MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY
and
CENTER FOR BIOLOGICAL INFORMATION PROCESSING
WHITAKER COLLEGE

A.I. Memo No.1114
C.B.I.P. Paper No. 35

June 1989

Parallel and deterministic algorithms for MRFs: surface reconstruction and integration

Davi Geiger and Federico Girosi

Abstract

In recent years many researchers have investigated the use of Markov random fields (MRFs) for computer vision. They can be applied for example in the output of the visual processes to reconstruct surfaces from sparse and noisy depth data, or to integrate early vision processes to label physical discontinuities. Drawbacks of MRFs models have been the computational complexity of the implementation and the difficulty in estimating the parameters of the model.

In this paper we derive deterministic approximations to MRFs models. One of the considered models is shown to give in a natural way the graduate non convexity (GNC) algorithm proposed by Blake and Zisserman. This model can be applied to smooth a field preserving its discontinuities. A new model is then proposed: it allows the gradient of the field to be enhanced at the discontinuities and smoothed elsewhere. All the theoretical results are obtained in the framework of the mean field theory, that is a well known statistical mechanics technique. A fast, parallel and iterative algorithm to solve the deterministic equations of the two models is presented, together with experiments on synthetic and real images. The algorithm is applied to the problem of surface reconstruction is in the case of sparse data. We also describe a fast algorithm that solves the problem of aligning the discontinuities of different visual models with intensity edges via integration.

This paper describes research done within the Center for Biological Information Processing, in the Department of Brain and Cognitive Sciences, and at the Artificial Intelligence Laboratory. This research is sponsored by a grant from the Office of Naval Research (ONR), Cognitive and Neural Sciences Division; by the Alfred P. Sloan Foundation; by the Artificial Intelligence Center of Hughes Aircraft Corporation (S1-801534-2); and by the NATO Scientific Affairs Division (0403/87). Support for the A. I. Laboratory's artificial intelligence research is provided by the Advanced Research Projects Agency of the Department of Defense under Army contract DACA76-85-C-0010, and in part by ONR contract N00014-85-K-0124.

1 Introduction

In order to give a viewer information about a three dimensional scene many algorithms have been developed on several early vision processes, such as edge detection, stereopsis, motion, texture, and color . This information refers to properties of the scene as shape, distance, color, shade or motion, and it is usually noisy and sparse: more processing is then necessary to extract the relevant information and fill in sparse data. In recent years many researchers [9][12][6] [4][5] have investigated the use of Markov Random Fields (MRFs) for early vision. MRFs models can generally be used for the reconstruction of a function starting from a set of noisy sparse data, such as intensity, stereo, or motion data. They have also been used to integrate early vision processes to label physical discontinuities. Two fields are usually required in the MRFs formulation of a problem: one represents the function that has to be reconstructed, and the other is associated to its discontinuities. The essence of the MRFs model is that the probability distribution of the configuration of the fields, given a set of data, is given as a Gibbs distribution. The model is then specified by an "energy function", that can be modeled to embody the a priori information about the system. In the standard approach an estimate of the field and its discontinuities is given by the configuration that maximizes the probability distribution, or equivalently that minimizes the energy function. Since the discontinuity field is a discrete valued field (it assumes only the values 0 or 1) this becomes a combinatorial optimization problem, that can be solved by methods of the Monte Carlo type (simulated annealing[10], for example).
The MRFs formulation is appealing because describes a system by local interactions and allows to capture many features of the system of interest by simply adding appropriate terms in the energy function. However, it has two main drawbacks: the amount of computer time needed for the implementation and the difficulty in estimating the parameters that control the relative weight of the various terms of the energy function.

In this paper we propose a deterministic approach to MRFs models. It consists in explicitly writing down a set of equations from which we can compute estimates of the mean values of the field f and the line process. We study two MRFs models, the second being an extension of the first, and we show that in each case the deterministic equations lead to a fast, parallel and iterative algorithm that can be used to obtain estimates of the fields

everywhere.

A natural framework for this approach is the equilibrium statistical mechanics, since it deals with systems with many degrees of freedom described by Gibbs probability distributions. In principle, statistical mechanics allows us to derive the mean statistical values of the field (mean field equations) as *explicit* functions of the data and the parameters of the model: an algorithm implementing this would consist just of one step. Since the analytical computations required to obtain these explicit expressions are usually too hard, some approximations have to be made. We use the *mean field approximation*, a well known statistical mechanics tool, to obtain an approximated solution, that is given in implicit form by a set of non linear equations. We call these equation *deterministic* to underline the deterministic character of the whole procedure.

We concentrate our attention on the partition function Z, that is the sum of the probability distribution over all the possible field configurations, since it is known to contain all the information about the system. The idea underlying our approach is first to eliminate the line process degrees of freedom from Z: we will show that in so doing the effect of their interaction with the field can be simulated by a temperature dependent "effective potential" that depends only upon f. Its use is fundamental in the derivation of the deterministic equations, and gives useful insights on the role and the significance of the parameters.

An advantage of such an approach is that the solution of the deterministic equations is faster than the Monte Carlo techniques, fully parallelizable and feasible of implementation on analog networks. The possibility of writing a set of equations is also useful for a better understanding of the nature of the solution and of the parameters of the model.

We discuss two different MRFs models. The energy function of the first model has been already studied by several authors[2][11][13][12]. It is interesting to notice that the GNC algorithm, proposed by Blake and Zisserman [2] in *ad hoc* manner, arises naturally in the framework of statistical mechanics. This estabilish a connection between MRFs and deterministic algorithm already used in vision.

In the second model we define an energy function with an extra term, that establishes an interaction between the line process at neighborhood sites. This interaction can stimulate the creation of a line at a particular site if a line at a neighborhood site has been created. This term allows the data

to be smoothed and, at the same time, the contrast at the discontinuities to be enhanced. Typical edge detection features like hysteresis, threshold and suprathreshold, which do not arise from the previous energy function, will arise naturally from the model. We point out that this model is a generalization of the previous one, that can be recovered by simply setting an appropriate parameter to zero.

These two methods can be applied to dense data and, with small modification, to sparse data as well. The problem of surface reconstruction (and image restoration) from sparse data is addressed and an algorithm to perform these tasks is obtained and implemented. We also outline an algorithm that solves the problem of aligning the discontinuities of different visual models with intensity edges via integration.

The paper is organized in the following way: Chapter 2 presents an overview of MRFs in vision. Chapter 3 discusses the deterministic approximation of MRFs for the three energy functions mentioned above. Chapter 4 shows how to estimate the parameters of the models. In Chapter 5 some results are described. Chapter 6 discusses applications for sparse data and integration of visual modules with intensity. Chapter 7 concludes the paper.

2 MRFs for smoothing fields and detecting discontinuities

Here we briefly summarize how MRFs have been used in computer vision. A more extensive discussion is given in Geman and Geman [9], Marroquin[12], Chou[4], Gamble and Poggio[6], Gamble, Geiger, Poggio, Weinshall [5].

Consider the problem of approximating a surface given sparse and noisy depth data, on a regular 2D lattice of sites. We think the surface as a field (surface-field) defined in the regular lattice, such that the value of this field at each site of the lattice is given by the surface height at this site. The Markov property asserts that the probability of a certain value of the field at any given site in the lattice depends only upon neighboring sites. According to the Clifford-Hammersley theorem, the prior probability of a state of the field f has the Gibbs form:

$$P(f) = \frac{1}{Z_f} e^{-\beta U(f)} \tag{2.1}$$

where f is the field, e.g. the surface-field, Z_f is the partition function, $U(f) = \sum_i E_i(f)$ is an energy function that can be computed as the sum of local contributions from each lattice site i, and β is a parameter that is called the inverse of the natural temperature of the field. If a sparse observation g for any given surface-field f is given and a model of the noise is available then one knows the conditional probability $P(g|f)$. Bayes theorem then allows to write the posterior distribution:

$$P(f|g) = \frac{P(g|f)P(f)}{P(g)} \equiv \frac{1}{Z} e^{-\beta E(f|g)} \ . \tag{2.2}$$

As a simple example, when the surfaces (surface-fields) are expected to be smooth and the noise is Gaussian, the energy is given by

$$E(f|g) = \sum_i [\gamma_i (f_i - g_i)^2 + \alpha \sum_{j \epsilon N_i} (f_i - f_j)^2], \tag{2.3}$$

where $\gamma_i = 1$ or 0 depending on whether data are available or not and N_i is a set of sites in an arbitrary neighborhood of the site i. The maximum of the posterior distribution or other related estimates of the "true" data-field value can not be computed analytically, but sample distributions of the field with the probability distribution of (2.2) can be obtained using Monte Carlo techniques such as the Metropolis algorithm [14]. These algorithms sample the space of possible values of the surface-field according to the probability distribution $P(f|g)$.

One of the main attractions of MRFs models is that they can deal directly with discontinuities. Geman and Geman [9] introduced the idea of another field, the line process, located on the dual lattice, and representing explicitly the presence or absence of discontinuities that break the smoothness assumption (2.2). The dual lattice is another lattice coupled with the data field lattice such that for each site of the data field lattice there are two sites of the dual lattice, one site corresponding to the vertical line and the other one to the horizontal line. The associated prior energy then becomes:

$$E_i(f, l) = \sum_{j \epsilon N_i} (f_i - f_j)^2 (1 - l_{ij}) + \sum_C V_C(l_{ij}) \tag{2.5}$$

where l_{ij} is the element of the binary field l located between site i, j. The term $V_C(l_{ij})$, where C is a clique defined by the neighborhood system of the line

process (binary field), reflects the fact that certain configurations of the line process are more likely to occur than others. Depth discontinuities are usually continuous and non-intersecting, and rarely consist of isolated points. These properties of physical discontinuities can be enforced locally by defining an appropriate set of energy values $V_C(l_{ij})$ for different configurations of the line process ([9], [13],[6].) In our models the cliques will be simplified to the nearest neighbors.

Two basic problems have arisen in using MRFs to solve vision problems. The first one is the amount of computer time used in the Metropolis algorithm or in simulated annealing [10]. The second problem is to how estimate the parameters of the energy function since they have been estimated in an ad hoc manner.

