E_n/(kT)\right)$$

This sum runs over all possible spin configurations. Using eq. (9) one can compute at a given temperature T the mean value over all possible configurations of some macroscopic physical parameter A:

$$\langle A \rangle = \sum_n A_n \, Pr(E(\vec{\sigma}) = E_n) = \frac{\sum_n A_n \, \exp\left(-E_n/(kT)\right)}{Z(T)} \tag{10}$$

Unfortunately, the partition function $Z(T)$ is usually impossible to compute. Nevertheless, when the system is described by eq. (1) one can use eq. (10) which, with an additionnal hypothesis, is a basic equation for mean field approximation and for mean field annealing.

In order to introduce the mean field annealing algorithm, we first introduce the somehow more classical mean field approximation method which has been used to solve optimization problems

[13, 15, 9]: it is a simple analytic approximation of the behaviour of interacting spin systems in thermal equilibrium. We start by developing eq. (1) around σ_i:

$$E(\sigma_1, \sigma_2, \ldots, \sigma_N) = \Phi_i \, \sigma_i - \frac{1}{2} \sum_{k=1, k \neq i}^{N} \sum_{j=1, j \neq i}^{N} J_{kj} \sigma_k \sigma_j - \sum_{k=1, k \neq i}^{N} \delta_k \sigma_k$$

Φ_i is the total field which affects the spin σ_i:

$$\Phi_i = - \left(\sum_{j=1}^{N} J_{ij} \, \sigma_j + \delta_i \right)$$

The *mean field* $\langle \Phi_i \rangle$ affecting σ_i is computed from the sum of the fields created on the spin σ_i by all the other spins "frozen" in their mean states, and of the external field δ_i viewed by the spin σ_i. The mean state of a spin $\langle \sigma_i \rangle$ is the mean value of the σ_i's computed over all possible states that may occur at the thermal equilibrium. We obtain:

$$\langle \Phi_i \rangle = - \left(\sum_{j=1}^{N} J_{ij} \, \langle \sigma_j \rangle + \delta_i \right) \tag{11}$$

We introduce now the following approximation [19]: The system composed of N interacting spins is viewed as the union of N systems each composed of a single spin. Such a single-spin system $\{\sigma_i\}$ is subject to the mean field $\langle \Phi_i \rangle$ created by all the other single-spin systems. Let us study such a single spin system. It has two possible states: $\{-1\}$ or $\{+1\}$. The probability for the system to be in one of these states is given by the Boltzmann distribution law, eq. (9):

$$P(X_i = \sigma_i^0) \approx \frac{\exp\left(-\langle \Phi_i \rangle \, \sigma_i^0 / T\right)}{P(X_i = 1) + P(X_i = -1)}, \;\; \sigma_i^0 \in \{-1, 1\} \tag{12}$$

X_i is the random variable associated with the value of the spin state. Notice that in the case of a single-spin system the partition function (the denominator of the expression above) has a very simple analytical expression. By combining eq. (10), (11), and (12), the mean state of σ_i can now be easily derived:

$$\begin{aligned} \langle \sigma_i \rangle &\approx \frac{(+1) \, \exp\left(-\langle \Phi_i \rangle / T\right) + (-1) \, \exp\left(\langle \Phi_i \rangle / T\right)}{\exp\left(\langle \Phi_i \rangle / T\right) + \exp\left(-\langle \Phi_i \rangle / T\right)} \\ &= \tanh\left(\frac{\sum_{j=1}^{N} J_{ij} \, \langle \sigma_j \rangle + \delta_i}{T} \right) \end{aligned} \tag{13}$$

