Incorporating knowledge via regularization theory: applications in vision and image processing.

David Suter and Harvey A. Cohen
La Trobe University
Bundoora, Australia 3083

Abstract

In computer vision one is faced with the task of discovering the nature of objects that produced the light intensity distribution received by the camera. In order to solve this problem, one usually uses knowledge of the optics, the reflectance properties of surfaces, and of the structure of the objects one is expecting to find in a given scene. In many problems in image processing one is trying to use knowledge of the imaged objects, or of the image formation process, to reconstruct a better image from a degraded image.

A proper computational formulation of these problems recognises that they involve inverse processes that are mathematically ill-posed. This allows one to derive systematically, computational schemes based upon regularization theory that ensure existence of solution and stability of inversion process.

Such approaches often suggest particular types of algorithms for efficient solution to the problems. These, in turn, often are suggestive of implementations that are highly parallel, require only local information and simple operations. These implementations have the essential features of neural network approaches to computation; we present simulations of these.

Keywords and phrases: Computer Vision, Image Processing, Regularization, Ill-Posed, Neural Network.

1 Introduction

This paper will outline the application of regularization theory to the solution of several problems in image processing and in early vision. Emphasis will be placed on demonstrating how this provides a framework for problem solution at the three levels of abstraction: problem formulation, algorithm development, and implementation. This corresponds to the three levels of Marr's theory of vision ([2] pp. 24). Indeed, we can see that many of the ad hoc assumptions invoked by earlier workers in vision can be seen as attempts to regularise the inversion process; in particular, the often invoked smoothness constraints correspond to standard Tikhonov stabilizing functionals.

Section **2** introduces regularisation theory and its connection with incorporating knowledge. Then in section 3, the stochastic extension of standard regularization theory is "shown to be related to statistical physics and neural nets. For concreteness, we illustrate this with a simple image restoration example. Section **4** presents some of our work in investigating these approaches in image restoration and computer vision. The last section discusses some of the problems still to be solved, and some of the limitations our work has shown.

The framework presented, naturally links the mathematics of inversion with the physics of large scale systems and the emerging neural network theory.

2 Inverse Problems: regularization using prior information.

In this section we outline how regularization theory provides a framework for formulation of the problems of inverse problems in image processing and vision. The next section extends this standard theory to the more recently proposed probabilistic regularization theory; and thus demonstrates how this problem formulation suggests algorithms and implementations through the connection with theories of statistical physics and neural networks.

The most general formulation of inverse problems is that one is given the sensed data g, which is produced through the action of an operator A, acting upon the data we wish to recover f:

$$Af = g.$$

In image processing f is a corrupted version of g and A is the imaging process. In the vision case the sensed data is the raw pixel intensities, the data of interest f describes the objects of interest in the scene (e.g. locus of object surfaces), and the operator A is typically a composition of operators describing the image forming process. A problem involving inversion of this operator is well-posed [12] [3] if we can ensure existence, uniqueness, and stability of solutions. For a variety of reasons, failure of one or more of these conditions is common and thus the problem is ill-posed. For example A may not be full rank (so that the solution is not unique - extra information may be required to restrict the solution space), or A may be invertible but ill-conditioned (thus small changes in data lead to large deviations in solution - which can be disastrous in the presence of noise), or A may be of rank greater than the number of degrees of freedom (the system is overdetermined and thus a solution may not exist if any of the measurements contain noise). Typically, the imaging process does not preserve all information (it is a 2-D projection of a 3-D world and some objects are partially occluded) and thus the problem is ill-posed in the first sense given. Additional constraints and assumptions restrict the solution space, but in the presence of noise, the problem can then become ill-posed in either of the last two senses.

Standard regularization theory provides mathematical tools that enable one zc... turn an ill-posed problem into a well posed problem. One method that has been applied often in early vision **[4] [5]** is to replace the above problem with one of finding the solution that minimizes:

$$|Af - g| + \lambda |Pf|$$

where the first term measures fidelity with data and the second term involves an operator that measures suitability of the solution according to additional constraints (e.g. smoothness of solution). The last term restricts the solution to an admissable family of solutions within which the reformulated problem has a unique solution.

It is through a choice of the functional P that one incorporates prior information, constraints, and expectations of the form of the solution. The constant λ weights the relative importance of these effects and thus reflects our confidence in the data in solely determining the solution. An extension of standard regularization theory that views the functionals as likelihood functionals [1] often leads naturally to a Markov Random Field model that is highly suggestive of algorithms and implementations for this problem solution. This will be outlined in the next section.

3 Neural Nets, Statistical Physics and Optimization

In this section we extend the standard regularization formalism to the probabilistic regularization approach. In so doing, not only are we provided with a more flexible method of problem formulation, but we establish a link between this level and the levels of algorithm and implementation through the links with the theories of statistical physics and neural networks.

3.1 Statistical Regularization

We will now concentrate on problems in image restoration and in vision but we emphasise that the methods have more general applications, particularly in the broader field of signal processing [21][22]. The essential common element is that such problems are inverse ill-posed problems.

When a problem is ill-posed a statistical reformulation of the problem (e.g. maximum a posteriori, or MAP, estimate) may lead to a well-posed one [40]. A commonly proposed stochastic reformulation uses a markov random field model. These models have found much use in the theory of statistical mechanics and thus also in the simulation of neural networks.