We propose to approximate the solution of the problem formulated in the MRFs frame with its "average solution." The mean-field theory (which is explained in the next chapter) allows us to find deterministic equations analytically for MRFs whose solution approximates the solution of the statistical problem.

3 A deterministic approximation of MRFs

3.1 The Line Process for two dimensions

The line process was described in chapter 2 for the one dimensional case. Now we generalize it to two dimensions. In this case we define a horizontal line process h_{ij} and a vertical line process v_{ij}. The line process h_{ij} connects the site (i,j) to the site $(i,j-1)$, while v_{ij} connects the site (i,j) to the site $(i-1,j)$. This is illustrated in figure 1. It is important to notice that the horizontal line process is determined by the gradient of the field in the vertical direction, and the vertical line process is determined by the gradient of the field in the horizontal direction. In this way we can reduce the two dimensional problem to two one dimensional problems, provided that the horizontal and vertical line processes do not interact. At the end we can add both line processes to get the edge map. This approach will be exploited in the next chapters. We point out that the fields f_{ij}, h_{ij} and v_{ij} are defined in the same lattice instead of having f_{ij} in one lattice and h_{ij}, v_{ij} in a dual lattice.

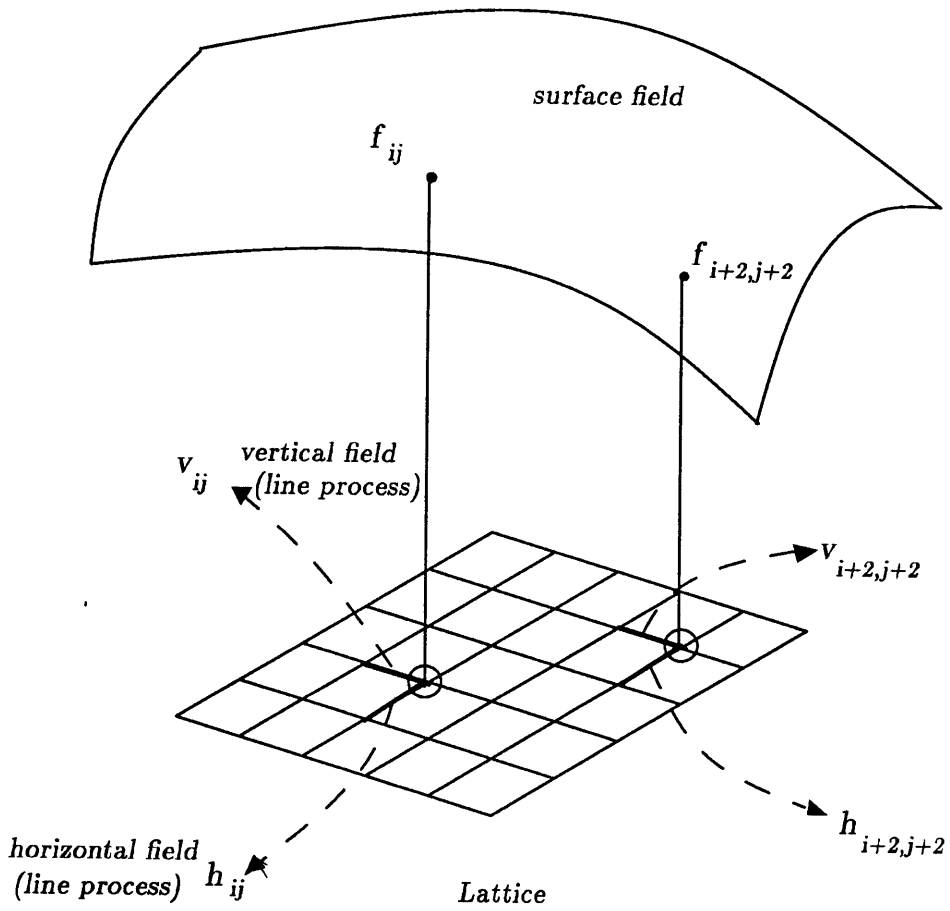

Figure 1: *The surface field f, the horizontal line process h and the vertical line process v are represented in two sites, (i,j) and (i+2, j+2), of the lattice. There are three fields defined in each site of the lattice as opposed to having one field in a lattice and the other two fields in a dual lattice.*

3.2 The Weak Membrane energy

Or how to smooth the field but not at the discontinuities

The Weak Membrane energy is given by

$$E_1(f,h,v) = E_{fg}(f) + E_{fl}(f,h,v) + E_l(h,v) \tag{3.1}$$

where

$$E_{fg}(f) = \sum_{i,j} (f_{i,j} - g_{i,j})^2 \tag{3.1a}$$

$$E_{fl}(f,h,v) = \alpha \sum_{i,j} [(f_{i,j} - f_{i,j-1})^2(1 - h_{i,j}) + (f_{i,j} - f_{i-1,j})^2(1 - v_{i,j})] \tag{3.1b}$$

$$E_l(h,v) = \gamma \sum_{i,j} (h_{i,j} + v_{i,j}) \tag{3.1c}$$

and α and γ are positive valued parameters.

The first term, as in the previous case, enforces closeness to the data and the second one contains the interaction between the field and the line processes: if the horizontal or vertical gradient is very high at site (i,j) the corresponding line process will be very likely to be active ($h_{i,j} = 1$ or $v_{i,j} = 1$), to make energy decrease and signal a discontinuity. The third term takes into account the price we pay each time we create a discontinuity and is necessary to prevent the creation of discontinuities everywhere.

3.3 Mean field theory and Weak Membrane

We assume that there is uncertainty in the model and then the statistical framework can be used to incorporate the uncertainty as described by (2.1). Using mean field theory it can be shown (see Appendix A) that

$$\overline{f_{ij}} = g_{ij} - \frac{1}{2}\frac{\partial F}{\partial g_{ij}} \tag{3.2}$$

where $F = -\frac{1}{\beta} ln Z$ is the free energy. Z is the partition function defined as

$$Z = \sum_{\{f\}} e^{-\beta E(f)}$$

where $\sum_{\{f\}}$ means the sum over all the possible configurations $\{f\}$ of the system.
The computation of the partition function is equivalent, in this case, to the evaluation of a multi-dimensional integral which can not be explicitly solved, due to the interaction between all the variables. Even if an exact solution can not be found, we can still obtain a good solution making use of the so called *mean field approximation.* The mean-field approximation is a general tool used in statistical mechanics that consists in substituting the interaction among the fields at different locations by the interaction of the field at each site with the mean field value at different locations (see appendix A)
After this step the partition function factors into the product of single-site partition functions, that can usually be computed

$$Z^{mf} = \prod_i [\sum_{\{f\}} e^{-\beta E_i^{mf}}]$$

.
Now the partition function, and then the free energy are functions of the mean values of the fields, which are still unknown. These values, however, are usually related to the derivatives of the free energy with respect to some external field, as in (3.2). Substituting the free energy with its mean-field approximation, we obtain a set of implicit and self-consistent equations for the fields, usually called *mean-field equations.* In this sense the mean-field equations give a deterministic solution to the MRFs problem, since they relate explicitly the mean values of the field to the data. They could be implemented by a deterministic network.
In the case of (3.1) the partition function becomes

$$Z = \sum_{\{f\}} e^{-\beta \sum_{i,j}[(f_{i,j}-g_{i,j})^2+\alpha({\Delta^h_{i,j}}^2+{\Delta^v_{i,j}}^2)]} \sum_{\{h,v\}} e^{-\beta \sum_{i,j}[h_{i,j}G^h_{i,j}+v_{i,j}G^v_{i,j}]} \qquad (3.3)$$

where $G^h_{i,j} = \gamma - \alpha {\Delta^h_{i,j}}^2$, $G^v_{i,j} = \gamma - \alpha {\Delta^v_{i,j}}^2$, $\Delta^h_{i,j} = f_{i,j} - f_{i-1,j}$ and $\Delta^v_{i,j} = f_{i,j} - f_{i,j-1}$.

3.3.1 Averaging Out the Line Process

The contribution of the line process to the partition function can be exactly computed. Indeed the line process term in (3.3) is the partition function of two spin systems (h and v) in an external field (G^h and G^v) with no interaction between neighboring sites. Then each spin contributes to the partition function independently from the others and its contribution is $(1 + e^{-\beta G^h_{i,j}})$ for the horizontal field and a similar factor for the vertical one. The partition function can then be rewritten as

$$Z = \sum_{\{f\}} e^{-\beta \sum_{i,j}[(f_{i,j}-g_{i,j})^2+\alpha({\Delta^h_{i,j}}^2+{\Delta^v_{i,j}}^2)]} \prod_{ij}(1 + e^{-\beta G^h_{i,j}})(1 + e^{-\beta G^v_{i,j}}) \quad (3.4) .$$

The capability of computing the line process contribution to Z allows us to obtain a solution for the mean values of the fields h and v once a solution for the field is provided. As an intermediate step, after some algebra derived in appendix B we obtain

$$\bar{h}_{i,j} =< \frac{1}{1 + e^{\beta G^h_{i,j}}} > \quad and \quad \bar{v}_{i,j} =< \frac{1}{1 + e^{\beta G^v_{i,j}}} > \qquad (3.5)$$

where $< ... >$ means the statistical mean value. In order to get a deterministic solution for the line process we need to do some approximations. We will assume then that we can neglect the statistical fluctuations of the field and so we replace the value of the field in (3.5) with its mean value. This is in essence the mean-field approximation.

$$\bar{h}_{i,j} = \frac{1}{1 + e^{\beta(\gamma-\alpha(\bar{f}_{i,j}-\bar{f}_{i-1,j})^2)}} \quad and \quad \bar{v}_{i,j} = \frac{1}{1 + e^{\beta(\gamma-\alpha(\bar{f}_{i,j}-\bar{f}_{i,j-1})^2)}} \qquad (3.6).$$

In the zero temperature limit ($\beta \to \infty$) (3.6) becomes the Heaviside function (1 or 0) and the interpretation is simple: when the horizontal or vertical gradient ($\bar{f}_{i,j} - \bar{f}_{i,j-1}$ or $\bar{f}_{i,j} - \bar{f}_{i-1,j}$) are larger than a threshold ($\sqrt{\frac{\gamma}{\alpha}}$) a vertical or horizontal discontinuity is created, since the price to smooth the function at that site is too high. This leads to a clear interpretation of the parameter γ, as it will be discussed in section 4.2.

