Image processing by simulated annealing

by P. Carnevali
L. Coletti
S. Patarnello

It is shown that simulated annealing, a statistical mechanics method recently proposed as a tool in solving complex optimization problems, can be used in problems arising in image processing. The problems examined are the estimation of the parameters necessary to describe a geometrical pattern corrupted by noise, the smoothing of bi-level images, and the process of halftoning a continuous-level image. The analogy between the system to be optimized and an equivalent physical system, whose ground state is sought, is put forward by showing that some of these problems are formally equivalent to ground state problems for two-dimensional Ising spin systems. In the case of low signal-to-noise ratios (particularly in image smoothing), the methods proposed here give better results than those obtained with standard techniques.

Introduction

Simulated annealing is a technique recently introduced [1, 2] to solve very complex optimization problems. A quick description of this method is given here to make this paper self-contained, but the reader is referred to [1] for a complete discussion.

Consider the problem of minimizing the function $E(x_i)$ of the many variables x_i, i.e., of looking for the values of x_i that yield the absolute minimum of the function $E(x_i)$. The basic idea of simulating annealing consists of treating the system to be optimized as a physical system described by the degrees of freedom x_i, with the energy given by $E = E(x_i)$. One then looks for the state of minimum energy of the physical system, i.e., what physicists call the *ground state*.

With simulated annealing, the ground state is reached by simulating a slow cooling of the physical system, starting from a very high temperature T down to $T = 0$. The cooling must be slow enough that the system does not get stuck into thermodynamically metastable states that are *local* minima of $E(x_i)$. This slow cooling process (called annealing from the analogy with metallurgic processes) is simulated using a standard method proposed by Metropolis et al. [3].

For a given temperature, the Metropolis method is a way to sample states of the physical system with the Boltzmann distribution

$$f = e^{-\frac{E}{T}}, \tag{1}$$

which is the distribution that properly describes the state of thermodynamical equilibrium for a given temperature T. One starts with a random configuration x_i. One then chooses (again, randomly) a small perturbation Δx_i in the system and calculates the energy change ΔE caused by the perturbation

$$\Delta E = E(x_i + \Delta x_i) - E(x_i). \tag{2}$$

If $\Delta E < 0$, then the perturbation is "accepted," for it means that it is energetically favorable for the system; otherwise, it is accepted with probability $e^{-\Delta E/T}$. When the perturbation is accepted, one continues the process with the perturbed state $x_i + \Delta x_i$ replacing the old one; otherwise a new perturbation Δx_i is attempted. It can be shown that the sequence of states obtained in this way is distributed according to (1). The Metropolis method is widely used in physics to study numerically the thermodynamical properties of large systems that cannot be treated with analytical methods.

In simulated annealing, one starts with a high value of T, so that the probability of the system being in a given state is independent of the energy of that state. One then slowly reduces T, by making sure that at each new value of T enough steps of the Metropolis procedure are made to guarantee that thermodynamical equilibrium has been reached. One continues the procedure until $T = 0$. If the cooling has been slow enough, the final state reached is the ground state of the physical system being considered; i.e., the values of x_i so obtained realize the absolute minimum of the function E. In practice, in many cases one is not really interested in finding the absolute minimum. Rather, in many interesting situations the minimum configuration is highly degenerate. In other words, there are many minima with values of E very close to its absolute minimum value, and one looks for one of the very many of them.

For problems of constrained minimum, the method can still be applied. One simply has to make sure that all the perturbations Δx_i that are generated during the Metropolis procedure continue to satisfy the constraints of the problem. In particular, the constraints could consist of prescribing discrete values for the x_i. Thus, simulated annealing applies as well in problems of discrete optimization (as shown by Kirkpatrick [1] in his various examples).

In this paper, we study the application of simulated annealing to various optimization problems arising in image processing. The method seems to be well suited for problems involving very low signal-to-noise ratios. In addition, we show that the analogy between the system to be optimized and a physical system can be even stronger than that implied by the process of simulated annealing. In particular, we show that various problems involving bi-level images (bitmaps) are formally equivalent to ground state problems for two-dimensional Ising spin systems imbedded in an external field.