We consider now the whole set of single-spin systems. We therefore have N equations of the form:

$$\mu_i = \tanh\left(\frac{\sum_{j=1}^{N} J_{ij} \, \mu_j \; + \; \delta_i}{T} \right) \tag{14}$$

where $\mu_i = \langle \sigma_i \rangle$. The problem of finding the mean state of interacting spin system at thermal equilibrium is now mapped into the problem of solving a system of N coupled non-linear equations, i.e., eq. (14). In the general case, an analytic solution is rather difficult to obtain. Instead, the

solution for the vector $\vec{\mu} = [\mu_1, \cdots, \mu_N]$ may well be the *stationary solution* of the following system of N differential equations:

$$\tau \frac{d\mu_i}{dt} = \tanh\left(\frac{\sum_{j=1}^{N} J_{ij}\,\mu_j \,+\, \delta_i}{T}\right) - \mu_i \tag{15}$$

where τ is a time constant introduced for homogeneity. In the discrete case the temporal derivative term can be written as:

$$\left.\frac{d\mu_i}{dt}\right)_{t_n} = \frac{\mu_i^{n+1} - \mu_i^n}{\Delta t} + o(\Delta t) \tag{16}$$

where μ_i^n is the value of μ_i at time t_n. By substituting in eq. (15) and by choosing $\tau = \Delta t$, we obtain an iterative solution for the system of differential equations described by eq. (15):

$$\forall\, i \in \{1, \ldots, N\},\ \mu_i^{n+1} = \tanh\left(\frac{\sum_{j=1}^{N} J_{ij}\,\mu_j^n \,+\, \delta_i}{T}\right),\ n \geq 1 \tag{17}$$

where μ_i^n is an estimation of $\langle \sigma_i \rangle$ at time t_n or at the n'th iteration.

Starting with an initial solution $\vec{\mu^0} = [\mu_1^0, \cdots, \mu_N^0]$ the convergence is reached at the n'th iteration such that $\vec{\mu^n} = [\mu_1^n, \cdots, \mu_N^n]$ becomes stationary. The physicists would say that the thermal equilibrium has been nearly reached. When the vector $\vec{\delta}$ is the null vector, one can start with a vector $\vec{\mu^0}$ close to the obvious unstable solution $[0, \cdots, 0]$. In practice, even when the vector $\vec{\delta}$ in not null one starts with an initial configuration obtained by adding noise to $[0, \cdots, 0]$. For instance the μ_i^0's are chosen randomly in the interval $[-10^{-5}, +10^{-5}]$. This initial state is plausible for an interacting spin system in a heat bath at high temperature. In fact the spin values (+1 or –1) are equally likely at high temperatures and hence, $\langle \sigma_i \rangle \ = 0$ for all spin i. During the iterative process, the μ_i's converge to values in between -1 or $+1$.

Two convergence modes are possible:

- *Synchronous mode.* At each step of the iterative process **all** the μ_i^n's are updated using the μ_i^{n-1}'s previously calculated.
- *Asynchronous mode.* At each step of the iterative process a spin μ_i^n is randomly selected and updated using the μ_i^{n-1}'s.

In practice, the asynchronous mode produces better results because the convergence process is less subject to oscillations frequently encountered in synchronous mode. In order to obtain a solution for the vector $\vec{\sigma}$ from the vector $\vec{\mu}$, one simply looks at the signs of the μ_i^n's. A positive sign implies that the probability that the corresponding spin has a value of +1 is greater than 0.5:

$$\begin{cases} \text{if } 1/2(1 + \mu_i^n) > 0.5 \text{ then} & \sigma_i = +1 \\ \text{else} & \sigma_i = -1. \end{cases}$$

A practical difficulty with mean field approximation is the choice of the temperature T at which the iterative process must occur. To avoid such a choice one of us [10] and other authors [20] have proposed to combine the mean field approximation process with an annealing process giving rise to mean field annealing. Hence, rather than fixing the temperature, the temperature is decreased during the convergence process: thus the μ_i's tend to $+1$ or -1 as the system converges.

V.A Estimating the MFA parameters and good initial conditions

The MFA equations contain two parameters to be determined: the initial temperature T_{init} and the decreasing factor of the temperature between two steps, $decT$. Moreover, the initial values of the μ_i^n's should be independent of the image. The aim is to avoid a trial-and-error process: the algorithm should be a black box from a user's perspective.