We notice that (3.6) can account for diagonal lines. We show with the example of a rectangle in a square lattice (see figure 2a). The threshold ($\sqrt{\frac{\gamma}{\alpha}}$)

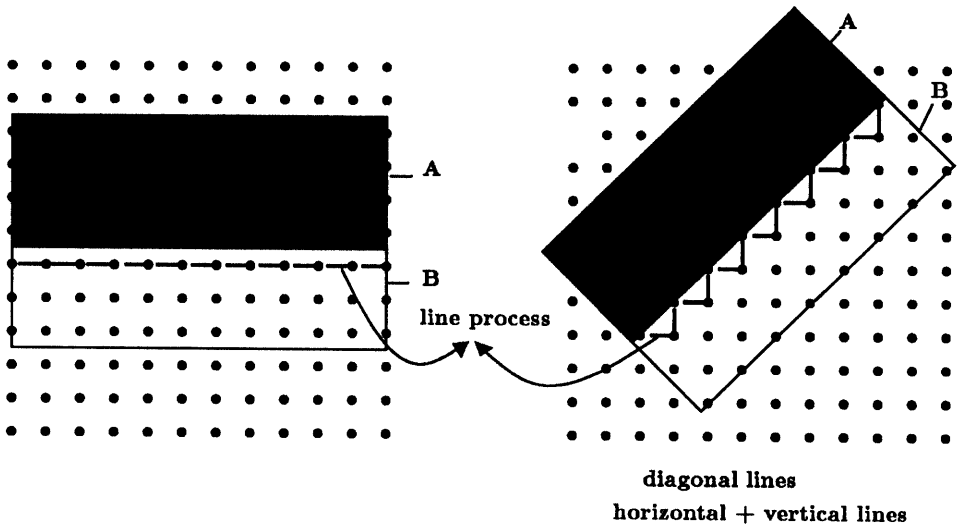

Figure 2: *a) At the discontinuity,* $\Delta_{i,j}^{h\,2} = (B-A)^2$ *and* $\Delta_{i,j}^{v\,2} = 0$. *A chain of horizontal line processes are created for* $(\frac{\gamma}{\alpha} < (B-A)^2)$. *b) At the discontinuity,* $\Delta_{i,j}^{h\,2} = (B-A)^2$ *and* $\Delta_{i,j}^{v\,2} = (B-A)^2$. *A chain of horizontal and vertical line processes are created for* $(\frac{\gamma}{\alpha} < (B-A)^2)$.

to create a discontinuity line is invariant under rotations of the rectangle with respect to the lattice (see figure 2b). The diagonal line is perceived from a chain of horizontal and vertical lines (like a stair case). The important result obtained from (3.6) is that although the total line created in figure 2b is $\sqrt{2}$ times bigger than in figure 2a the threshold to create a line has been kept the same.

3.3.2 The Effective Potential

We discuss how the interaction of the field f with itself has changed after the line process has been eliminated from the partition function. From (3.4) we notice that the partition function can be rewritten as

$$Z = \sum_{\{f\}} e^{-\beta(E_{fg}(f)+E_{eff}(f))}$$

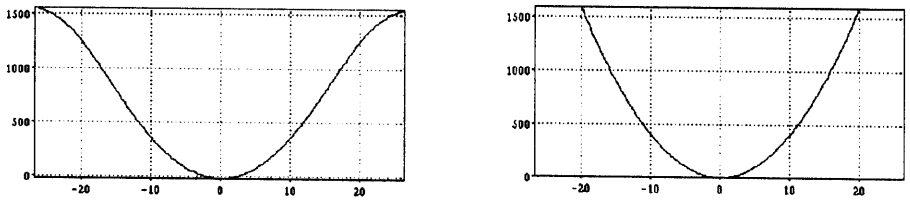

Figure 3: *The effective potential is shown as a function of* Δ^h_{ij}. *(For illustration we set* $\Delta^v_{ij} = 0$.*) a) For* $\beta = 0.002$. *b) Zero temperature limit* ($\beta \to \infty$).

where

$$E_{eff}(f) = \sum_{i,j} \alpha({\Delta^h_{i,j}}^2 + {\Delta^v_{i,j}}^2) - \frac{1}{\beta} ln[(1 + e^{-\beta G^h_{i,j}})(1 + e^{-\beta G^v_{i,j}})]$$

and $E_{fg}(f)$ is given by (3.1a).
This is the partition function of a system composed of one continuous valued field, whose energy is $E_{fg} + E_{eff}$. We interpret this result as the **effect** of the interaction of the line processes with the field f. This effect can be simulated by modifying appropriately the interaction of the field with itself, substituting the smoothing term in the energy function with a new temperature dependent potential.
In figure 3 the effective potential is depicted for different temperatures. It simulates the effect of the line processes on the field f. Notice that the energy function is still the sum of local interactions between first neighbors. For the zero temperature limit one can see in figure 3 that the smoothing term is active only when the gradient is smaller than a threshold, proportional to the ratio between γ and α. When the temperature is different from zero the border (threshold) of the smoothing region is no longer well defined due to thermal noise. It corresponds to an adaptive threshold that depends on the temperature.

3.3.3 A Deterministic Solution for f

The introduction of the effective potential allows us to obtain a set of deterministic equations for the field f. The mean field theory, as explained in section 3.1, can be applied to the partition function of equation (3.4) to obtain a set of mean-field equations depending on the temperature. In terms of statistical mechanics the data term plus the effective potential represent the free energy of the system. The mean field solutions are obtained by minimizing the free energy. The set of deterministic equations can be written as

$$\frac{\partial}{\partial f_{i,j}}(E_{fg} + E_{eff}(f)) = 0$$

and after some computation

$$\begin{aligned} \bar{f}_{i,j} = \; & g_{i,j} - \alpha(\bar{f}_{i,j} - \bar{f}_{i,j-1})(1 - \bar{v}_{i,j}) + \alpha(\bar{f}_{i,j+1} - \bar{f}_{i,j})(1 - \bar{v}_{i,j+1}) \\ & -\alpha(\bar{f}_{i,j} - \bar{f}_{i-1,j})(1 - \bar{h}_{i,j}) + \alpha(\bar{f}_{i+1,j} - \bar{f}_{i,j})(1 - \bar{h}_{i+1,j}) \end{aligned} \tag{3.7}$$

where $\bar{h}_{i,j}$ and $\bar{v}_{i,j}$ are given by formula 3.6.
Equation (3.7) gives the field at site i, j as the sum of data at the same site, plus an average of the field at its neighbor sites. This average takes in account the difference between the neighbors. The larger is the difference, the smaller is the contribution to the average. This is captured by the term $(1-l_{i,j})$, where $l_{i,j}$ is the line process. At the zero temperature limit ($\beta \to \infty$) the line process becomes 1 or 0 and then only terms smaller than a threshold must be taken in account for the average. This interpretation helps us in understanding the role of the α and γ parameters, as it will be discussed in chapter 4. Notice that the form of (3.7) is suitable for the application of a fast, parallel and iterative scheme of solution.

3.3.4 The Effective Potential and the Graduate Non Convexity Algorithm

We have to point out that this energy function has been studied by Blake and Zisserman [2], in the context of edge detection and surface interpolation. They do not derive the results from the MRFs formulation but they simply minimize the Weak Membrane energy function. From a statistical mechanics point of view the mean-field solution does not minimize the energy function,

but this becomes true in the zero temperature limit, so their approach must be recovered from the MRFs formulation in this limit. This is indeed the case, and it is easy to show that the effective potential becomes the Blake and Zisserman potential when β goes to infinity. In order to obtain the minimum of the energy function E_1 Blake and Zisserman introduce the GNC (Graduate Non Convexity) algorithm, which is different from our deterministic scheme, but can be embedded in the MRFs framework in a natural way. Let us review briefly the GNC algorithm. The main problem with the Weak Membrane Energy is that is not a convex function and a gradient descent method can not be applied to obtain the minimum because one could be trapped in a local minimum. In order to solve this problem Blake and Zisserman introduce a family of energy functions $E^{(p)}$, depending continuously on a parameter p, $p\epsilon[0,1]$, such that $E^{(1)}$ is convex, $E^{(0)} \equiv E_1$ and $E^{(p)}$ are non convex for $p\epsilon[0,1)$. Gradient descent is successively applied to the energy function $E^{(p)}$ for a prescribed decreasing sequence of values of p starting from $p = 1$, and this procedure is proved to converge for a class of given data. The construction of the family of energy functions $E^{(p)}$is *ad hoc* and uses piecewise polynomials. In our framework, a family of energy functions with such properties is naturally given by ${E_{eff}}^{(T)}$ where T is the temperature of the system. The GNC algorithm can then be interpreted as the tracking of the minimum of the energy function as the temperature is lowered to zero (like a deterministic annealing). In this way the approach of Blake and Zisserman can be viewed as a deterministic solution of the MRFs problem, even if it does not fully exploit the possibility of obtaining deterministic equations for the surface and discontinuity fields. The results obtained by applying this method to edge detection, for example, are good, and a pattern of meaningful discontinuities can usually be recovered [2]. However, sometimes the full set of discontinuities is not obtained: when the gradient of the image brightness is under the threshold, a discontinuity may not be detected, even though it would be necessary, for example, to close a contour.
If interaction between lines (self interaction of the line-field) is introduced in the energy function this problem can be overcome, as we discuss next.

Figure 4 summarizes the procedure used before to obtain deterministic algorithms from the statistical methods.

