Parallel and deterministic algorithms from MRFs: surface reconstruction and integration

Davi Geiger[1,2] and Federico Girosi [1]

[1] Massachusetts Institute of Technology
Artificial Intelligence Laboratory, and

[2] Siemens Corporate Research, Inc,
755 College Road East, Princeton, NJ 08540

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. The input to these systems are usually noisy and some times sparse as well as its output and then more processing is necessary to extract the relevant information and fill in sparse data. In this way the problem of surface reconstruction starting from a set of noisy sparse data is prototypical for vision. In recent years many researchers[10][13][5] [3][4] have investigated the use of Markov random fields (MRFs) for early vision. 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 resides on Bayes theory with local interaction between the fields, where the posterior probability distribution for the configuration of the fields, given a set of data, is given as a Gibbs distribution. The model is then specified by the a priori information about the system and a conditional probability of the data given the fields. In the standard approach an estimate of the field and its discontinuities is given by the configuration that maximizes the probability distribution. This becomes a combinatorial optimization problem, that can be solved by methods of the Monte Carlo type (simulated annealing[11], for example). The MRFs formulation has two main drawbacks: the amount of computer time needed for the implementation and the difficulty in estimating the parameters of the model.

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 use the *mean field approximation* and the *saddle point approximation*, both well known statistical mechanics tools, 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.

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 study the Weak Membrane model that has been already studied by several authors[2][12][14][13]. It is interesting to notice that the GNC algorithm, proposed by Blake and Zisserman [2], arises naturally in the framework of statistical mechanics. This estabilish a connection between MRFs and deterministic algorithm already used in vision. The model is applied to dense data

and 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 that can be used for the integration of different modules.

The paper is organized in the following way: section 2 presents an overview of MRFs in vision. section 3 discusses the deterministic approximation of MRFs for the three energy functions mentioned above. Section 4 dicusses the issue of parameters estimation. In section 5 some results are described including sparse data. Section 6 shows applications for the integration of visual modules with intensity edges. Section 7 concludes the paper.

2 MRFs and the Weak Membrane model

Here we briefly summarize how MRFs are applied to the Weak Membrane model. A more extensive discussion is given in Geman and Geman [10], Marroquin[13], Chou[3], Gamble and Poggio[5], Gamble, Geiger, Poggio, Weinshall [4] and Geiger [9].

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 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:

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

Geman and Geman [10] 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).

As a simple example, when the surfaces (surface-fields) are expected to be smooth but not at the discontinuities and the noise is Gaussian, the energy, for the one dimensional case, is given by

$$E(f, l|g) = \sum_i \{\lambda_i (f_i - g_i)^2 + \alpha \sum_{j \epsilon N_i} [(f_i - f_j)^2 (1 - l_{ij})] + \sum_C V_C(l_{ij})\}, \tag{2.3}$$

where $\lambda_i = 1$ 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. 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 ([10], [14]). In our models the cliques will be simplified to the nearest neighbors.

2.1   The Line Process for two dimensions

In this case we define a horizontal line process h_{ij} and a vertical line process v_{ij}. We point out that another possible approach is considered by Geiger and Yuille [8] and Geiger[9], where the line process is a scalar quantity in two dimensions. Any how, 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 $(i - 1, j)$.

2.2   The Weak Membrane model

A special case of (2.3) is the Weak membrane model given in two dimensions, for sparse data, by

$$E_1(f, h, v) = E_{fg}(f) + E_{fl}(f, h, v) + E_l(h, v) \qquad (2.4)$$

where

$$E_{fg}(f) = \sum_{i,j} \lambda_{ij}(f_{i,j} - g_{i,j})^2 \quad , \quad E_l(h, v) = \gamma \sum_{i,j} (h_{i,j} + v_{i,j}) \qquad (2.4a)$$

$$E_{fl}(f, h, v) = \alpha \sum_{i,j} \left[(f_{i,j} - f_{i,j-1})^2 (1 - v_{i,j}) + (f_{i,j} - f_{i-1,j})^2 (1 - h_{i,j}) \right] \qquad (2.4b)$$

α and γ are positive valued parameters and λ_{ij} is 1 when there is data and zero otherwise.