Problems in image restoration are clearly inverse problems since one is trying to recover the original signal from the degraded one. Many of the problems in vision are also inverse ones in that we are given observations (the image data),

[1] Indeed, classical spline fitting can be viewed in a probabilistic bayesian manner (see for e.g. Carmody [8]); thus the early work of Terzopoulos in interpolating sparse stereo depth maps [39] can be viewed as regularization of the problem.

Section 2 introduces regularisation theory and its connection with incorporating knowledge. Then in section 3, the stochastic extension of standard regularization theory is `shown to be related to statistical physics and neural nets. For concreteness, we illustrate this with a simple image restoration example. Section 4 presents some of our work in investigating these approaches in image restoration and computer vision. The last section discusses some of the problems still to be solved, and some of the limitations our work has shown.

The framework presented, naturally links the mathematics of inversion with the physics of large scale systems and the emerging neural network theory.

2 Inverse Problems: regularization using prior information.

In this section we outline how regularization theory provides a framework for formulation of the problems of inverse problems in image processing and vision. The next section extends this standard theory to the more recently proposed probabilistic regularization theory; and thus demonstrates how this problem formulation suggests algorithms and implementations through the connection with theories of statistical physics and neural networks.

The most general formulation of inverse problems is that one is given the sensed data g, which is produced through the action of an operator A, acting upon the data we wish to recover f:

$$Af = g.$$

In image processing f is a corrupted version of g and A is the imaging process. In the vision case the sensed data is the raw pixel intensities, the data of interest f describes the objects of interest in the scene (e.g. locus of object surfaces'), and the operator A is typically a composition of operators describing the image forming process. A problem involving inversion of this operator is well-posed [12] [3] if we can ensure existence, uniqueness, and stability of solutions. For a variety of reasons, failure of one or more of these conditions is common and thus the problem is ill-posed. For example A may not be full rank (so that the solution is not unique - extra information may be required to restrict the solution space), or A may be invertible but ill-conditioned (thus small changes in data lead to large deviations in solution - which can be disastrous in the presence of noise), or A may be of rank greater than the number of degrees of freedom (the system is overdetermined and thus a solution may not exist if any of the measurements contain noise). Typically, the imaging process does not preserve all information (it is a 2-D projection of a 3-D world and some objects are partially occluded) and thus the problem is ill-posed in the first sense given. Additional constraints and assumptions restrict the solution space, but in the presence of noise, the problem can then become ill-posed in either of the last two senses.

Standard regularization theory provides mathematical tools that enable one to turn an ill-posed problem into a well posed problem. One method that has been applied often in early vision **[4]** [5] is to replace the above problem with one of finding the solution that minimizes:

$$|Af - g| + \lambda|Pf|$$

where the first term measures fidelity with data and the second term involves an operator that measures suitability of the solution according to additional constraints (e.g. smoothness of solution). The last term restricts the solution to an admissable family of solutions within which the reformulated problem has a unique solution.

It is through a choice of the functional P that one incorporates prior information, constraints, and expectations of the form of the solution. The constant A weights the relative importance of these effects and thus reflects our confidence in the data in solely determining the solution. An extension of standard regularization theory that views the functionals as likelihood functionals [1] often leads naturally to a Markov Random Field model that is highly suggestive of algorithms and implementations for this problem solution. This will be outlined in the next section.

3 Neural Nets, Statistical Physics and Optimization

In this section we extend the standard regularization formalism to the probabilistic regularization approach. In so doing, not only are we provided with a more flexible method of problem formulation, but we establish a link between this level and the levels of algorithm and implementation through the links with the theories of statistical physics and neural networks.

3.1 Statistical Regularization

We will now concentrate on problems in image restoration and in vision but we emphasise that the methods have more general applications, particularly in the broader field of signal processing [21][22]. The essential common element is that such problems are inverse ill-posed problems.

When a problem is ill-posed a statistical reformulation of the problem (e.g. maximum a posteriori, or MAP, estimate) may lead to a well-posed one [40]. A commonly proposed stochastic reformulation uses a markov random field model. These models have found much use in the theory of statistical mechanics and thus also in the simulation of neural networks.

Problems in image restoration are clearly inverse problems since one is trying to recover the original signal from the degraded one. Many of the problems in vision are also inverse ones in that we are given observations (the image data),

[1]Indeed, classical spline fitting can be viewed in a probabilistic bayesian manner *(see* for e.g. Carmody [8]); thus the early work of Terzopoulos in interpolating sparse stereo depth maps [39] can be viewed as regularization of the problem.

have some knowledge of the transformations from object surface to data (such as reflectance models and camera models); and we wish to recover the surfaces that produced this data. A closely related set of problems in image processing involve the recovery of signals degraded by such processes as optical blur and sensor noise.

A unique solution to these problems does not generally exist due to the paucity of data and due to noise. One then requires a regularised solution, or a maximum likelihood solution. Using Bayes theorem we have that the probability of the solution $F = f$ given the (corrupted) data $G = g$ *is:*

$$P(F|G) = \frac{P(G|F)P(F)}{P(G)} .$$

Thus, maximising the posterior *P (FIG)* is equivalent to maximising the logarithm of the right hand side:

$$lnP(G|F) + lnP(F) - lnP(G).$$

Now the second term is a probability measure on the class of solutions and the traditional methods of choosing splines of a certain degree of smoothness can be seen as attempts at specifying this probability distribution. Such prior models usually invoke a smoothness assumption. A typical such choice is to require that the solution minimize $\sum_{ij}(f_i - f_j)^2$ for i, j neighbors, which is a discrete version of the minimum integral of the first derivative squared. This is equivalent to choosing the distribution:

$$P(f) = \frac{1}{Z}\exp(-\textstyle\sum_{i,j}(f_i - f_j)^2$$

In the language of statistical physics Z is a normalisation constant called the partition function, and the terms in the summation are clique potentials. The assumption of Gaussian noise, of standard deviation σ, allows us to derive:

$$lnP(G|F) = \frac{-1}{2\sigma^2}\sum_i (f_i - g_i)^2).$$

Since *lnP(G)* is assumed constant, the MAP estimate can be found by minimizing

$$lnP(G|F) = \frac{-1}{2\sigma^2}\sum_i (f_i - g_i)^2).$$

A comparison of this with section 2 shows that we have a probabilistic interpretation of the regularization terms with $\lambda = \sigma^2$.The terms involve only nearest neighbour information and thus this formulation is highly suggestive of the various relaxation algorithms based upon markov random field simulations such as the metropolis, the heat bath, or the gibbs sampler method [24]. Such algorithms map naturally onto implementations involving highly parallel simple processors such as the connection machine or neural network type architectures such as the boltzmann machine. These algorithms and some implementations will be discussed in the next section.

3.2 Neural Nets

The relationship between neural net models and the physics of large scale magnetic spin systems has become widely appreciated since the work of Hopfield [16]. The magnetic interaction between spins i and j, given by T_{ij}, is the analog of the synaptic coupling. Thus the Hamiltonian governing the equivalent magnetic spin system, provides an energy interpretation to the dynamics of neural nets. Such an analogy was exploited originally only for neural net models of associative memory. However, since the system evolves in a manner that finds local minima of the overall energy surface, it was soon realised that such systems are capable of performing optimization tasks. By now the relationship between statistical mechanics, neural networks, and optimization, is a major area of research. A more complete introduction can be found in any of the recent books [17][18] [19] [20].

Thus, when we find that a problem is naturally formulated in terms of the minimization of a cost function involving many variables, we can exploit this analogy to map the problem onto a statistical physics system. This allows us to take advantage of the sophisticated methods developed in statistical physics for either reasoning about the behaviour analytically, or for performing Monte Carlo simulations. Moreover, we can then exploit the relationship between the simulated system and models of neural networks in order to propose viable hardware implement at ions.

4 Image Restoration - Computer Vision, as Optimization Problems

In this section we outline previous work in casting problems in machine vision and image processing as stochastic regularisation problems. Then we present our own work in image restoration and stereo surface reconstruction.

A common problem in image processing is the restoration of a picture degraded by noise. For example fig. 1. shows a one dimensional profile of an ideal step edge and the same edge corrupted by noise. The noise elements are seen to be those that are different from the average signal level. Thus the standard techniques of image processing involve a type of spatial averaging (or equivalently, suppression of high frequencies in the fourier transform domain); this tends to reduce the noise but also blurs the edge. Similar problems occur when attempting to segment the image, the edges of regions correspond to large changes in intensity and are thus found by locating areas of high spatial derivatives. However, the noise also produces large spurious responses to the derivative edge detectors. A smoothing step can be carried out, before the edge detection, in order to remove noise, but this also reduces the delectability of true edges. Humans have little difficulty in separating those regions of large intensity change due to isolated and random noise from those brought about by discontinuities of an underlying (otherwise rather smooth) signal. What is required is a characterization of our prior knowledge of the smoothness of the underlying signal and the tendency of edges to form extended contours in two dimensions. This prior information can be included naturally by using a regularization formulation.

Geman and Geman [23] were the first to formulate the general image restoration problem in this manner and solve it by statistical physics methods. They used the more simple smoothness functional of Ising potentials and also introduced a coupled line field that allowed the smoothness potential to be turned off at estimated discontinuities. We present results of simulations of this process later but first show how this has been adapted to solve problems in computer vision.

Poggio et al [5] were the first to point out that many additional constraints imposed upon solutions to problems in vision could be viewed as regularization procedures. In particular, the smoothness constraints imposed on solutions to optic flow, shape from shading, and stereo reconstruction are seen as regularization functionals. The smoothing by gaussian masks before edge detection can also be viewed as an attempt at regularizing an ill-posed problem [6]. It is interesting that this approach was first suggested purely on neurophysiological basis in the DOG filters of Marr [2].

It is know becoming widely acknowledged that many of the problems of machine vision; such as recovery of optical flow, shape from shading and stereo surface reconstruction, can be formulated in this systematic manner [25] [26] [27] [28]. Moreover, this approach promises to provide us with a systematic approach to integrating the various visual modules by coupling together several random fields (statistical systems) [29]. Space permits us to consider only the reconstruction of surfaces from a sparse depth map (such as could be obtained by stereo or by laser range finding).

Hutchinson and Koch [30] have proposed a hybrid analog and digital network to minimize the following surface functional:

$$E(f,l) = \sum_{i,j}(f_i - f_j)^2(1 - l_{ij}) + \frac{1}{\lambda}\sum_i (f_i - g_i)^2 + \sum_i V^C(l)$$

where the terms are a measure of smoothness, closeness of fit, and number of discontinuities respectively. This can be seen as being of the form outlined above except that there are terms associated with a coupled line process field. The term $(1 - l_{i,})$ becomes zero, allowing a discontinuity between f_i and f_j if there is a discontinuity line between them ($l_{ij} = 1$). Such a possibility can be estimated by a separate module (such as an intensity discontinuity detector). The last term is one imposing a probability measure on the configuration of lines introduced by imposing a clique potential $V^C(l)$.