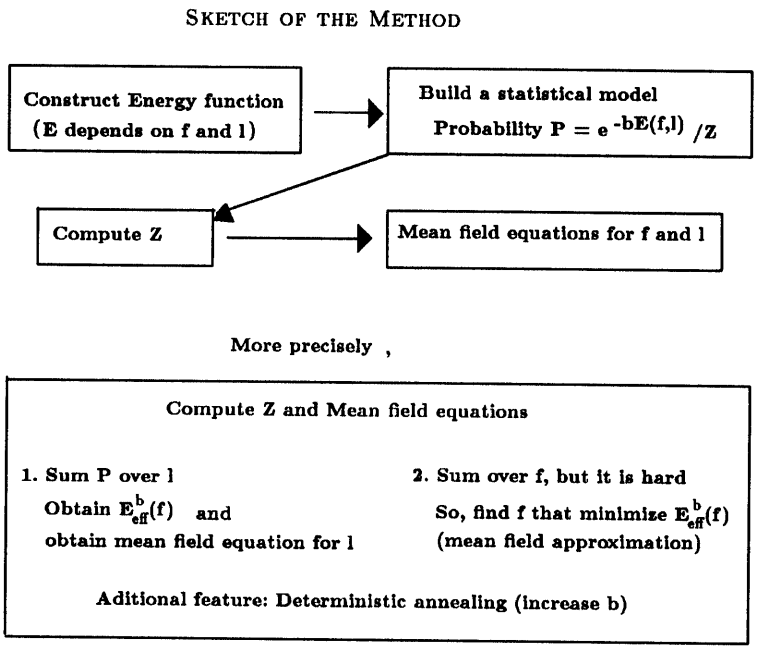

Figure 4: *Overview of the method*

3.4 Improving the weak membrane model

Or how to smooth the field while enhancing the differences at the discontinuities

So far we have not exploited an important physical constraint of images, namely the smoothness of the discontinuity field. Isolated discontinuities are very unlikely to occur and, on the contrary, the presence of a discontinuity at a site makes more likely the presence of a discontinuity at a neighboring site. This smoothness constraint on the discontinuity field can be incorporated in the model by simply adding a new term to the energy function. Thus, the total energy becomes

$$E_2 = E_1 + E_{ll}$$

where E_1 is given by (3.1) and we define the new term

$$E_{ll} = -\epsilon\gamma \sum_{i,j} [h_{i,j}h_{i,j-1} + v_{i,j}v_{i-1,j}]$$

and ϵ is a new parameter whose exact meaning and estimation will be explained in the next section. To make more evident the meaning of the new term we notice that

$$E_l + E_{ll} = \sum_{ij} \gamma(1-2\epsilon)[h_{i,j}^2 + v_{i,j}^2] + \epsilon\gamma[(h_{i,j} - h_{i,j-1})^2 + (v_{i,j} - v_{i-1,j})^2] \quad (3.8)$$

where E_l is given by (3.1c). From (3.8) it is clear that ϵ is related to the degree of smoothness of the discontinuity field and so must be a positive number. In order to keep positive the price we pay for creating a discontinuity and to prevent the line processes from being active everywhere, ϵ has to be less than 1.

We notice here that this model is a simplified version of other possible potentials. In particular, the neighbor size considered here is at most of two pixels where Gamble & Poggio (1987) for example, have discussed more sophisticated cliques composed of larger neighbors.

3.4.1 Averaging Out the Line Process

Again, as in the Weak Membrane energy case we calculate the contribution of the line process in the partition function. The partition function can be written as

$$Z = \sum_{\{f\}} e^{-\beta \sum_{i,j}[(f_{i,j}-g_{i,j})^2+\alpha({\Delta^h_{i,j}}^2+{\Delta^v_{i,j}}^2)]} \times \\ \times \sum_{\{h,v\}} e^{-\beta \sum_{i,j}[h_{i,j}(G^h_{i,j}-\epsilon\gamma\frac{h_{i,j-1}+h_{i,j+1}}{2})+v_{i,j}(G^v_{i,j}-\epsilon\gamma\frac{v_{i-1,j}+v_{i+1,j}}{2})]} \quad (3.9)$$

where the Gs have been defined in section 3.3. We rewrite the line process term by substituting $h_{i,j}(G^h_{i,j} - \epsilon\gamma h_{i,j-1})$ by $h_{i,j}(G^h_{i,j} - \epsilon\gamma\frac{h_{i,j-1}+h_{i,j+1}}{2})$. Notice that this change is just a rearrangement of the terms on the sum over i, j and therefore does not change the value of Z. We also notice that the line process term of (3.9) is the partition function of many one-dimensional Ising spin models in an external field. For a constant external field an exact solution can be obtained by the method of the transfer matrix (see Parisi [15]). This method can be still applied to our case (not constant external field), giving a compact expression for the partition function. However, the expression we obtain is highly non local for the field f and not useful for fast computations. Because of it we apply the mean field approximation and replace $h_{i,j-1}$ by $\bar{h}_{i,j-1}$, $h_{i,j+1}$ by $\bar{h}_{i,j+1}$, $v_{i-1,j}$ by $\bar{v}_{i-1,j}$, and $v_{i+1,j}$ by $\bar{v}_{i+1,j}$ in (3.9). Notice that at this step the mean field approximations refer to the line process and not to the field f (that is kept "frozen"). The sum over the line process configurations (3.9) can then be computed, giving

$$Z^{mf-hv} = \sum_{\{f\}} e^{-\beta \sum_{i,j}[(f_{i,j}-g_{i,j})^2+\alpha({\Delta^h_{i,j}}^2+{\Delta^v_{i,j}}^2)]} \times \\ \times \prod_{ij}(1 + e^{-\beta(G^h_{i,j}-\epsilon\gamma\frac{\bar{h}_{i,j-1}+\bar{h}_{i,j+1}}{2})})(1 + e^{-\beta(G^v_{i,j}-\epsilon\gamma\frac{\bar{v}_{i-1,j}+\bar{v}_{i+1,j}}{2})})$$

The partition function is now similar to the previous one (3.4) and the zero-temperature mean-field equations for the line process field can then be derived in the same approximation:

$$\bar{h}_{i,j} = \sigma_\beta(\alpha(\bar{f}_{i,j} - \bar{f}_{i,j-1})^2 - \gamma + \epsilon\gamma\frac{\bar{h}_{i,j-1} + \bar{h}_{i,j+1}}{2}) \quad \text{and}$$

$$\bar{v}_{i,j} = \sigma_\beta(\alpha(\bar{f}_{i,j} - \bar{f}_{i-1,j})^2 - \gamma + \epsilon\gamma\frac{\bar{v}_{i-1,j} + \bar{v}_{i+1,j}}{2}) \quad (3.10)$$

where $\sigma_\beta(x) = \frac{1}{1+e^{\beta x}}$ is the sigmoid function. Notice that now equations (3.10) form a set of non linear equations and the solution for the line process is not as simple as before.

3.4.2 Hysteresis, Suprathreshold and Threshold

The solution obtained in (3.10) for the line process can deal with the problem of "streaking". Streaking is the breaking up of an edge contour caused by fluctuations above and below a threshold along a contour [3]. This is a common problem with thresholded detectors, in particular with the scheme implemented by the Weak Membrane energy. For the third energy the thresholding is done with hysteresis. This is the way that Canny's detector works [3]. The hysteresis phenomenon is evident in (3.10). With the creation of a horizontal line at a neighbor site, say $\{i, j-1\}$ ($h_{i,j-1} = 1$), the energy necessary for creating a line at a site $\{i, j\}$ decreases by $\frac{\epsilon\gamma}{2}$. On the other hand, if two lines are created at $\{i, j-1\}$ and $\{i, j+1\}$ then the energy decreases by $\epsilon\gamma$. A low threshold (threshold) and a high threshold (suprathreshold) arise naturally. The suprathreshold for creating a line is given by $\frac{\gamma}{\alpha} \leq (\overline{\Delta^h_{i,j}})^2$, in this case a line is always created. In the same way the lowest threshold is given by $\frac{\gamma}{\alpha}(1-\epsilon) \leq (\overline{\Delta^h_{i,j}})^2 \leq \frac{\gamma}{\alpha}$, in this case a horizontal line will be created at site $\{i, j\}$ if a horizontal line has been created in the sites $\{i, j-1\}$ and $\{i, j+1\}$.
At higher temperatures the threshold and suprathreshold are not so well - defined, and what we have are adaptive thresholds that depend on the values of the field. These points will become clearer when we discuss the effective potential.
We point out that some hysteresis may occur on the Weak Membrane energy. However it will happen just on special cases and without much of control. A consistent framework to enhance and smooth images requires at least two thresholds. The Weak Membrane energy is a one-threshold model and therefore can not fully exhibit hysteresis.

3.4.3 The Effective Potential

By analogy to section 3.3, the effective potential is given by

$$E_{eff}(f) = \sum_{i,j}\Big\{\alpha({\Delta^h_{i,j}}^2 + {\Delta^v_{i,j}}^2) - \frac{1}{\beta} ln[(1 + e^{-\beta(G^h_{i,j} - \epsilon\gamma\frac{\bar{h}_{i,j-1}+\bar{h}_{i,j+1}}{2})})(1 + e^{-\beta(G^v_{i,j} - \epsilon\gamma\frac{v_{i-1,j}+v_{i+1,j}}{2})})]\Big\} \quad (3.11)$$

An exact computation by means of the transfer matrix method shows that this potential is highly non local and its practical use is limited. What we will do next is to use approximation techniques to obtain a potential that is local , up to the first neighbors.
Without losing any generality we now dismiss the contribution of the vertical process since it is independent of the horizontal process contribution. In order to obtain an expression for E_{eff} in terms of the gradient field Δ_h we investigate (3.10). Since the discontinuity field is smooth we first approximate the term $\frac{\bar{h}_{i,j-1}+\bar{h}_{i,j+1}}{2}$ by $\bar{h}_{ij}$ and (3.10) at the zero temperature becomes

$$\bar{h}_{i,j} = \theta(\alpha(\bar{f}_{i,j} - \bar{f}_{i,j-1})^2 - \gamma + \epsilon\gamma\bar{h}_{i,j}) \quad (3.12)$$

The solution of (3.12) depends on the value of $\overline{\Delta^h_{i,j}}^2 = (\bar{f}_{i,j} - \bar{f}_{i,j-1})^2$ as follows

$$\bar{h}_{i,j} = \begin{cases} 0 & \text{if} \quad (\overline{\Delta^h_{ij}})^2 \leq \frac{\gamma}{\alpha}(1-\epsilon) \\ 0 \text{ or } 1 & \text{if} \quad \frac{\gamma}{\alpha}(1-\epsilon) \leq (\overline{\Delta^h_{ij}})^2 \leq \frac{\gamma}{\alpha} \\ 1 & \text{if} \quad (\overline{\Delta^h_{ij}})^2 \geq \frac{\gamma}{\alpha} \end{cases}$$