In the following section we analyze the problem of the recognition of a regular pattern of rectangles out of an initial configuration corrupted by noise. The next section is devoted to the process of smoothing a bi-level image by means of the annealing procedure. In this context we find that the "cost function" for this problem can be easily related to the energy of a statistical model defined on a lattice (the given bitmap), known as the Ising model. As we have already mentioned, this procedure works remarkably well in the presence of high noise levels. Finally, in the fourth section we generalize this analogy by showing that the problem of halftoning a gray-level image is again equivalent to a spin system in which the internal interaction is of a different nature than the previous one. Correspondingly, the annealing procedure has to be a bit more accurate, since the physical system turns out to have a more complex ground state. We are also able to provide quantitative estimates of the gray and spatial resolutions in terms of the physical quantities involved in the spin system.

Estimation of parameters

The first problem of image processing we want to solve using simulated annealing is that of parameter estimation.

Suppose we have an image represented by the rectangular array of real numbers α_{ij}, with $1 \le i \le N_x$, $1 \le j \le N_y$. In addition, assume that a parametric model for the image is available. In other words, we know that the given image represents a scene containing objects of shapes and sizes that are partially known and that can be precisely described in terms of the parameters $x_l (1 \le l \le N_p)$. This model is assumed to be well known, so that, given any set of values x_l, it is possible to calculate the corresponding image $\alpha_{ij} = \alpha'_{ij}(x_l)$ that would be observed, in the absence of noise, for a scene described by the values x_l for the parameters.

The problem then consists of estimating "good" values x_l that describe the given image. Since noise or other sources of error can be present, it cannot be expected that a precise fit of the observed image can be obtained; i.e., in general one will not be able to find a set of values x_l^* such that

$$\alpha'_{ij}(x_l^*) = \alpha_{ij}\,. \tag{3}$$

Rather, one will introduce a measure E of the difference between the left and the right side of (3) and minimize it.

For example, choosing the L^2 norm to measure the distance between α and α', one has

$$E(x_l) = \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} [\alpha'_{ij}(x_l) - \alpha_{ij}]^2, \tag{4}$$

whereas the expression for E in the case that an L^1 norm is preferred is

$$E(x_l) = \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} |\alpha'_{ij}(x_l) - \alpha_{ij}|\,. \tag{5}$$

Thus the problem of parameter estimation is a minimization problem that can be solved using simulated annealing.

As an example, we chose to apply simulated annealing to the following parameter estimation problem. The given image α_{ij} is assumed to be a bi-level image (bitmap); i.e., it can only have the values 0 and 1, with 0 conventionally representing white (or background) and 1 black (or foreground). The image is a square one, with $N_x = N_y = N$. The scene observed is assumed to be a white sheet of paper, with an unknown number of black rectangles on it. The rectangles are at unknown positions on the sheet and their sizes are also unknown. The rectangles are allowed to overlap each other. To simplify the programming, it is assumed that the rectangles have their sides parallel to the sides of the image. In this way, the lth rectangle is a black area composed of the pixels with $x_{\min}^{(l)} \le i \le x_{\max}^{(l)}$ and $y_{\min}^{(l)} \le j \le y_{\max}^{(l)}$ and each rectangle is described by the four parameters $x_{\min}^{(l)}$, $x_{\max}^{(l)}$, $y_{\min}^{(l)}$, and $y_{\max}^{(l)}$. If N_R is the (unknown) number of rectangles, the image α'_{ij} corresponding to a given set of rectangles is given by

$$\alpha'_{ij} = \sum_{l=1}^{N_R} \theta(i - x_{\min}^{(l)})\theta(x_{\max}^{(l)} - i)\theta(j - y_{\min}^{(l)})\theta(y_{\max}^{(l)} - j), \tag{6}$$

where the sum is understood to be a boolean sum (OR) and θ is a step function of integer argument defined to be 0 if the argument is negative and 1 otherwise.