Taking partial derivatives with respect to f_i leads to a discrete approximation of the laplacian operator from the first term. Minimization with respect to this term is thus a membrane fitting process [31]. Their hybrid machine has two cycles: minimization of the functional at fixed line process configuration, followed by a cycle where the effect of flipping each line process is evaluated. They note that such a scheme is only guaranteed to find a local rather than global minimum. A similar hybrid machine is proposed for the solution of the optical flow field [32][34].

4.1 Image Restoration

Our experiments in image restoration aim to investigate the quality of the various proposed stochastic algorithms and their sensitivity to choice of parameters. For comparison purposes we use an image similar to that in the original work of Geman and Geman [23].

The original formulation of Geman and Geman [23] used simulated annealing [20] to find an approximation to the MAP estimate. Others have suggested collecting statistics at a fixed temperature [25] or a mean field approximation that is deterministic [38]. In these methods one has to pick parameters associated with either the annealing schedule or with the relative weighting of terms in the cost function. We have space here to show only a few representative samples of our simulations for different choices of parameters and iteration method.

Fig. 2(a) shows an image of rectangles degraded by noise. This is similar to one of the test patterns used by Geman and Geman [23]. An acceptable image is obtained without the lineprocess field (fig. 2(b)), but is improved by the addition of the lineprocess (fig. 2(c)). Fig. 2(d) shows, however, the effect of poor choice of σ. The next figure displays the results of a Marroquin Mean Field Algorithm: figs. 3 (b), 3 (c) and (d) show results for the deterministic scheme (2,3 and 10 iterations respectively). Finally, note that the edge field not only improves restoration but allows us to estimate the location of edges in noisy images. The combination of surface regularization and discontinuity detection has been suggested as the most reliable method for edge detection in machine vision [31]. However, our preliminary results confirm the need for accurate methods of determining parameters as well as the need for these parameters to be spatially varying [29].

4.2 Stereo Surface Reconstruction

A common approach has been to formulate the solution to a problem in vision as a solution to a variational formulation (extrema of a functional). The variational formulation leads to Euler-Lagrange equations that can be discretized to yield a set of equations that are usually easy to solve in an iterative manner. Sometimes, however, the iterative schemes so derived may have poor convergence properties; whereas another formulation may lead more naturally to iterative schemes with better convergence behaviour. An analysis of shape from shading along these lines may be found in Horn and Brooks [1].

An alternative approach is to avoid solving by the Euler-Lagrange methods, and perform direct gradient descent on the functional. In particular, stochastic highly parallel methods based upon Markov Random Fields (MRF) offer general purpose solution methods that are efficient on highly parallel fine grained architectures. Moreover, the stochastic element of these methods avoids the solution being confined to merely local minima of the functional. Such an approach emphasises the close relationship between these problems in vision and recent theoretical advances in optimization and neural networks inspired by advances in statistical physics.

In the stereo surface reconstruction problem, one is given two images of a scene where corresponding points in the two images are displaced by an amount that depends upon the distance from the baseline between the two image planes and the surface upon which physical points lie. This displacement or disparity, determines the actual distance through a simple triangulation calculation. However, before this calculation can take place one has to find which features correspond in the two images. The features are generally simple (local brightness or edges) and numerous (otherwise the depth map obtained would be sparse); thus one has many potential matches. Selecting the best match is often formulated as an optimization problem, one tries to quantify the quality of the match and tries to restrict the number of potential matches by assuming depth generally varies smoothly except at isolated discontinuities. This introduces two terms into the cost function. We illustrate this with the results of simulations involving random dot stereograms (see figure 4a). Again we map our data compatibility and smoothness terms, of our regularised formulation, onto magnetic interactions of a statistical physics system. This gives us a natural neural net like implementation.

The Hamiltonian for the spin glass equivalent of the neural net is similar to that given in [37] and [36]. We consider that the system is a Q layer network of $N \times N$ spins. The spin at layer r is 1 if the disparity at that site is r (0 otherwise). In a direct simulation the proportion of time that a spin takes on the value 1 is a measure of the probability that that site has such a disparity. We have used both the Ising potentials and the minimum squared first derivative (Section 2) for ensuring smoothness of the results. The closeness of fit term is based upon a 3x3 correlation.

An annealing schedule can be used to find an optimal or most probable configuration [20]. However, instead of simulating directly the underlying spin system we have adopted the methods of approximation of Marroquin [38]. Here, the evolution of the network is deterministic and within 20 iterations one typically converges upon a "mean field" approximation (see fig. 4b).

The two random dot pictures can also be regarded as a pair from a temporal sequence rather than a spatial sequence; the stereo pair becomes a motion sequence. Thus similar techniques can be used to segment moving textured surfaces from a textured background. We can see the power of the human visual system in this respect when we observe a camouflaged animal blend into its background when it stops moving but is detectable when it is moving. Thus our visual system is capable of performing such matching functions effectively in a very short space of time; it is possible that algorithms similar to those discussed here are used.

5 Future Research

There are a number of "free" parameters that must be chosen in the simulations (those imposing the relative weighting of cost function terms and those controlling the course of the iteration updates). One of the aims of the simulations we have performed is to gain a better understanding of the influence of each parameter and of the range of acceptable values. For example, the relative weighting of continuity and correlation terms greatly influences the number and size of disparity levels in the solution to the stereogram when noise is present: the higher the continuity weighting the fewer are the levels resulting (since those having a significant majority coerce neighbors to be similarly aligned). This leads to a cleaner image but may eliminate small regions of true disparity. Another aim of the simulations is to compare the quality of the results with those using annealing schedules or different updating strategies.