These results are illustrated in figure 5. We notice that in the region $\frac{\gamma}{\alpha}(1-\epsilon) < (\Delta^h_{ij})^2 < \frac{\gamma}{\alpha}$ both solutions are possible , thus reflecting a weakness in the assumption we made. A possible way to obtain a unique solution is to consider an average of the two solutions such that the higher the gradient of f, the higher is the value of h (more likely for creating a line). We interpret the values of h between 1 and 0 as the likelihood of creating a line. We approximate the solution of the mean field equation for $h_{i,j}$ at any temperature (β) by the analytical expression

$$\bar{h}_{i,j} = \frac{\alpha}{\gamma}(\Delta^h_{ij})^2 - \frac{1}{\beta\gamma} ln(1 + e^{-\beta(\gamma - \alpha(\Delta^h_{ij})^2)}) + \frac{1}{\beta\gamma} ln(1 + e^{-\beta\gamma}) \quad (3.13)$$

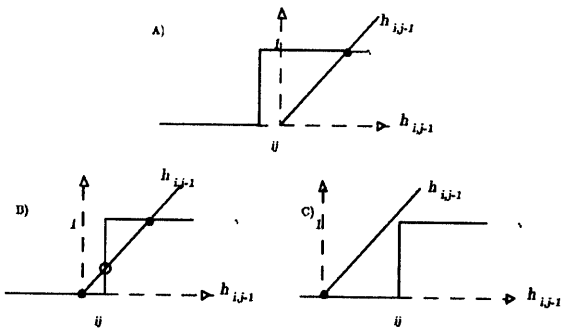

Figure 5: *The three figures illustrate the solution of (3.12) according to different values of the gradient field f* (Δ_{ij}^h) *a)The line process is on* $(h_{ij} = 1)$ *b) Two solution (three solution for temperature different then zero) c) Shows the solution of no line-process* $(h_{ij} = 0)$

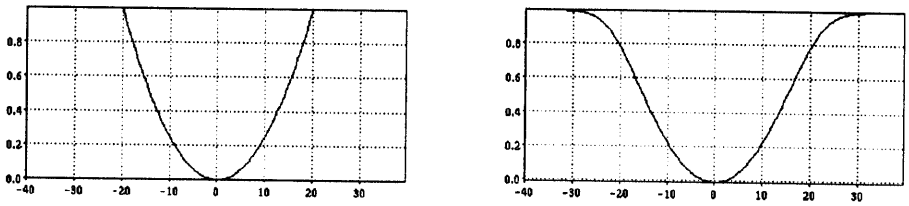

Figure 6: $\bar{h}_{i,j}$ *as a function of the gradient field* (Δ_{ij}^h). *Two solutions of (3.13). a)* $\beta = 1$ *b)* $\beta = 0.002$

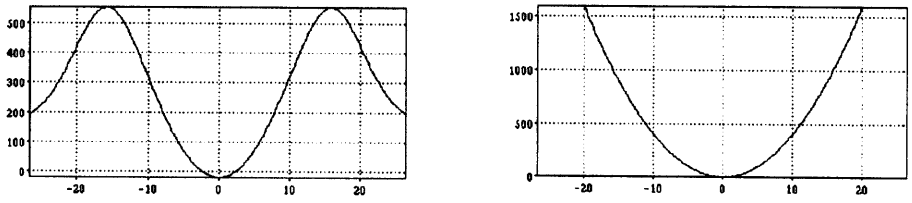

Figure 7: *The third energy potential for $\beta = 0.002$ and different values for the parameter ϵ. a) For $\epsilon = 0.9$ b) For $\epsilon = 0$ the potential becomes the effective potential of the Weak Membrane energy.*

that has the desired property and should be a reasonable guess for the dependence of h_{ij} on the gradient of the field f.
Figure 6 shows two solutions of (3.13) for different temperatures. Substituting the value of h_{ij} given by (3.13) in (3.11) we then obtain the third energy potential, that is plotted in Figure 7 for some values of β and ϵ.
The effect of the third energy potential on the mean field $\bar{f}$ can be understood by analyzing the "force" (derivative) of the potential (see Figure 8). The third energy effect can be described as follows:

1. Smooth the field for small values of the gradient $((\Delta_{ij}^h)^2 < (\Delta_0)^2)$. (Positive derivative of the potential.)

2. Enhance the discontinuity for $(\Delta_0)^2 < (\Delta_{ij}^h)^2 < \frac{\gamma}{\alpha}$, in other words, forcing the value of the field at a site to be pulled away from the value of the field at the considered neighbors. (Negative derivative of the potential.)

3. Neither smoothing nor enhancing the field for $(\Delta_{ij}^h)^2 > \frac{\gamma}{\alpha}$. (The derivative of the potential is zero.)

We point out that the third energy potential derived above is a local approximation of the mean field approximation over the discontinuity field.

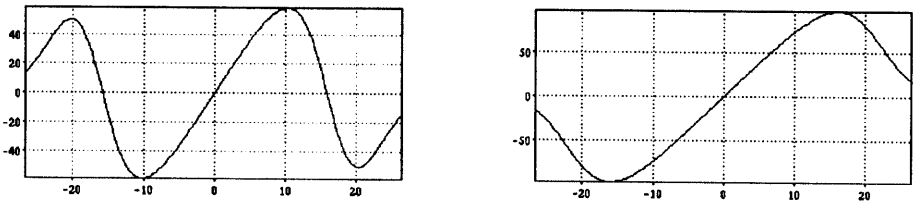

Figure 8: *The third energy force for β= 0.002 and different values of the parameter ϵ a)For ϵ=0.9 . b) For ϵ = 0 the force is the same as for the Weak Membrane energy.*

3.4.4 A deterministic solution for f

Here again the same arguments of section 3.3.3 apply to obtain a deterministic solution for the field f. After some computations we then obtain:

$$\begin{aligned}\bar{f}_{i,j} = \; & g_{i,j} - \alpha\overline{\Delta^v_{i,j}}(1-\bar{v}_{i,j}) + \alpha\overline{\Delta^v_{i,j+1}}(1-\bar{v}_{i,j+1}) \\ & -\alpha\overline{\Delta^h_{i,j}}(1-\bar{h}_{i,j}) + \alpha\overline{\Delta^h_{i+1,j}}(1-\bar{h}_{i+1,j}) \\ & +\alpha\epsilon\overline{\Delta^h_{i,j}}\bar{h}_{i,j}\sigma_\beta(\gamma-\alpha(\overline{\Delta^h_{i,j}})^2) \\ & -\alpha\epsilon\overline{\Delta^h_{i,j+1}}\bar{h}_{i,j+1}\sigma_\beta(\gamma-\alpha(\overline{\Delta^h_{i,j+1}})^2) \\ & +\alpha\epsilon\overline{\Delta^v_{i,j}}\bar{v}_{i,j}\sigma_\beta(\gamma-\alpha(\overline{\Delta^v_{i,j}})^2) \\ & -\alpha\epsilon\overline{\Delta^v_{i,j+1}}\bar{v}_{i,j+1}\sigma_\beta(\gamma-\alpha(\overline{\Delta^v_{i,j+1}})^2) \qquad (3.14)\end{aligned}$$

4 Parameters

The parameters α, γ, ϵ and β must be estimated in order to develop an algorithm that smoothes, enhances and finds the discontinuities of the given data-field.

4.1 The parameter α

The parameter α controls the balance between the "trust" in the data and the smoothing term. The noisier are the data the less you want to "trust" it so α is larger, the less noisy are the data the more you "trust" it so α should be smaller. To estimate α various mathematical methods are available. The

generalized cross validation method introduced by Wahba [17] and the standard regularization method described by Tikhonov [16, 1] were studied by Geiger and Poggio [7] to estimate α. The standard regularization method gives [7] an equation relating α with the signal-to-noise ratio that for the continuous case is

$$\int_0^\infty \int_0^\infty \alpha \frac{S(\omega,\nu)(\omega^2+\nu^2)^2}{(1+\alpha(\omega^2+\nu^2))^3} d\omega\, d\nu = \int_0^\infty \int_0^\infty \frac{N(\omega\,\nu)(\omega^2+\nu^2)}{(1+\alpha(\omega^2+\nu^2))^3} d\omega\, d\nu \quad (4.1)$$

where $S(\omega,\nu)$ is the spectral density (power spectrum) of the signal $f(x,y)$ and $N(\omega,\nu)$ is the spectral density of the noise $v(x,y)$, assuming that $v(x,y)$ is stationary.

When an estimation of the noise is not available, it is necessary to use the generalized cross validation method (GCV). GCV states that the optimal value of α can be obtained by minimizing the functional (here in one dimension)

$$V(\alpha) = \frac{1}{n} \sum_{i=0}^{n} \frac{[f_{n,\alpha}(t_i) - g_i]^2}{(1 - a_{kk}(\alpha))^2} \omega_k^2(\alpha) \quad (4.2)$$

where $f_{n,\alpha}(t_i)$ is the smoothed solution given by (3.4),

$$\omega_k(\alpha) = (1 - a_{kk}(\alpha))/(1 - \frac{1}{n} \sum_{j=0}^{n} a_{jj}(\alpha))$$

and $a_{kk}(\alpha) = \frac{\partial}{\partial f_k}(f_{n,\alpha})(t_k)$.
For the purpose of smoothing images both methods give about the same values of α [7]. Better results should be obtained if the estimation of α is done locally (for a small neighborhood).

4.2 The parameter γ

From (3.6) one can see that $\sqrt{\frac{\gamma}{\alpha}}$ is the threshold for creating a line in the Weak Membrane energy. Let's call ξ the value $\sqrt{\frac{\gamma}{\alpha}}$. From the expression of the effective potential we notice that if the gradient Δf_{ij} is above ξ there is no smoothing and if the gradient is below ξ then smoothing is applied. The

parameter ξ defines the resolution of the system and we explain this with two examples.

In the first example suppose we are working with the stereo module, so the data field is a depth-field. In this case ξ is the threshold for the changes in depth to be called a depth discontinuity. This value is determined according to the resolution of the stereo system available and/or to the desired resolution that one is interested.