The unknowns of the problem are N_R and the $4N_R$ parameters describing the rectangles. Using (4) or (5) to define E is equivalent, since for these bi-level images we have $(\alpha' - \alpha)^2 = |\alpha' - \alpha|$. Thus one obtains

$$E = \sum_{i=1}^{N} \sum_{j=1}^{N} \left| \alpha_{ij} - \sum_{l=1}^{N_R} \theta(i - x_{\min}^{(l)})\theta(x_{\max}^{(l)} - i) \times \theta(j - y_{\min}^{(l)})\theta(y_{\max}^{(l)} - j) \right| . \quad (7)$$

Note that this expression represents nothing other than the number of pixels on which α and α' are in disagreement.

However, this expression for E is useless because of the variable number of degrees of freedom (which is $4N_R + 1$). By taking a large enough number of rectangles, one can obtain a solution with $E = 0$ for any given α (in fact, one such solution consists of having a 1×1 rectangle for each black pixel of the given image). To overcome this problem, one needs to penalize solutions involving too many rectangles, in such a way that if two solutions give the same value of E using (7), the solution with the smaller N_R is to be preferred. This is easily achieved by adding to the expression for E a term proportional to N_R. The new expression for E then becomes

$$E = \kappa N_R + \sum_{i=1}^{N} \sum_{j=1}^{N} \left| \alpha_{ij} - \sum_{l=1}^{N_R} \theta(i - x_{\min}^{(l)})\theta(x_{\max}^{(l)} - i) \times \theta(j - y_{\min}^{(l)})\theta(y_{\max}^{(l)} - j) \right| . \quad (8)$$

The parameter κ is interpreted as follows: For small values of κ, small rectangles present in the input image α are considered to be "real." When κ is made larger, small rectangles start to be considered as noise because their inclusion in the solution would increase E rather than decrease it. In fact, suppose the inclusion of a rectangle causes an improvement of n pixels in the agreement between α and α'. Then the second term of (8) decreases by n. However, the corresponding change in E is $\kappa - n$, and this is energetically favorable (negative) only if $n > \kappa$. Thus, only rectangles implying an improvement of at least κ pixels in the agreement between α and α' are included in the solution; κ then acts as a control parameter describing, roughly, how much noise one is willing to keep in the solution.

The solution of the problem of the rectangles using simulated annealing has been implemented for a 128×128 bitmap. The input image α_{ij} has been constructed starting from a noiseless configuration including 50 rectangles (shown in **Figure 1**) and adding random noise to it: Each pixel has been changed from 0 to 1 (or vice versa) with probability r, r taking the values 0.2, 0.3, and 0.4. The images obtained in this way are shown in **Figure 2**.

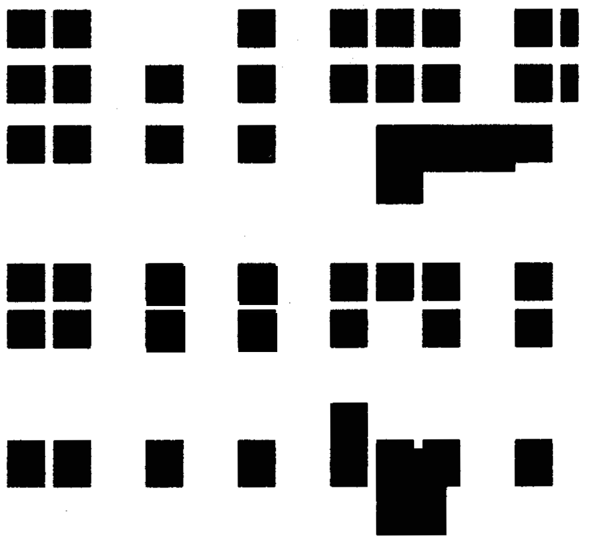

Figure 1

Parameter estimation: noiseless configuration with 50 rectangles.