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.

The maximum of the posterior distribution (MAP) 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 [15]. These algorithms sample the space of possible values of the surface-field according to the probability distribution $P(f|g)$.

3   A deterministic approximation of MRFs

3.1   Mean field theory and Weak Membrane

We assume that there is uncertainty in the model and that (2.4) should be understood within the context of (2.2). We then propose to estimate the mean field values from the statistical model. The mean field value of f, $\bar{f}$, is given by

$$\bar{f}_{lk} = \frac{1}{Z} \sum_{f,h,v} f_{lk} e^{-\beta(\sum_{i,j}[\lambda_{ij}(f_{ij}-g_{ij})^2 + E_{fl} + E_l])}$$

where $\sum_{f,h,v}$ represents the sum over all possible configurations of the fields f, h and v. A configuration of the field f is a possible set of values assumed by f in the lattice. Z is given by

$$Z = \sum_{f,h,v} e^{-\beta E_1(f,h,v)}$$

From these definitions the following equality is derived:

$$-2\beta\lambda_{lk}(\bar{f}_{lk} - g_{lk}) = \frac{1}{Z}\frac{\partial Z}{\partial g_{lk}} \tag{3.1}$$

We still have to compute the partition function Z. In the case of (2.4) Z becomes

$$Z = \sum_{\{f\}} e^{-\beta\sum_{i,j}[\lambda_{ij}(f_{i,j}-g_{i,j})^2 + \gamma]} \sum_{\{h,v\}} e^{-\beta\sum_{i,j}[(h_{i,j}-1)G^h_{i,j} + (v_{i,j}-1)G^v_{i,j}]} \tag{3.2}$$

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 $\Delta^v_{i,j} = f_{i,j} - f_{i,j-1}$.

3.1.1 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 (3.2) 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

$$Z = \sum_{\{f\}} e^{-\beta\sum_{i,j}[\lambda_{i,j}(f_{i,j}-g_{i,j})^2 + \gamma]} \prod_{ij}(1 + e^{\beta G^h_{i,j}})(1 + e^{\beta G^v_{i,j}}) \tag{3.3} .$$

3.1.2 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 (3.3) we notice that the partition function can be rewritten as

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

where

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

and $E_{fg}(f)$ is given by (2.4a). 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
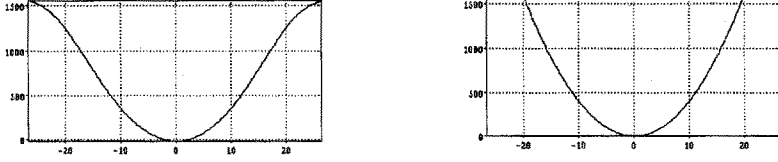
Figure 1: *A slice of the effective potential as a function of Δ_{ij}^h. a) For $\beta = 0.002$. b) Zero temperature limit ($\beta \to \infty$).*

appropriately the interaction of the field with itself, substituting the smoothing term in the energy function with a new temperature dependent potential.

In figure 1 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 1 that the smoothing term is active only when the gradient is smaller than a threshold, proportional to the ratio between γ and α.

3.1.3 The saddle point approximation

The saddle point approximation substitutes the partition function in (3.3) by its largest term, so

$$Z \approx max_f e^{-\beta \sum_{ij}[\lambda_{ij}(f_{ij}-g_{ij})^2 + E_{eff}(f)]} = e^{-\beta \sum_{ij}[\lambda_{ij}(<f_{ij}>-g_{ij})^2 + E_{ij}^{eff}(<f>)]} \qquad (3.4)$$

where $< f >$ minimizes the free energy

$E_{fg} + E_{eff}(f) = \sum_{ij} \lambda_{ij}(f_{ij} - g_{ij})^2 + \gamma - \frac{1}{\beta}ln[(1 + e^{\beta G_{i,j}^h})(1 + e^{\beta G_{i,j}^v})]$. Under this approximation $< f >$ is the mean field of f.