In the second example the field is the intensity data and the parameter ξ represents the threshold for detecting edges. This value is somehow arbitrary, and probably context dependent. A variation of $\xi = 16$ units to $\xi = 32$ units for an 8-bit array image is likely to be in agreement with the human threshold. The exact value of ξ depends on the attention of the observer and/or the sensitivity of the system.

For the second energy the absolute threshold to create an edge is also given by ξ. However when there is support from neighbor edges this threshold can be lowered to $\xi \times (\sqrt{(1-\epsilon)}$ (see (3.10)). This suggests setting γ so to guarantee that the highest noise gap is smoothed (we are assuming that noise gaps are isolated features). In this case the ϵ parameter will be set to assure that at the edges the images will not be smoothed but perhaps enhanced. In this case the value of $\xi = 30$ may be desired for an 8-bit array.

4.3 The parameter ϵ

The parameter ϵ allows the energy to be more general by controlling the amount of propagation of the line. So, once a line is created, the price to pay for creating another line next to it will be lowered by the amount of $\gamma\epsilon$. In other words, from (3.11) one can see that the difference in the energy for when a line has been created and when no line has been created at pixel $(i-1, j)$, is given by $\gamma\epsilon$. This is what characterizes the threshold and suprathreshold or the hysteresis phenomena [3]. The threshold is given by $\sqrt{\frac{\gamma(1-\epsilon)}{\alpha}}$ and the suprathreshold by $\sqrt{\frac{\gamma}{\alpha}}$. The parameter ϵ varies from 0 to 1.0 and when is zero reduces the third energy to the Weak Membrane energy. When $\epsilon = 1.0$ lines are created everywhere, since once a line is created there is no cost in creating another one and then it propagates indefinitely (to the image limits). How much does one want to propagate a line? How much should the difference between the threshold and the suprathreshold be? Which exact value of ϵ

should be chosen? According to the discussion above about γ we conclude that ϵ should be chosen (assuming we already have α and γ) to guarantee that $\sqrt{\frac{\gamma(1-\epsilon)}{\alpha}}$ is below the desired edge threshold. Again it depends on what one wants to do with the data. For detecting edges of objects in a scene one wants to have high suprathreshold and threshold (usually object boundaries exhibit high gradients) and ϵ large (bigger than 0.5) so that all the object boundaries are detected, included the exceptional boundary pixels with a smaller gradient.

4.4 The parameter β

The parameter β controls the uncertainty of the model. The smaller is β the more inaccurate is the model. This suggests that for solving the mean field equations a rough solution can be obtained for a small value of β (high uncertainty) and thereafter we can increase β (small uncertainty) to obtain more accurate solutions. This can be called deterministic annealing.
We also notice that β multiply the first term on (3.1) to give

$$\beta E_{fg}(f) = \sum_{i,j} \beta (f_{i,j} - g_{i,j})^2$$

which suggests that the inverse of β gives the standard deviation of the noise. We keep in mind that β also multiplies the other parameters of the model and therefore by estimating β by the amount of noise the estimation of the other parameters have to be scaled by this factor.

5 Results

5.1 Implementation

For the implementation the zero temperature limit equations have provided results as good as the deterministic annealing with a faster computational time. We do not have proof of convergence but only suggestive experimental results.
In order to find the mean field solution for the third energy, we solved (3.14) together with (3.10) in a coupled and iterative way. More precisely:
For simplicity, let's write (3.14) as

$$\bar{f}_{i,j} = F(g_{i,j}, \bar{f}_{i,j}, \bar{f}_{i-1,j}, \bar{f}_{i,j-1}, \bar{f}_{i+1,j}, \bar{f}_{i,j+1}, h_{i,j}, h_{i-1,j}, v_{i,j}, v_{i,j-1})$$

and (3.10) as

$$h_{i,j} = H(\bar{f}_{i,j}, f_{i,j-1}, h_{i,j-1}, h_{i,j+1}) \quad v_{i,j} = V(\bar{f}_{i,j}, f_{i-1,j}, v_{i-1,j}, v_{i+1,j})$$

The algorithm solves these equations iteratively by updating (for the parallel case) in the following way

$$\bar{f}_{ij}^{n+1} = \bar{f}_{ij}^{n} - \omega\left[\bar{f}_{i,j}^{n} - F(g_{i,j}, \bar{f}_{i,j}^{n}, ..., \bar{f}_{i,j+1}^{n}, h_{i,j}^{n}, ..., v_{i,j-1}^{n})\right] \tag{5.1}$$

and

$$h_{i,j}^{n+1} = H(\bar{f}_{i,j}^{n}, f_{i-1,j}^{n}, h_{i,j-1}^{n}, h_{i,j+1}^{n}) \quad \text{and} \quad v_{i,j}^{n+1} = V(\bar{f}_{i,j}^{n}, f_{i,j-1}^{n}, v_{i-1,j}^{n}, v_{i+1,j}^{n})$$

where ω controls the rate of change and the index n indicates the step of the iterative process.
For a serial implementation we first update the even sites (like the white squares in a chess board) and then the odd sites (like the black squares). Typically the algorithm has converged in 10 iterations which takes about 1 minute for images of 64×64 pixels on a Symbolics 3600.

5.2 Synthetic Images

We first tested the algorithm on a synthetic step edge image with noise added. This is a good test-image since locally many edges on real images are of this type.
The step edge image is an 8-bit array of 64×64 pixels with a step intensity of 140 units (see Figure 9a). White noise, with standard deviation 30, is added to the step edge (see Figure 9b). We apply the algorithm described by (5.1) to reconstruct the original image. First we use $\epsilon = 0$, in order to have the Weak Membrane energy. We set $\alpha = 4$ (typical smoothing parameter) and $\gamma = 24000$ (so that we do not smooth the edge). The result after 20 iterations is shown in Figure 9c. We notice that the step edge starts to be smoothed before all the noise has been smoothed. In Figure 9d we set ϵ to be

0.9 and $\gamma = 45000$, and the result is more striking. Because we set the ratio $\frac{\gamma}{\alpha}$ to be higher than that, all the noise has been smoothed away. However, at the edge there is support from the neighbor edges to lower the threshold to a factor $(1 - \epsilon) = 0.1$. Therefore the noise is suppressed and the edges are not smoothed but on the contrary are enhanced. We conclude that for a noisy step edge, the third energy is able of retrieving and enhancing the edge (see figure 9 1d) while the Weak Membrane energy retrieves the edge with worse performance (see figure 9c).

5.3 Real Images

When we apply the second energy algorithm to a real still life image the result is an enhancement of specular edges, shadow edges and some other contours while smoothing out the noise (see Figure 10). This result is obtained consistentently in all the images we used and also with other fields besides intensity, e.g. "color" images. In our case "Color" images are 8-bit images representing the ratio between the red array and the sum of red and green arrays.

6 Surface reconstruction and alignment via integration

6.1 Sparse Data and Surface Reconstruction

Until now we dealt with dense data defined on allthe sites of the lattice. This is not always the case, especially if the field is the output of a stereo or motion module. In that case data are given on a subset of the lattice sites and the data-field interaction term (E_{fg}) in the energy function has to be modified in the following way

$$E_{fg} = \sum_{i,j}(f_{i,j} - g_{i,j})^2 \Rightarrow \sum_{i,j}(f_{i,j} - g_{i,j})^2 \gamma_{i,j}.$$

Here $\gamma_{i,j}$ is a flag that is one if a datum is given at site (i, j) and 0 otherwise. This slight modification has no effect on the theoretical results, and only some changes take place in the deterministic equations for the surface field: the term enforcing closeness to data disappears when a datum

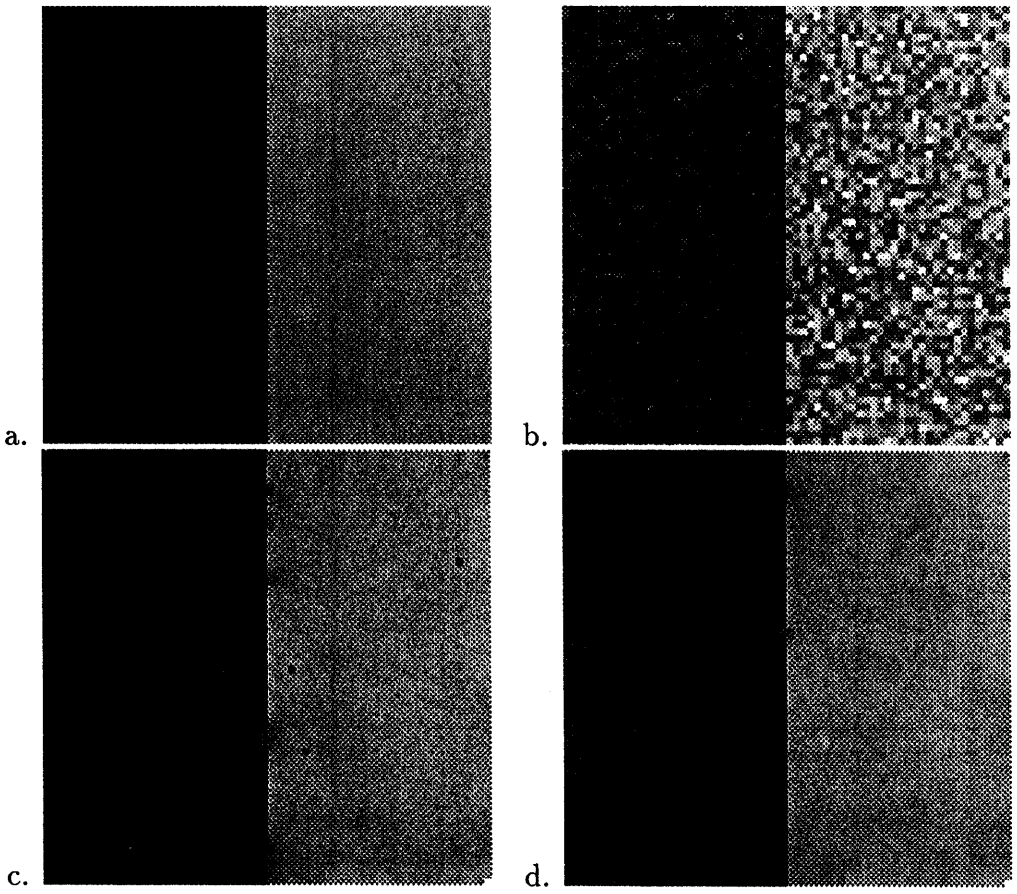

Figure 9: *a) Step edge with 140 grey value units for the step. 1b) White noise with standard deviation 30 grey value units has been added. c) The noisy image after 20 iterations for* $\alpha = 4$, $\gamma = 24000$, $\epsilon = 0$. *d) The noisy images after 20 iteration for* $\alpha = 4$, $\gamma = 45000$, $\epsilon = 0.9$