In each of these images, the average value of α is r in originally white areas, $1 - r$ in originally black areas. The difference between the two average values is $1 - 2r$. The standard deviation of α in both the originally white and originally black areas is given by $\sqrt{r(1 - r)}$. Thus we can define a signal-to-noise ratio as the ratio of the difference between the two average values and the standard deviation, given by $(1 - 2r)/\sqrt{r(1 - r)}$. This gives signal-to-noise ratios of 1.5, 0.87, and 0.41 for the three values of r considered.

The annealing process has been started using a random configuration of rectangles, like the one shown in **Figure 3**. At each step, a change has been picked up randomly among the following classes of possible changes:

- Creation of a new rectangle at a random position and with a random size (smaller than a fixed maximum size).
- Removal of an existing rectangle.
- Stretching of one side of an existing rectangle in one direction by a small, random number of pixels.
- Splitting of an existing rectangle into two rectangles, leaving a one-pixel-wide strip between the two. The splitting can occur at a random point and can be either in the horizontal or in the vertical direction.

Results of the simulated annealing process for the three values of r considered are shown in **Figure 4**. Despite the low signal-to-noise ratios considered, reconstruction of the original image is quite good: Even for the almost hopeless case $r = 0.4$ (corresponding to a signal-to-noise ratio of only

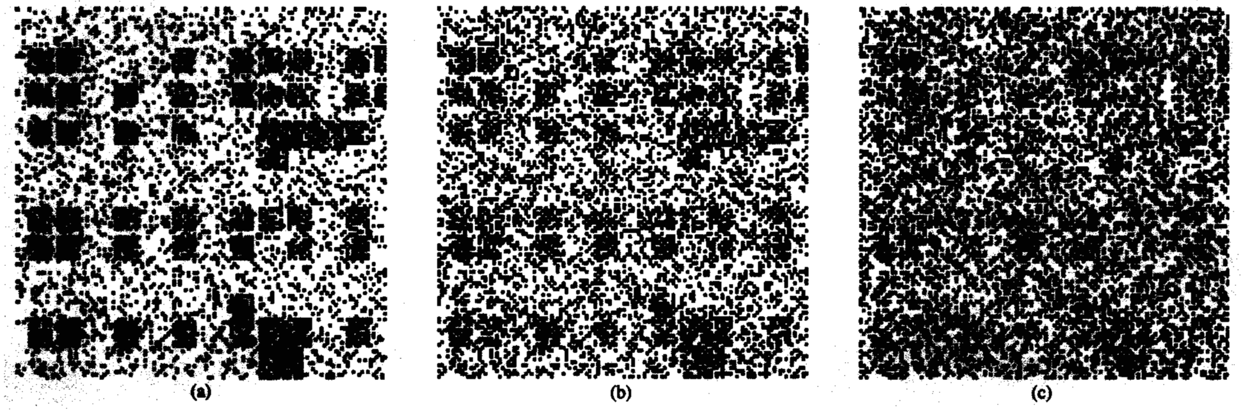

Figure 2

Parameter estimation: noise-corrupted configurations. The image in Fig. 1 has been corrupted by adding random noise as explained in the text. Signal-to-noise ratios are 1.5, 0.87, and 0.41 in (a), (b), and (c), respectively.

0.41), many features of the original image are reconstructed correctly.

Image smoothing and Ising spin systems

The second problem we want to study is that of smoothing images in a sense to be made precise below. We show that a large class of smoothing methods similar to those proposed in [4] are equivalent to ground state problems for two-dimensional Ising spin systems.

For an introduction on Ising spin systems, we refer the reader to textbooks (such as, for example, [5] or [6]). Here we only mention some of the points we need for the following.