The initial temperature. The spin system onto which we have mapped the figure/ground discrimination problem typically has two phases; at high enough temperatures, according to the Boltzmann distribution, all the states reachable by the system are equally likely. As the temperature is lowered down, a phase transition occurs at $T = T_c$ and as $T \rightarrow 0$ the μ_i values represent a specific decision made as to the solution to the problem. At convergence, these values verify:

$$\frac{1}{N}\sum_{i=1}^{N} \mid \mu_i \mid = 1 \tag{18}$$

Untill now, we don't know how to calculate the critical temperature of an interacting spin system when the external field is not null (i.e., when there is at least one δ_i that is not null). Practically, we use random asynchronous dynamics (which is the dynamics the most similar to the behaviour of a spin system). We start with a high enough initial temperature determined by the μ_i distribution: Each μ_i has to be around 0 at the beginning and during several iterations; otherwise, the initial temperature is increased.

Annealing schedule. As already mentioned, two different annealing schedules are possible:

- Van den Bout & Miller [20] tried to estimate theoretically the critical temperature and then they perform two sets of iterations: one iteration process at this critical temperature until a near equilibrium state is reached and another iteration process at a temperature value that is close to 0. However, the critical temperature is quite difficult to estimate and the behaviour of the spin system predicted by their analytical approximation is not in accordance with the computational experiments.

- Instead of estimating the critical temperature, we prefer the following schedule. Initially the temperature has a high value and as soon as every spin has been updated at least once, the temperature is decreased to a smaller value. Then the temperature continues to decrease at each step of the convergence process. This does not guarantee that a near equilibrium state is reached at each temperature value but when the temperature is small enough then the system is frozen in a good stable state. Consequently, the convergence time is reduced since at low temperatures the convergence to a stationary solution is accelerated. One of us has successfully used this strategy to solve hard, np-complete graph combinatorial problems [9], [10].

The initial values of the μ_i's. When the vector $\vec{\delta}$ (i.e., the external field) is the null vector, one can start with a vector $\vec{\mu^0}$ close to the obvious unstable solution $[0, \cdots, 0]$. In practice, when the vector $\vec{\delta}$ is not null one starts with an initial configuration obtained by adding noise to $[0, \cdots, 0]$.

For instance the μ_i^0's are chosen randomly in the interval $[-10^{-5}, +10^{-5}]$. This initial state is plausible for a spin system in a heat bath at high temperature. In fact the spin values (+1 or –1) are equally likely at high temperatures and hence, $\langle \sigma_i \rangle = 0$ for all spin i (see equation (9)). During the iterative process, the μ_i's converge to values in between -1 or $+1$.

The mean field annealing algorithm is outlined in the Appendix.

VI Examples

We tested these algorithms over a wide variety of images. Figure 5 and Figure 6 show two such images. These images are preprocessed as follows. Edges are first extracted using the Canny/Deriche operator [4]. A small set of connected edges are grouped to form an edgel as follows: The tangent direction associated with such an edgel is computed by fitting a straight line in the least-square sense to the small set of connected edges. Then this small set of edges is replaced by an edgel, i.e., the fitted line. The position of the edgel is given by its midpoint, its direction is given by the direction of the line, and its contrast is given by the average contrast of the edges forming the edgel. Figure 7 and Figure 8 show the input data of our experiments.

Figure 5: Image #2.

Figure 6: Image #3.

Figure 2 shows the result of applying mean field annealing to the synthetic data (image #1). Figure 9 and Figure 10 show the result of applying mean field annealing to the real images.

VII Discussion

In this paper we attacked the problem of figure-ground discrimination with special emphasis on the problem of separating image data into curve and noise. We proposed a global approach using combinatorial optimization. We suggested a mathematical encoding of the problem which takes into account such image properties as cocircularity, smoothness, proximity, and contrast and this encoding fits the constraints of the statistical modelling of interacting spin systems. We described mean field annealing which is a deterministic approximation of simulated annealing and to which it compares favourably both in terms of robustness and efficiency. Indeed, equivalent sequential

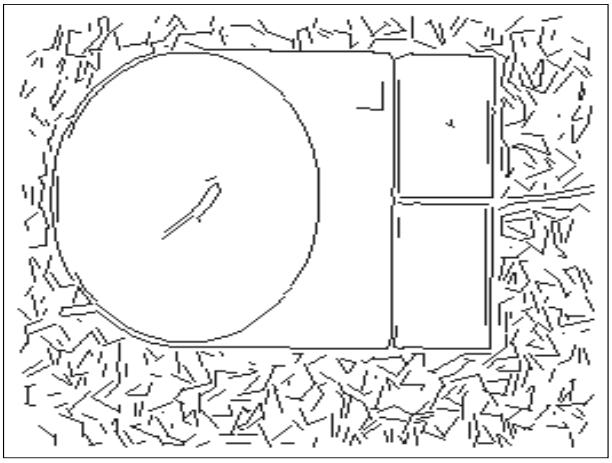

Figure 7: Image #2: Set of edgels.