3.1.4 MF equations for h, v, and f

Once we obtained the partition function in (3.4) we can now compute the mean field values of h, v, and f.

● *Solving for the line process, h and v*

In analogy to (3.1) we can derive the mean field equation for h to be

$$\bar{h}_{lk} = 1 - \frac{1}{\beta} \frac{\partial ln Z}{\partial G_{lk}^h}$$

and similarly we find the mean field equation for v. So after some algebra we derive

$$\bar{h}_{i,j} = \frac{1}{1 + e^{\beta(\gamma - \alpha(\bar{f}_{i,j} - \bar{f}_{i-1,j})^2)}} \quad \text{and} \quad \bar{v}_{i,j} = \frac{1}{1 + E^{\beta(\gamma - \alpha(\bar{f}_{i,j} - \bar{f}_{i,j-1})^2)}} \qquad (3.5).$$

where $\bar{f} = < f >$ is the mean field solution (under the saddle point approximation) that we calculate next. It is interesting to notice that the mean value of the line process can vary continuously from 0 to 1. Moreover the value of l that is defined everywhere in the lattice, stresses the strength of the edges. This can be used to decide the existence or not of an

edge and to analyze its shape. In the zero temperature limit ($\beta \to \infty$) (3.5) 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.

• *Solving for the field f*

The mean field solutions for f 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

$$\lambda_{ij}\bar{f}_{i,j} = \begin{array}{l} \lambda_{ij}g_{i,j} - \alpha(\bar{f}_{i,j} - \bar{f}_{i,j-1})(1 - \bar{v}_{i,j}) + \alpha(\bar{f}_{i,j+1} - \bar{f}_{i,j})(1 - \bar{v}_{i,j+1}) \\ -\alpha(\bar{f}_{i,j} - \bar{f}_{i-1,j})(1 - \bar{h}_{i,j}) + \alpha(\bar{f}_{i+1,j} - \bar{f}_{i,j})(1 - \bar{h}_{i+1,j}) \end{array} \qquad (3.6)$$

where $\bar{h}_{i,j}$ and $\bar{v}_{i,j}$ are given by (3.5).

Equation (3.6) 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 section 4. Notice that the form of (3.6) is suitable for the application of a fast, parallel and iterative scheme of solution. In order to solve (3.6) we can introduce a damping like force ($\frac{\partial f}{\partial t}$) to the effective potential such that the fixed point of the dynamic equation

$$\frac{\partial f}{\partial t} = -\frac{\partial}{\partial f_{i,j}}(E_{fg} + E_{eff}(f)) \qquad (3.7)$$

is the mean field solution. This is a gradient descent algorithm.

3.1.5 The effective potential and the graduated 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 (graduated non convexity) algorithm that 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, as in (3.7), 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.

4 Parameters

The parameters α, γ, and β must be estimated in order to develop an algorithm that smoothes 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] give good pratical results. For a more detailed analysis see, for example, Geiger and Poggio [7].

4.2 The parameter γ

From (3.5) one can see that $\sqrt{\frac{\gamma}{\alpha}}$ is the threshold for creating a line in the Weak Membrane energy. For the stereo module, where the data field is a depth-field, $\sqrt{\frac{\gamma}{\alpha}}$ 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. For the intensity data, the parameter $\sqrt{\frac{\gamma}{\alpha}}$ represents the threshold for detecting edges. This value is somehow arbitrary, and probably context dependent. The exact value depends on the attention of the observer and/or the sensitivity of the system.

4.3 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.