An alternative to parameter choice through simulation is parameter estimation by either learning from examples or through theoretical derivations. Others have suggested some approaches to this **[14] [10] [13].** We intend to investigate these and similar schemes.

Furthermore, more elaborate models can be made. For example, a hierarchical model can speed convergence of the solution in a manner analogous to the multigrid methods of Terzopoulos [39]. In such a scheme sparse points (but now allowing the correlations to be based upon large neighborhood averages) can be used to identify the major disparity levels and to obtain a coarse approximation to the solution, with lower levels providing more detail. Finally, a more complex model involving coupled nets can simulate the co-operative effect between different visual modules in obtaining a more accurate and robust depth perception [29]. We thus intend simulating complex coupled systems of the type outlined in Section 4.

In the long term we hope to devise architectures that are capable of performing successful variants of these approaches in real time. The highly parallel nature of the simple calculations involved map naturally onto a massively parallel machine such as the connection machine. However, a cheaper alternative is to geometrically partition the problem onto an array of transputers [9] [35]. It is also possible to design specific VLSI architectures for both standard and probabilistic regularization algorithms [11].

6 Conclusion

Th concepts of regularization theory have given a comprehensive framework to formulation of the problems in vision: at all three levels of problem, algorithm, and implementation. Furthermore, the mathematical theory of regularization provides a useful theory for incorporating prior knowledge, constraints, and expectations of solution. The exciting prospect is that the analogies with statistical physics and neural networks will continue to provide fruitful ideas for development of machine intelligence and for understanding human intelligence.

References

[1] B. K. P. Horn and M. J. Brooks "The Variational Approach to Shape from Shading," *Computer Vision, Graphics, and Image Processing*, Vol. 33, pp. 174-208, 1986.

[2] D. Marr *Vision,* Freeman, San Francisco, 1982.

[3] V. A. Mozorov "Methods for Solving Incorrectly Posed Problems," Springer-Verlag, New York, 1984.

[4] D. Terzopoulos "Regularization of Inverse Visual Problems Involving Discontinuities," *IEEE* Trans. *Pattern Analysis and Machine Intelligence,* Vol. 8 No. 4, pp. 413-424, June 1986.

[5] T. Poggio and C. Koch "Ill-posed problems in early vision: from computational theory to analogue networks," *Proc. R. Soc. Land.,* Vol. B 226, pp. 303-323, 1985

[6] M. Bertero, P. Brianzi, E. R. Pike, and L. Rebolia "Linear regularizing algorithms for positive solutions of linear inverse problems," *Proc. R. Soc. Land.,* A, Vol. 425, pp. 257-275, 1988.

[7] A. N. Tikhonov, A. V. Goncharskii, V. V. Stepanov, and I. V. Kochikov, "Ill-posed image processing problems," *Soy. Phys. Dokl.,* Vol. 32 No. 6, pp. 456-458, June 1987.

[8] T. J. Carmody "Diagnostics for Multivariate Smoothing Splines," *J. of Statistical Planning and Inference,* Vol. 19, pp. 171-186, 1988.

[9] D. Suter and X. Deng "Neural Net Simulation on Transputers," *Proc. IEEE Systems, Man, and Cybernetics Conf.,* Beijing, Aug. 1988, (to appear).

[10] D. Shulman and J. Aloimonos "Boundary Preserving Regularization: Theory Part 1," University of Maryland, Computer Vision Lab., Tech Report CS-TR-2011, April 1988.

[11] S. Y. Kung "VLSI Array Processors," Prentice-Hall, London, 1988.

[12] A. N. Tikhonov and A. V. Goncharsky (Eds.) "Ill-Posed Problems in the Natural Sciences," MIR Publishers, Moscow, 1987.

[13] D. Shulman and J. Y. Aloimonos " (Non)-rigid motion interpretation: a regularized approach," *Proc. R. Soc. Lond. B,* pp. 217-234, 1988.

[14] S. P. Luttrell "The use of Markov random field models to derive sampling schemes for inverse texture problems," Inverse *Problems,* Vol. 3, pp. 289-300, 1987.

[15] S. P. Luttrell and C. J. Oliver "The role of prior knowledge in coherent image processing," *Phil. Trans. R. Soc. Lond. A,* Vol. 324, pp. 297-298, 1988.

[16] J. J. Hopfield, "Neural Networks and Physical Systems with Emergent Collective Computational Abilities," *Proc. Natl. Acad. Sci. USA,* Vol. 79, pp. 2554-2558, 1982.

[17] M. Mezard, G. Parisi, and M. Virasoro, "Spin Glass Theory and Beyond," World Scientific, Singapore, 1987.

[18] D. Chowdhury, "Spin Glasses and Other Frustrated Systems," World Scientific, Singapore, 1987.

[19] J. L. van Hernmen and I. Morgenstern (Ed.), "Heidelberg Colloquim on Glassy Dynamics," Lecture Notes in Physics 275, Springer Verlag, Berlin, 1987.

[20] P. J. M. van Laarhoven and E. H. L. Aarts, "Simulated Annealing: Theory and Applications," Reidel, Dordrecht, 1987.

[21] W. Jeffrey and R. Rosner, "Neural Network Processing as a Tool for Function Optimization," *AIP Conference Proc. 151 Neural Networks for Computing,* Snowbird, Utah U.S.A., pp. 241-246, 1986.