Figure 8: Image #3: Set of edgels.

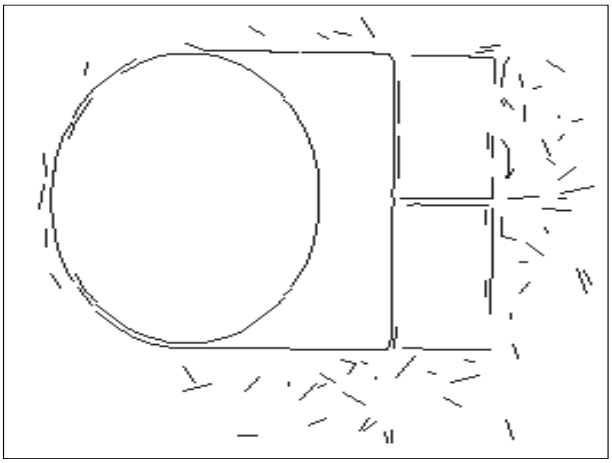

Figure 9: The result of applying MFA to image #2.

Figure 10: The result of applying MFA to image #3.

implementations of both algorithms (mean field annealing and simulated annealing) reveal that the former is of the order of 50 times faster than the latter [8]. Both simulated annealing and mean field annealing may be implemented on a fine-grained parallel machine – such an implementation is enforced by the analogy between the energy associated with an interacting spin system (such as ours) and the energy of a recursive neural network. With such an analogy, a neuron (or a processor) is associated with an edgel and each neuron communicates with all the other neurons. With such a parallel implementation and in synchronous mode (see section V) mean field annealing is N times faster than in asynchronous mode (where N is the number of edgels). Therefore the synchronous parallel implementation of mean field annealing is expected to be $50N$ times faster than the parallel implementation of simulated annealing.

In the future we plan to continue to try to improve our method in order to be able to eliminate all the noisy edgels, even those that are close to a shape and included in this shape by our current encoding. We also plan to try to solve other aspects of the image segmentation problem such as the feature grouping problem. We intend to extend the approach advocated in this paper to other

vision problems such as matching and reconstruction.

Appendix: The MFA algorithm

The following algorithm applies whenever the cost function associated with the combinatorial optimization problem to be solved is of the form of eq. (1).

1. Fix the convergence mode (synchronous or asynchronous).
   Fix the initial temperature: $T = T_{init} = 100$.
   Fix the temperature decreasing factor between two consecutive iterations: $decT = 0.9995$.
   For all i, μ_i^0 is set to a random value in the interval $[-10^{-5}, 10^{-5}]$.
   Initialize an iteration counter: $iter \leftarrow 0$.

2. $iter \leftarrow iter + 1$.
   Case of an asynchronous convergence mode:

   (a) $\forall i$, $\mu_i^{old} = \mu_i^{iter-1}$.

   (b) Scan at random the spins μ_i and iteratively update each spin μ_i see n in the scan (see equation (17)):

      i. Calculate
      $$\mu_i^{new} = \tanh\left(\frac{\sum_{j=1}^{N} J_{ij}\,\mu_j^{old} \;+\; \delta_i}{T}\right).$$

      ii. Update μ_i: $\mu_i^{old} \leftarrow \mu_i^{new}$.

   (c) $\forall i$, $\mu_i^{iter} = \mu_i^{old}$.

   Case of a synchronous convergence mode: for all spins μ_i,
   $$\mu_i^{iter} = \tanh\left(\frac{\sum_{j=1}^{N} J_{ij}\,\mu_j^{iter-1} \;+\; \delta_i}{T}\right).$$

3. If $iter \leq 5$,

   (a) decrease the temperature: $T \leftarrow decT \times T$,

   (b) goto step 2.