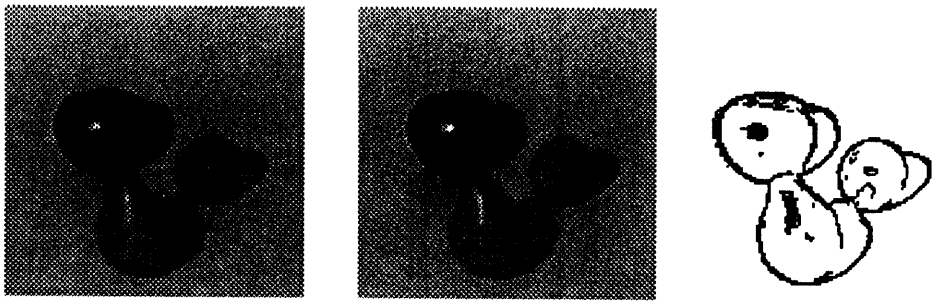

Figure 10: *a) Still life image 128 × 128 pixels. b) The image smoothed with* $\epsilon = 0.5$, $\gamma = 1400$ *and* $\alpha = 4$ *for 9 iterations c) The line process field (without thinning).*

is not given at a site. We rewrite here the deterministic solution for the field f in the case of the Weak Membrane Energy, since the third energy equations are modified in the same way. In the case of sparse data (3.7) becomes then

$$\begin{aligned}\bar{f}_{i,j}\gamma_{i,j} = \; & g_{i,j}\gamma_{i,j} - \alpha(\bar{f}_{i,j} - \bar{f}_{i,j-1})(1 - \bar{v}_{i,j}) + \alpha(\bar{f}_{i,j+1} - \bar{f}_{i,j})(1 - \bar{v}_{i,j+1}) \\ & -\alpha(\bar{f}_{i,j} - \bar{f}_{i-1,j})(1 - \bar{h}_{i,j}) + \alpha(\bar{f}_{i+1,j} - \bar{f}_{i,j})(1 - \bar{h}_{i+1,j})\end{aligned}$$

To apply this algorithm one needs first to fill in the data. We chose to fill in the data by averaging next neighbors and applying the algorithm at the same time. So at each step of the algorithm each lattice site is visited: if there is no field value the average neighbor value is taken; otherwise we apply the above algorithm. We notice that no action is taken if at a particular site there is no field value and no neighbor field value.
From one face image we produced sparse data by randomly suppressing 70 % of the data (see Figure 11). We then applied the third energy algorithm to sparse data. The parameters were kept the same as the other real image. The reconstruction from sparse data can be applied to depth data in which case it is usually called surface reconstruction. We used a stereo algorithm based on zero crossings to obtain depth data from the image shown below. Then we applied the algorithm to reconstruct the depth surface. The parameters were different according to the criteria for depth discontinuity.

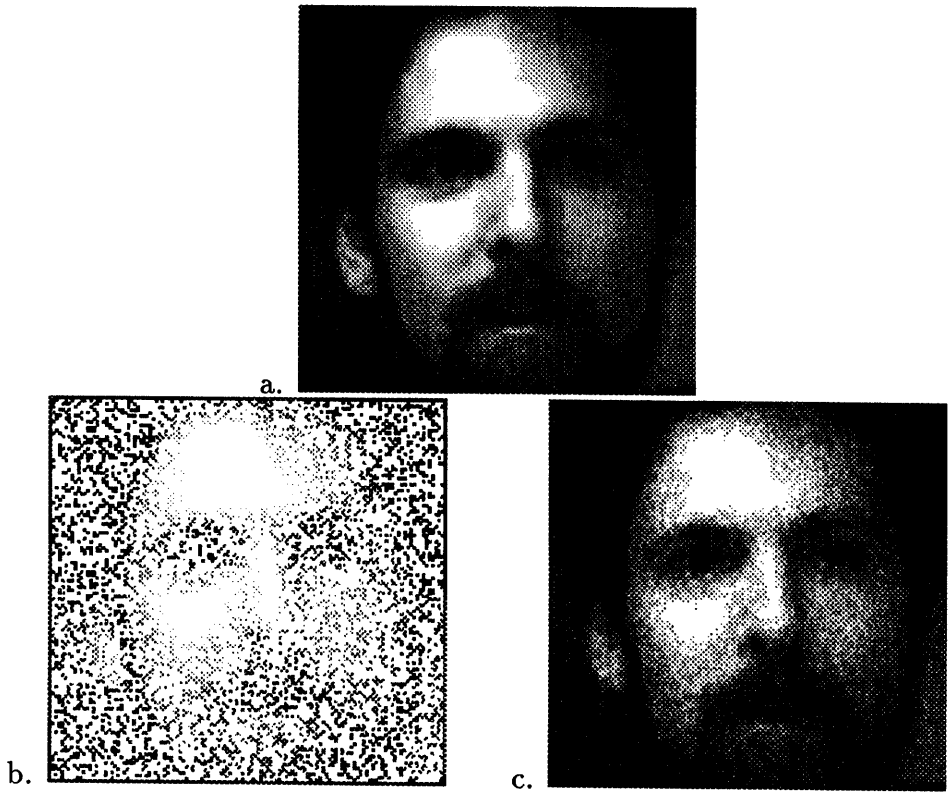

Figure 11: *a) A face image of 8-bits and 128 X 128 pixels. b) Randomly chosen 70 % of the original image. For display the other 30% are filled with white dots. c) The algorithm described above is applied to smooth and fill in at the same time with* $\epsilon = 0.9$, $\gamma = 1400$ *and* $\alpha = 4$ *for 70 iterations.*

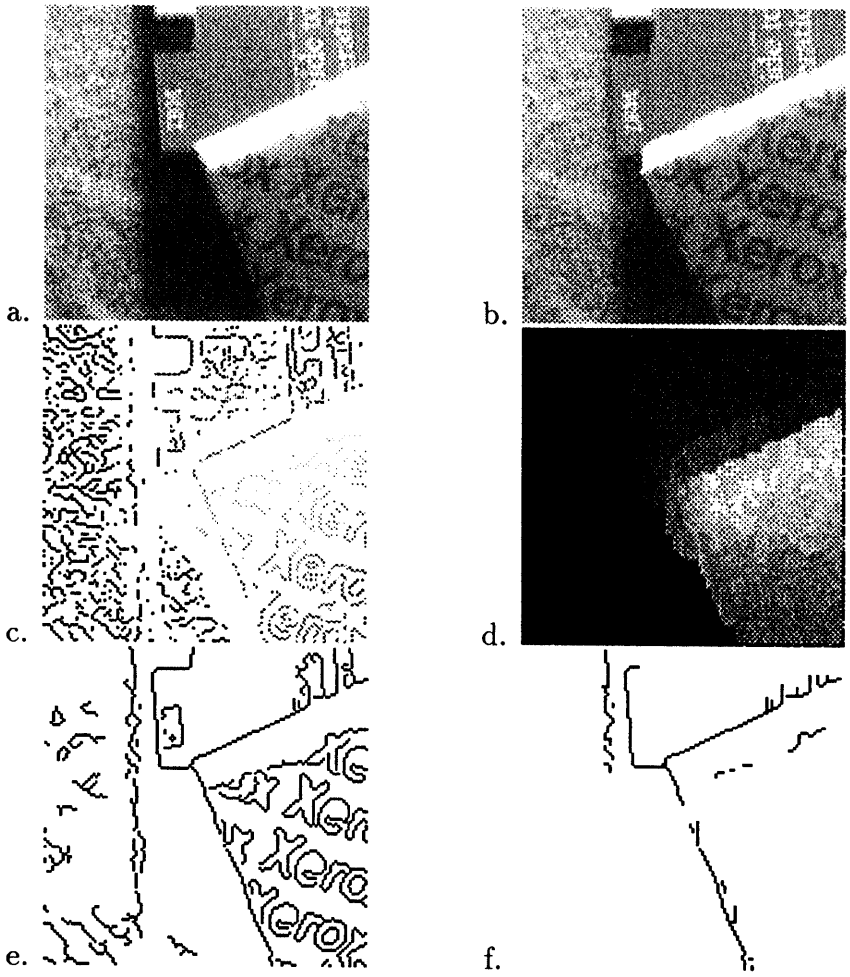

Figure 12: *a) The left image (8-bits and 128 × 128 pixels). b) The right image of the same scene c) Sparse depth data obtained by the stereo algorithm based on zero crossing. Intensity represents depth values. d) The algorithm is applied with* $\epsilon = 0.0$, $\gamma = 400$ *and* $\alpha = 4$ *to reconstruct the surface. e) Canny edges for figure a). f) The alignment algorithm is applied with the same parameters as d) and* $\delta = 384$. *Depth discontinuities aligned with the intensity edges are obtained.*

6.2 Alignment of visual modules with intensity edges

The integration of different visual modules to improve the detection of the discontinuities can also be addressed in this scheme. As suggested by Gamble & Poggio [6] , we can add the term $\delta\,(v_{ij}+h_{ij})(1-e_{ij})$ to the Weak Membrane Energy or the second energy. Here e_{ij} is an external field, for example the edge map that is coupled with the stereo field. For implementation purposes the only consequence of adding this term is the change of the global parameter γ into the local parameter $\gamma'_{ij} = \gamma - \delta(1 - e_{ij})$. R. Thau implemented this scheme in the Connection Machine using Canny's intensity edges for e_{ij}. We first took a pair of images (Figure 12) and used a stereo algorithm to find depth at the zero crossing positions. We then used the algorithm to reconstruct depth everywhere and detect depth discontinuities. As we expected depth discontinuities are aligned with the intensity edges.