[22] W. Jeffrey and R. Rosner, "Optimization Algorithms: Simulated Annealing and Neural Network Processing," *The Astrophysical Journal,* Vol. 310, pp. 473-481, Nov. 1986.

[23] S. Geman and D. Geman "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," *IEEE PAMI,* Vol. 6, pp. 721-741, 1984.

[24] J. L. Marroquin "Probabilistic Solution of Inverse Problems," PhD thesis, MIT, Cambridge MA, Sept. 1985.

[25] J. Marroquin, S. Mitter, and T. Poggio, "Probabilistic solution of ill-posed problems in computational vision," *J. Am. Stat. Assoc.,* Vol. 82, pp. 76-89, 1987

[26] D. Geman, "Stochastic model for boundary detection," *Image and Vision Computing,* Vol. 5 No. 2, pp. 61-66, May 1987.

[27] D. W. Murray, A. Kashko and H. Buxton, "A parallel approach to the picture restoration algorithm of Geman and Geman on an SIMD machine," *Image and Vision Computing,* Vol. 4 No. 3, pp. 133-142, Aug. 1986.

[28] B. F. Buxton and D. W. Murray, "Optic flow segmentation as an ill-posed and maximum likelihood problem," *Image and Vision Computing,* Vol. 3 No. 4, pp. 163-169, Nov. 1985.

[29] E. Gamble and T. Poggio, "Visual Integration and Detection of Discontinuities: The Key Role of Intensity Edges," "MIT **A.I.** Memo No. 970, Oct. 1987.

[30] J. M. Hutchinson and C. Koch, "Simple Analog and Hybrid Networks for Surface Interpolation," *AIP Conference Proc. 151 Neural Networks for Computing,* Snowbird, Utah U.S.A., pp. 235-240, 1986.

[31] A. Blake and A. Zisserman, "Visual Reconstruction," MIT Press, Cambridge Mass., 1987.

[32] J. M. Hutchinson, C. Koch, J. Luo and C. Mead, "Computing Motion using Analog and Binary Resistive Networks," *IEEE Computer*, pp. 52-63, March 1988

[33] D. M. Titterington "General structure of regularization procedures in image reconstruction," *Astron. Astrophys.*, Vol. 144, pp. 381-381, 1985.

[34] N. M. Grzyskracz and A. Yuille, "Motion Correspondence and Analog Networks," *AIP Conference Proc. 151 Neural Networks for Computing,* Snowbird, Utah U.S.A., pp. 200-205, 1986.

[35] B. M. Forrest, D. Roweth, N. Stroud, D.J. Wallace and G.V. Wilson, "Implementing Neural Network Models on Parallel Computers," *The Computer Journal,* Vol. 30 No. 5, pp. 413-419, 1987.

[36] E. Goles and G. Y. Vichniac, "Lyapunov Function for Parallel Neural Networks," *AIP Conference Proc. 151* Neural *Networks for Computing,* Snowbird, Utah U.S.A., pp. 129-134, 1986.

[37] R. Divko and K. Schulten, "Stochastic Spin Models for Pattern Recognition," *AIP Conference Proc. 151 Neural Networks for Computing,* Snowbird, Utah U.S.A., pp. 129-134, 1986.

[38] J. L. Marroquin, "Deterministic Bayesian Estimation of Markov Random Fields with Applications in Computational Vision," *Proc. 1st Int. Conf. on Computer Vision,* pp. 597-601, June 8-11, 1987.

[39] D. Terzopoulos, "Multi-Level Reconstruction of Visual Surfaces," MIT A.L memo 671, April 1982.

[40] J. Franklin "Well-Posed Stochastic Extensions of Ill-Posed Linear Problems,' *J. of Mathematical Analysis and Applications,* Vol. 31, pp. 682-716, 1970.

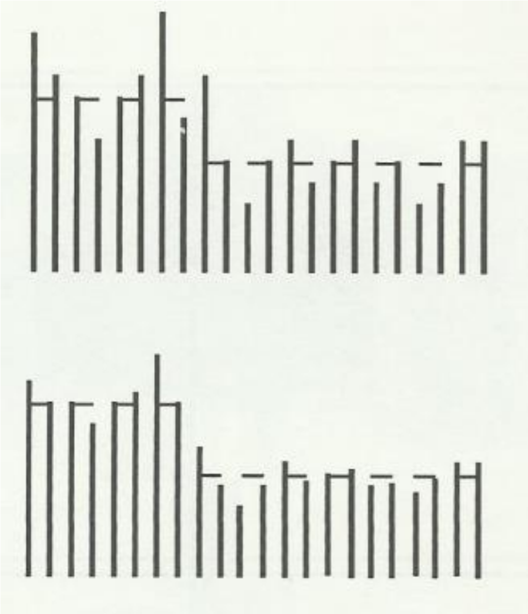

Fig. 1. Noise reduction by filtering.
(a) - Top.
A step edge (dashed line) corrupted by noise.
(b) - Bottom.
The step edge filtered with a running mean filter. The noise is reduced but the edge is smoothed. This not only produces blurred pictures in image restoration, but frustrates image segmentation in computer vision. Non-linear filters, such as running median, often perform better but are still unsuitable for segmentation purposes. Regularization approaches can provide approaches that smooth noise while preserving and locating edges.