Consider a two-dimensional network of points arranged in a square lattice with each point, labeled by its integer coordinates i, j, connected to its four nearest neighbors. Suppose we place in each point a particle with a magnetic moment (spin) and that each particle can be in one of two states, conventionally labeled $\mu = -1$ and $\mu = +1$ or called "spin down" and "spin up," respectively. Suppose that each particle interacts with its four neighbors, and assume that the interaction is translation-invariant and isotropic. Then the energy of the system can be written as

$$E = -\frac{J}{2} \sum_{i_1 j_1 i_2 j_2} C_{i_1 j_1 i_2 j_2} \mu_{i_1 j_1} \mu_{i_2 j_2}, \tag{9}$$

where $C_{i_1 j_1 i_2 j_2}$ is a connection matrix which is 1 if the points i_1, j_1 and i_2, j_2 are nearest neighbors to each other and 0 otherwise. This is a very simplified but physically reasonable model for a two-dimensional substance exhibiting ferromagnetic behavior which was introduced by Ising [7] in 1925. If the system is infinite (or if border effects are neglected), the expression for E can also be written as

$$E = -J \sum_{ij} (\mu_{ij}\mu_{i+1,j} + \mu_{ij}\mu_{i,j+1}), \tag{10}$$

from which it is clear that J (a positive constant) is the contribution to the total energy given by a pair of adjacent spins, the sign of the contribution being negative if the two adjacent spins are aligned (both up or both down) and positive otherwise. The system, in order to minimize its energy, tends to align all its spins in the same direction. Therefore, the ground state configuration is very simple, with all the spins of the lattice pointing in one direction (up or down). Antiferromagnetic models are also possible, in which J is negative, and they are considered in the following section.

An additional feature one can include in Ising models is the presence of an external magnetic field γ_{ij} which tends to align the spins in the direction prescribed by its sign. This introduces an additional term $-\frac{1}{2}\gamma_{ij}\mu_{ij}$ for each spin in the expression for the energy, so that we have

$$E = -\frac{1}{2} \sum_{ij} \gamma_{ij}\mu_{ij} - \frac{J}{2} \sum_{i_1 j_1 i_2 j_2} C_{i_1 j_1 i_2 j_2} \mu_{i_1 j_1} \mu_{i_2 j_2} \tag{11}$$

or

$$E = -\frac{1}{2} \sum_{ij} \gamma_{ij}\mu_{ij} - J \sum_{ij} (\mu_{ij}\mu_{i+1,j} + \mu_{ij}\mu_{i,j+1}). \tag{12}$$

The factor $\frac{1}{2}$ has been introduced simply for convenience and its only effect is to change the units in which the field γ is measured. It is clear that each spin is under the influence of two (possibly competing) forces: one due to the interaction with its neighbors and that tends to align the spin with its neighbors; the second, due to the external field, which tends to align the spin with the external field (in fact, the external field contribution to E is negative, i.e., energetically favorable, for spins which have μ with the same sign as γ). The introduction of such an external field strongly increases the complexity of the model under study; in particular, it may be very difficult to find the ground state if the value of γ_{ij} changes significantly with the position on the lattice.

An additional generalization could consist of introducing an interaction that does not involve only pairs of nearest neighbor points in the lattice. In this case the expression for E would contain a contribution for every pair of points that are close enough to each other. This can be obtained simply by keeping all the above expressions as they are, but allowing the connection matrix $C_{i_1j_1i_2j_2}$ to have values different from 0 (and not constrained to be 1) if the points (i_1, j_1) and (i_2, j_2) are close enough to each other. If one wants to keep an interaction term which is translation invariant, C must be of the form

$$C_{i_1j_1i_2j_2} = C(i_1 - i_2, j_1 - j_2), \tag{13}$$

with the function C being nonzero only for sufficiently low values of its arguments.

Since their introduction in the literature [7], a large amount of work has been done on these models. Onsager [8] found an exact analytical solution for the model described by (10) for all values of the coupling constant J; he also showed the existence of a critical temperature (for which an expression, depending only on J, can be written) at which a phase transition occurs. This has important physical interpretations, and is consistent with the behavior of ferromagnetic systems as a function of the temperature, in particular the existence of the Curie temperature at which ferromagnetism is known to disappear. Unfortunately, efforts to find analytical solutions for the model with an external field have been so far unsuccessful; even for the uniform case (γ_{ij} position independent) only approximate solutions are available.