7 Conclusion

We have used statistical mechanics tools to derive deterministic approximations of Markov random fields models. In particular we have studied an energy model that is suitable for image reconstruction or any field reconstruction. The model has been developed to include the following characteristics:

- the surface field is smoothed when its gradient is not too high,
- contrast will be enhanced where a discontinuity occurs (if it is not too large already),
- the discontinuity field is likely to be smooth (isolated discontinuities are inhibited),
- hysteresis and adaptive multiple threshold arise naturally from the model.
- three parameters are needed to specify the model,
- when one of the parameters (ϵ) is set to zero the *weak membrane energy* is recovered.
- An understanding of the role of the parameters is possible,

- The model can deal with sparse data and alignment of the discontinuities of different modules with the intensity edges.

We have shown that the deterministic algorithm of GNC can be regarded as an approximation of the gradient descent method with a deterministic annealing schedule to solve the mean field equations. This suggests a unified framework to connect different methods used on image segmentation, restoration and surface reconstruction. We will show in another paper[8] that several deterministic algorithms for image segmentation and reconstruction are approximations of two methods to solve the mean field equations: the gradient descent method discussed in this paper and the parameter space method discussed in [8]. We derived a deterministic solution for the mean values of the surface and discontinuity fields, consisting of a system of coupled nonlinear equations. An algorithm has been implemented to obtain a solution for this system: it is fully parallelizable, iterative and recursive, allowing efficient computation. It would be interesting to analyze other approximations or extensions of this model. For example, a model that includes interaction between the horizontal and vertical line processes could be developed to inhibit self-intersections of the discontinuity field. A term like $h_{ij}(1 - v_{ij})$ may be sufficient. An analysis of convergence of the algorithm would also be important.

Acknowledgments We are grateful to Tomaso Poggio for his guidance and support. Suggestions from Eric Grimson, Ellen Hildreth and Woody Yang were extremely valuable, Ed Gamble and Alessandro Verri were always around with helpful comments.

Appendix A

Given an energy function of the form

$$E(f) = \sum_i \{(f_i - g_i)^2 + [\sum_{j \epsilon N_i} V(f_i, f_j)]\}$$

where N_i is the set of neighbor sites of i, $V(f_i, f_j)$ is an interaction potential and g_i is an external field (data), the partition function is given by

$$Z = \sum_f e^{-\beta E(f)}$$

where the $\sum_f$ means a sum over all the possible configurations of the system, that is a multidimensional integral over all the variables f_i, in this case. The mean field equation for f_k is given by

$$\bar{f}_k = \frac{1}{Z} \sum_f f_k e^{-\beta(\sum_i \left\{(f_i - g_i)^2 + \sum_{j \epsilon N_i} V(f_i, f_j)\right\})} \qquad (A1)$$

From this definition, and the definition of the partition function Z , the following equality is derived:

$$\frac{1}{Z} \frac{\partial Z}{\partial g_k} = -2\beta(\bar{f}_k - g_k) \qquad (A2)$$

.

Defining as usual the free energy F as $F = -\frac{1}{\beta} ln Z$ we can now obtain, from (A2)

$$\bar{f}_k = g_k - \frac{1}{2} \frac{\partial F}{\partial g_k} \qquad (A3)$$

In the case of energy E the mean-field approximation consists in replacing the fields at the neighbor sites with their statistical mean value $\bar{f}_j$. Therefore the mean-field approximation E^{mf} to the energy function becomes

$$E^{mf}(f) = \sum_i \{(f_i - g_i)^2 + [\sum_{j \epsilon N_i} V(f_i, \bar{f}_j)]$$

This is a good approximation of the original energy when the fluctuations of the field f are small, which is the case in our experiments.

Appendix B

In this appendix we will derive an expression for the mean value of the h and v fields of a system described by the Weak Membrane Energy

$$E_1(f,h,v) = E_{fg}(f) + E_{fl}(f,h,v) + E_l(h,v).$$

Noticing that

$$E_1(f,h,v) = E_1(f) + \sum_{i,j}[h_{i,j}G^h_{i,j} + v_{i,j}G^v_{i,j}]$$

where the energy function E_1 has been defined in section 1, the partition function can be written as

$$Z = \sum_{\{f\}} e^{-\beta E_1(f)} \sum_{\{h,v\}} e^{-\beta \sum_{i,j}[h_{i,j}G^h_{i,j} + v_{i,j}G^v_{i,j}]}.$$

Let us consider just the field h. By definition the mean value $\bar{h}_{i,j}$ is

$$\bar{h}_{i,j} = \frac{1}{Z}\sum_{\{f\}} e^{-\beta E_1(f)} \sum_{\{h,v\}} h_{i,j} e^{-\beta \sum_{i,j}[h_{i,j}G^h_{i,j} + v_{i,j}G^v_{i,j}]}. \qquad (B1)$$

Then we can rewrite (B1) in the following way:

$$\bar{h}_{i,j} = -\frac{1}{\beta Z}\sum_{\{f\}} e^{-\beta E_1(f)} \frac{\partial}{\partial G^h_{i,j}} Z_{hv}(f) \qquad (B2)$$

where

$$Z_{hv}(f) = \sum_{\{h,v\}} e^{-\beta \sum_{i,j}[h_{i,j}G^h_{i,j} + v_{i,j}G^v_{i,j}]}.$$

Z_{hv} represents the contribution of the line process to the partition function of the whole system for each fixed configuration of the field f and can be seen as the partition function of the line process system when the field f is kept "frozen". We notice that the line process system is simply a classical, non-interactive spin system in an external field. Since there is not interaction between the spin variables, $Z_{hv}(f)$ can be easily computed, giving

$$Z_{hv}(f) = \prod_{i,j}(1 + e^{-\beta G^h_{i,j}})(1 + e^{-\beta G^v_{i,j}}).$$

We can now define the free energy of this spin system by the usual relation

$$Z_{hv}(f) = e^{-\beta F_{hv}(f)}$$

and substituting in (B2) we obtain the following equation

$$\bar{h}_{i,j} = \frac{1}{Z} \sum_{\{f\}} e^{-\beta[E_1(f)+F_{hv}(f)]} \frac{\partial}{\partial G^h_{i,j}} F_{hv}(f). \qquad (B3)$$

Before proceeding further we now notice that, having defined $F_{hv}(f)$, the partition function of the whole system can now be rewritten as

$$Z = \sum_{\{f\}} e^{-\beta[E_1(f)+F_{hv}(f)]}.$$

This can be regarded as the partition function of a system, composed just by the field f, whose energy function is $E_1(f) + F_{hv}(f)$. Now, indicating with $< ... >$ the statistical mean value over all the possible configurations of this system, (B3) can be rewritten now in the following way:

$$\bar{h}_{i,j} =< \frac{\partial F_{hv}(f)}{\partial G^h_{i,j}} >$$

and then

$$\bar{h}_{i,j} =< \frac{1}{1 + e^{\beta G^h_{i,j}}} > . \qquad (B4)$$

A similar expression holds for $\bar{v}_{i,j}$. The meaning of these equations is the following: if the field f is kept "frozen" the mean value of the horizontal line process is simply $\frac{\partial F_{hv}(f)}{\partial G^h_{i,j}}$, but to obtain the true mean value $\bar{h}_{i,j}$ we have to average $\frac{\partial F_{hv}(f)}{\partial G^h_{i,j}}$ over all the possible configurations of the field f, whose probability distribution is given by

$$P(f) = \frac{e^{[E_1(f)+F_{hv}(f)]}}{Z}$$

where the term $F_{hv}(f)$ is due to the interaction of the line process with the field f.

References

[1] M. Bertero, T. Poggio, and V. Torre. Ill-posed problems in early vision. Technical report. Also Proc. IEEE, in press.

[2] Andrew Blake and Andrew Zisserman. *Visual Reconstruction.* MIT Press, Cambridge, Mass, 1987.

[3] John F. Canny. Finding lines and edges in images. Technical Report TM-720, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 1983.

[4] Paul B. Chou and Christopher M. Brown. Multimodal reconstruction and segmentation with Markov random fields and HCF optimization. In *Proceedings Image Understanding Workshop*, pages 214–221, Cambridge, MA, February 1988. Morgan Kaufmann, San Mateo, CA.

[5] E. Gamble, D. Geiger, T. Poggio, and D. Weinshall. Integration of vision modules and labeling of surface discontinuities. *Invited paper submitted to IEEE Trans. Sustems, Man & Cybernetics*, December 1989.

[6] Edward B. Gamble and Tomaso Poggio. Visual integration and detection of discontinuities: The key role of intensity edges. A.I. Memo No. 970, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, October 1987.

[7] Davi Geiger and Tomaso Poggio. An optimal scale for edge detection. In *Proceedings IJCAI*, August 1987.

[8] Davi Geiger and Alan Yuille. A common framework for image segmentation and surface reconstruction. In preparation.

[9] Stuart Geman and Don Geman. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6:721–741, 1984.

[10] S. Kirkpatrick, C.D. Gelatt, and M.P. Vecchi. Optimization by simulated annealing. *Science*, 220:219–227, 1983.

[11] Christof Koch, Jose Marroquin, and Alan Yuille. Analog 'neuronal' networks in early vision. A.I. Memo No. 751, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 1985.

[12] Jose L. Marroquin. Deterministic Bayesian estimation of Markovian random fields with applications to computational vision. In *Proceedings of the International Conference on Computer Vision*, London, England, June 1987. IEEE, Washington, DC.

[13] Jose L. Marroquin, Sanjoy Mitter, and Tomaso Poggio. Probabilistic solution of ill-posed problems in computational vision. In L. Baumann, editor, *Proceedings Image Understanding Workshop*, pages 293–309, McLean, VA, August 1985. Scientific Applications International Corporation.

[14] N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller. Equation of state calculations by fast computing machines. *J. Phys. Chem*, 21:1087, 1953.

[15] G. Parisi. *Statistical Field Theory.* Addison-Wesley, Reading, Massachusets, 1988.

[16] A. N. Tikhonov and V. Y. Arsenin. *Solutions of Ill-posed Problems.* W.H.Winston, Washington, D.C., 1977.

[17] G. Wahba. Practical approximate solutions to linear operator equations when the data are noisy. *SIAM J. Numer. Anal.*, 14, 1977.