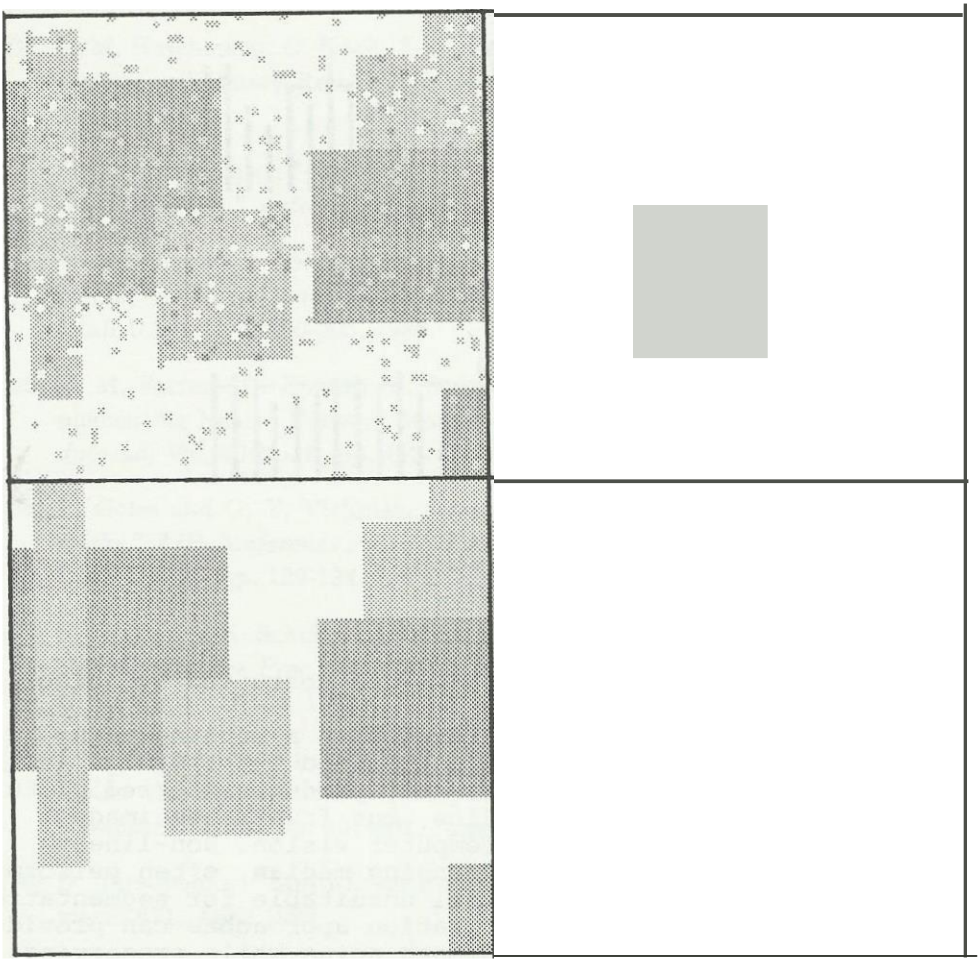

Fig. 2. Image Restoration Using Simulated Annealing.
(a) - Top left.
Original image corrupted with noise.
(b) - Top right.
Image after restoration using simulated annealing but with no lineprocesses.
(c) - Bottom left.
Image after annealing with lineprocesses.
(d) - Bottom right.
Image restoration with algorithm using incorrect estimation of noise level.