4. Test if the algorithm has converged into a configuration different from the initial one: if $\sum_{i=1}^{N} \sum_{l=iter-5}^{iter} \mid \mu_i^l \mid < 0,99 \times 6 \times N$, then the system has not converged:

   (a) decrease the temperature: $T \leftarrow decT \times T$,

   (b) goto step 2.

5. The system has converged: assign the final values of the σ_i's: $\forall i$, if $\mu_i^{iter} > 0$ then $\sigma_i = 1$ else $\sigma_i = -1$.

In practice we take $T_{init} = 10$ and $decT = 0.9995$. At convergence the values of the μ_i's are very close to either -1 or $+1$. In our experiments we have noticed that this is not the case with mean field approximation in which case the final values of the μ_i's are less disriminative.

References

[1] M. Brady and H. Asada. Smoothed local symmetries and their implementation. *International Journal of Robotics Research*, 3(3):36–61, 1984.

[2] J. Canny. A Computational Approach to Edge Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-8(6):679–698, November 1986.

[3] P. Carnevali, L. Coletti, and S. Patarnello. Image Processing by Simulated Annealing. *IBM Journal of Research and Development*, 29(6):569–579, November 1985.

[4] R. Deriche. Using Canny's criteria to derive a recursively implemented optimal edge detector. *International Journal of Computer Vision*, 1(2):167–187, 1987.

[5] D. Geiger and F. Girosi. Parallel and deterministic algorithms from mrf's: Surface reconstruction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 13(5):410–412, May 1991.

[6] D. Geiger and A. Yuille. A common framework for image segmentation. *International Journal of Computer Vision*, 6(3):227–243, 1991.

[7] D. Gutfinger and J. Sklansky. Robust classifiers by mixed adaptation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, PAMI-13(6):552–567, June 1991.

[8] L. Hérault and R. Horaud. Figure-ground discrimination: a combinatorial optimization approach. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 15(9):899–914, September 1993.

[9] L. Hérault and J.J. Niez. Neural Networks and Graph K-Partitioning. *Complex Systems*, 3(6):531–576, December 1989.

[10] L. Hérault and J.J. Niez. Neural Networks and Combinatorial Optimisation: A Study of NP-Complete Graph Problems. In E. Gelembe, editor, *Neural Networks: Advances and Applications*, pages 165–213. North Holland, 1991.

[11] M. Kass, A. Witkin, and D. Terzopoulos. Snakes: Active contour models. *International Journal of Computer Vision*, 1(4):321–331, January 1988.

[12] W. Kohler. *Gestalt Psychology*. Meridian, New-York, 1980.

[13] H. Orland. Mean field theory for optimization problems. *Journal de Physique – Lettres*, 46(17):L–763–L–770, September 1985.

[14] P. Parent and S.W. Zucker. Trace Inference, Curvature Consistency, and Curve Detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(8):823–839, August 1989.

[15] C. Peterson. A new method for mapping optimization problems onto neural networks. *International Journal of Neural Systems*, 1(1):3–22, 1989.

[16] C. Peterson. Track finding with neural networks. *Nuclear Instruments and Methods in Physics Research*, A279:537–545, 1989.

[17] T. J. Sejnowski and G. E. Hinton. Separating Figure from Ground with a Boltzmann Machine. In Michael Arbib and Allen Hanson, editors, *Vision, Brain, and Cooperative Computation*, pages 703–724. MIT Press, 1988.

[18] A. Shashua and S. Ullman. Structural Saliency: The Detection of Globally Salient Structures Using a Locally Connected Network. In *Proc. IEEE International Conference on Computer Vision*, pages 321–327, Tampa, Florida, USA, December 1988.

[19] H.E. Stanley. *Introduction to Phases Transitions and Critical Phenomena*. Oxford University Press, 1971.

[20] D.E. Van den Bout and T. K. Miller. Graph Partitioning Using Annealed Neural Networks. In *Int. Joint Conf. on Neural Networks*, pages 521–528, Washington D.C., June 1989.

[21] J. Zerubia and R. Chellappa. Mean field approximation using compound Gauss-Markov random field for edge detection and image restoration. In *Proceedings of ICASSP*, pages 2193–2196, Alburquerque, New Mexico, USA, April 1990.