5 Results

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
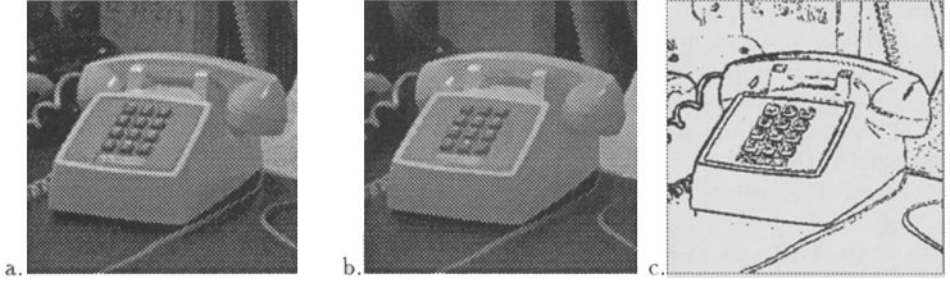
Figure 2: *a. 8-bit image of 256 X 256 pixels. b. The image smoothed with $\alpha = 1$, $\gamma = 19$, 10 iterations. c. The line process field for $\alpha = 0.1$, $\gamma = 19$, and 10 iterations.*
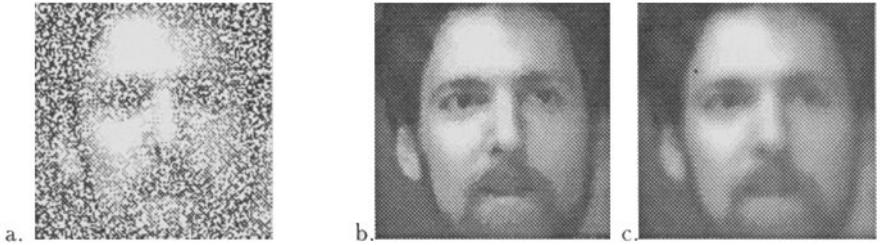

Figure 3: *a. Randomly chosen 50 % of an original image. For display the other 50% are filled with white dots. b. The algorithm described in (3.6) is applied to smooth and fill in at the same time with $\gamma = 208$ and $\alpha = 4$ for 10 iterations. c. For comparison and to stress the importance of the line process field we show the result of simply bluring the sparse data, when $h = v = 0$ everywhere*

convergence but only suggestive experimental results. In order to find the mean field solution we solved (3.6) using (3.7) together with (3.5) in a coupled and iterative way.

The algorithm is parallel depending upon the first neighbors. 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.

When we apply the 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 2).

From one face image we produced sparse data by randomly suppressing 50 % of the data (see Figure 3). We then applied the Weak Membrane model to sparse data. The parameters were kept the same as the other real image. We also compared the results with simply bluring the data (no line process).

The reconstruction from sparse data can be applied to depth data in which case it is usually called surface reconstruction.

6 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 [5] , we can add the term $\delta\,(v_{ij} + h_{ij})(1 - e_{ij})$ to the Weak Membrane model. 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 $\gamma'_{ij} = \gamma - \delta(1 - e_{ij})$.

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.

- We derived a deterministic solution for the mean values of the surface and discontinuity fields, consisting of a system of coupled nonlinear equations. The "key" step was the elimination (averaging out) of the binary fields (h and v) from the "energy" function. An algorithm has been implemented to obtain a solution for this system: it is fully parallelizable, iterative and recursive, allowing efficient computation.

- An understanding of the role of the parameters is possible,

- 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.

- We extended the model to deal with sparse data and alignment of the discontinuities of different modules with the intensity edges.

This work suggests a unified framework to connect different methods used on image segmentation, restoration and surface reconstruction. We 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]. In another paper [6] we analyze possible extensions to this model to include propagation of lines.

Acknowledgments We are grateful to Tomaso Poggio for his guidance and support.

References

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

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

[3] P. B. Chou and C. 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.

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

[5] E. B. Gamble and T. 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.

[6] D. Geiger and F. Girosi. Parallel and deterministic algorithms for mrfs: surface reconstruction and integration. A.I. Memo No. 1114, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, May 1989.

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

[8] D. Geiger and A. Yuille. A common framework for image segmentation and surface reconstruction. Harvard Robotics Laboratory Technical Report 89-7, Harvard, August 1989.

[9] Davi Geiger. *Visual models with statistical field theory*. PhD thesis, Massachusetts Institute of Technology, 1989.

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

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

[12] C. Koch, J. Marroquin, and A. Yuille. Analog 'neuronal' networks in early vision. *Proc. Natl. Acad. Sci.*, 83:4263–4267, 1985.

[13] J. 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.

[14] J. L. Marroquin, S. Mitter, and T. 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.

[15] 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.

[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.