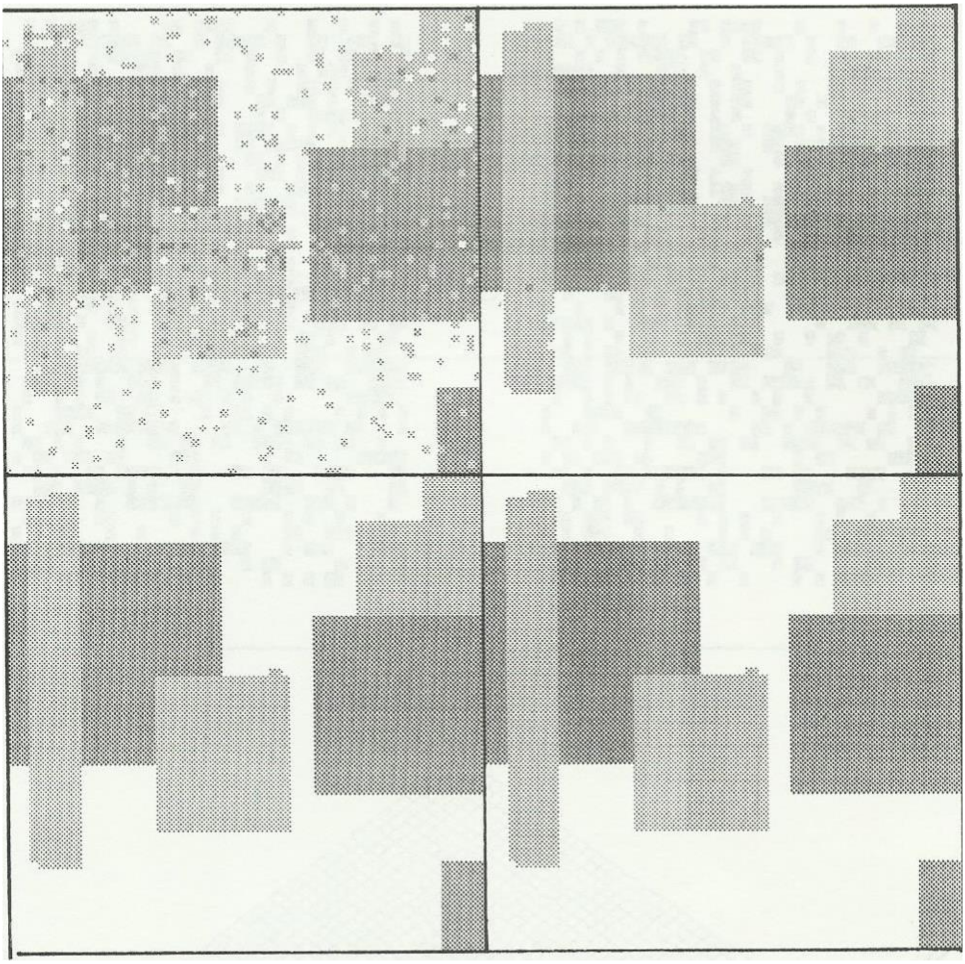

Fig. 3. Image Restoration Using Maximum Posterior Marginals.
(a) - Top left.
Original image corrupted with noise.
(b) - Top right.
Image after restoration using 2 iterations of the Mean Field algorithm.
(c) - Bottom left.
Image after restoration using 3 iterations of the Mean Field algorithm.
(d) - Bottom right.
Image after restoration using 10 iterations of the Mean Field algorithm.

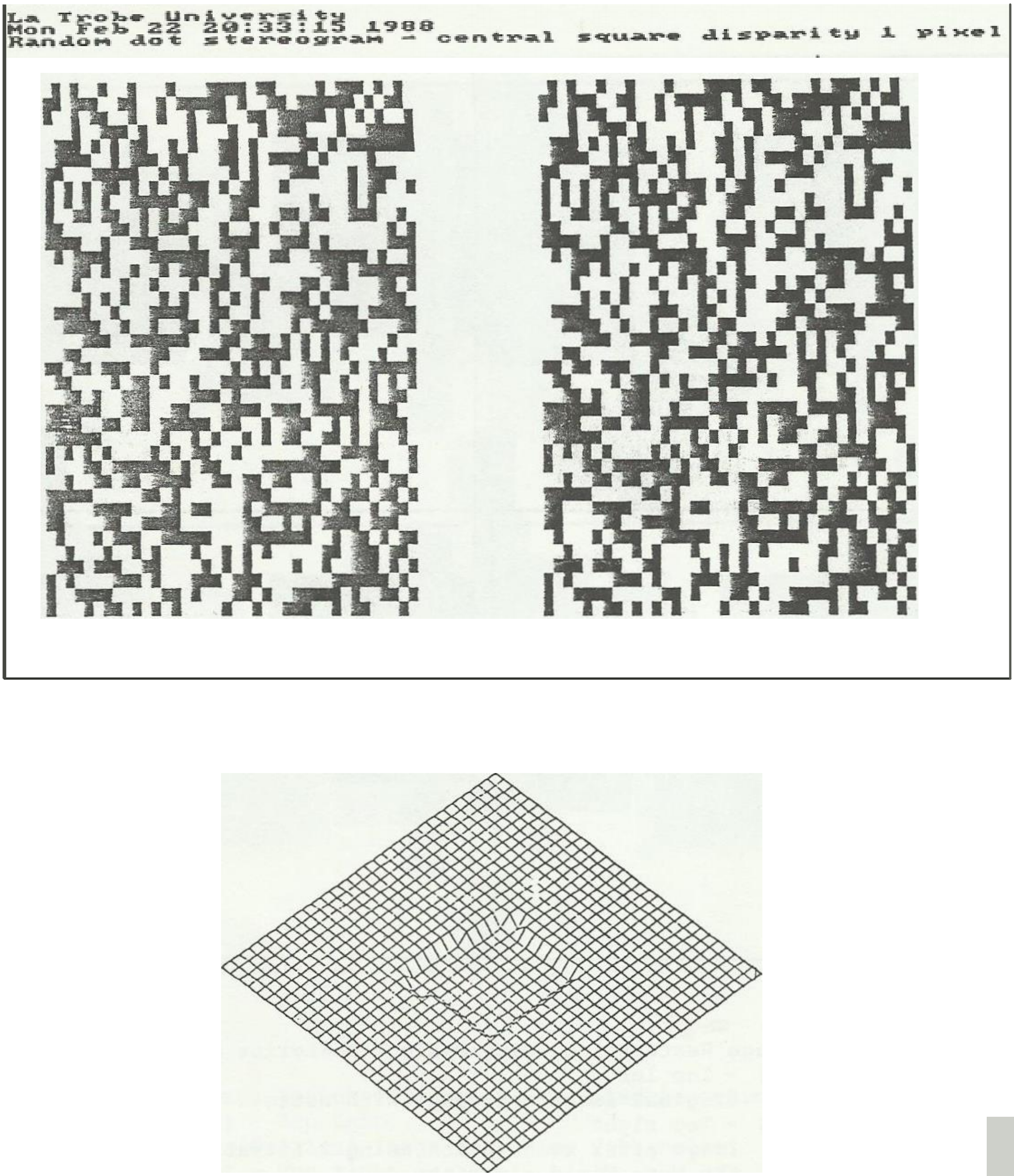

Fig. 4. The Mean Field Algorithm used to solve the stereo correspondance problem.

(a) - Top.
Two random dot input images, a small square in the centre is displaced in one image relative to the other.

(b) - Bottom.
Graph of the disparity solution discovered by the algorithm.