Figure 3

Parameter estimation: initial random configuration of rectangles.

We now turn to the problem of smoothing an image. We consider a bi-level (0, 1) square image α_{ij} with $1 \le i \le N$ and $1 \le j \le N$. Given α_{ij}, we want to find a smoothed image β_{ij}. Following the approach suggested in [4], β is chosen to be the image that minimizes a two-part cost function. The first part of the cost function, R, measures the "roughness" of the smoothed image β: Ideally the smoothed image should not be "rough" at all. The second part of the cost function, D, is a measure of the discrepancy between the smoothed image β

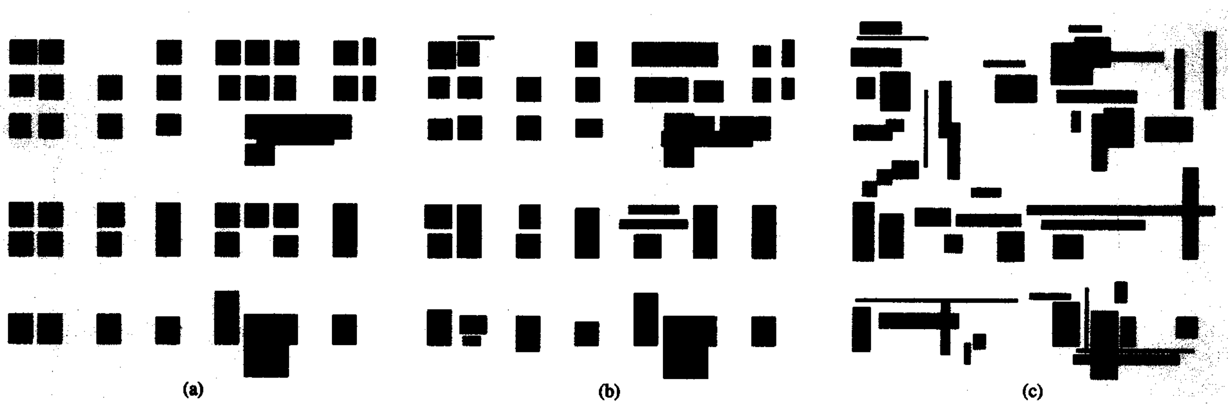

Figure 4

Parameter estimation: results of simulated annealing on images of Fig. 2.

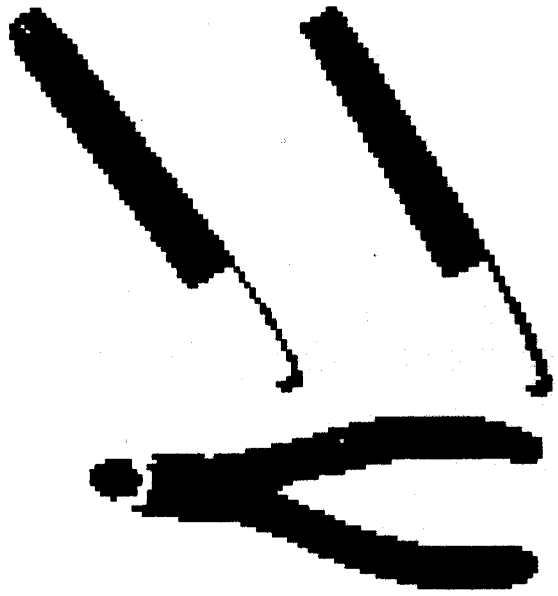

Figure 5

Image smoothing: noiseless bitmap.

and the original image α, and has the effect of preventing the solution β from differing too much from the given image to be smoothed. The cost function to be minimized is then taken to have the form

$$E = R + \lambda D, \tag{14}$$

where λ is a number that parametrizes the desired trade-off between roughness and discrepancy for the smoothed image β.

For D we can take either the L^1 or the L^2 distance between α and β: As we have already mentioned, they are equivalent for these bi-level images since, for any values for α and β in (0, 1), one has trivially $|\alpha - \beta| = (\alpha - \beta)^2$. Thus we can choose

$$D = \sum_{ij} (\alpha_{ij} - \beta_{ij})^2. \tag{15}$$

For R, the two simplest choices proposed in [4] (the digital Laplacian and the digital gradient magnitude) happen to be equivalent. Both choices result in

$$R = \sum_{i_1=1}^{N} \sum_{j_1=1}^{N} \sum_{i_2=1}^{N} \sum_{j_2=1}^{N} C_{i_1j_1i_2j_2}(\beta_{i_1j_1} - \beta_{i_2j_2})^2, \tag{16}$$

where, as in (9), $C_{i_1j_1i_2j_2}$ is a connection matrix. In this way the problem of finding β is well defined, and consists of minimizing E given by (14) for given α_{ij} and λ, with D and R defined as in (15) and (16).

This problem is formally equivalent to finding the ground state of an Ising system imbedded in an external field. To show this, we have to reformulate the problem in terms of a bi-level image with values $(-1, 1)$ instead of (0, 1). Thus we define

$$\gamma_{ij} = 2\alpha_{ij} - 1 \tag{17}$$

and

$$\mu_{ij} = 2\beta_{ij} - 1. \tag{18}$$

Since α and β can only take the values 0 and 1, γ and μ correspondingly take the values -1 and $+1$. We need to express D and R in terms of these new variables.

We first write the following identities, which can be easily verified:

$$\gamma_{ij}^2 = 1; \tag{19}$$

$$\mu_{ij}^2 = 1; \tag{20}$$

$$(\alpha_{ij} - \beta_{ij})^2 = \frac{1 - \gamma_{ij}\mu_{ij}}{2}; \tag{21}$$

$$(\beta_{i_1j_1} - \beta_{i_2j_2})^2 = \frac{1 - \mu_{i_1j_1}\mu_{i_2j_2}}{2}. \tag{22}$$

Using these identities, we get the following new expressions for D and R:

$$D = \frac{1}{2} N^2 - \frac{1}{2} \sum_{ij} \gamma_{ij}\mu_{ij}; \tag{23}$$

$$R = \frac{1}{2} \sum_{i_1j_1i_2j_2} C_{i_1j_1i_2j_2} - \frac{1}{2} \sum_{i_1j_1i_2j_2} C_{i_1j_1i_2j_2}\beta_{i_1j_1}\beta_{i_2j_2}. \tag{24}$$

The first term in R is simply half the number of neighbor pairs and equals $N^2(N^2 - 1)$. Thus,

$$R = N^2(N^2 - 1) - \frac{1}{2} \sum_{i_1j_1i_2j_2} C_{i_1j_1i_2j_2}\mu_{i_1j_1}\mu_{i_2j_2}. \tag{25}$$

Expression (14) now becomes

$$E = -\frac{1}{2} \sum_{i_1j_1i_2j_2} C_{i_1j_1i_2j_2}\mu_{i_1j_1}\mu_{i_2j_2} - \frac{\lambda}{2} \sum_{ij} \gamma_{ij}\mu_{ij}, \tag{26}$$

where we have dropped the constant terms which do not depend on μ. We can do that because addition of a term independent of μ does not change the point where E reaches its absolute minimum. This is also consistent with the well-known fact that the energy of a physical system can be redefined by a constant (or by a quantity independent from the degrees of freedom that describe the system) without changing the physics. Moreover, we can also multiply the energy by a positive constant without altering its physical features; in our case, if we choose the factor $1/\lambda$, we obtain the expression

$$E = -\frac{1}{2} \sum_{ij} \gamma_{ij}\mu_{ij} - \frac{1}{2\lambda} \sum_{i_1j_1i_2j_2} C_{i_1j_1i_2j_2}\mu_{i_1j_1}\mu_{i_2j_2}, \tag{27}$$

which is identical to (11) provided that $J = 1/\lambda$. Thus we have shown the complete equivalence of the smoothing

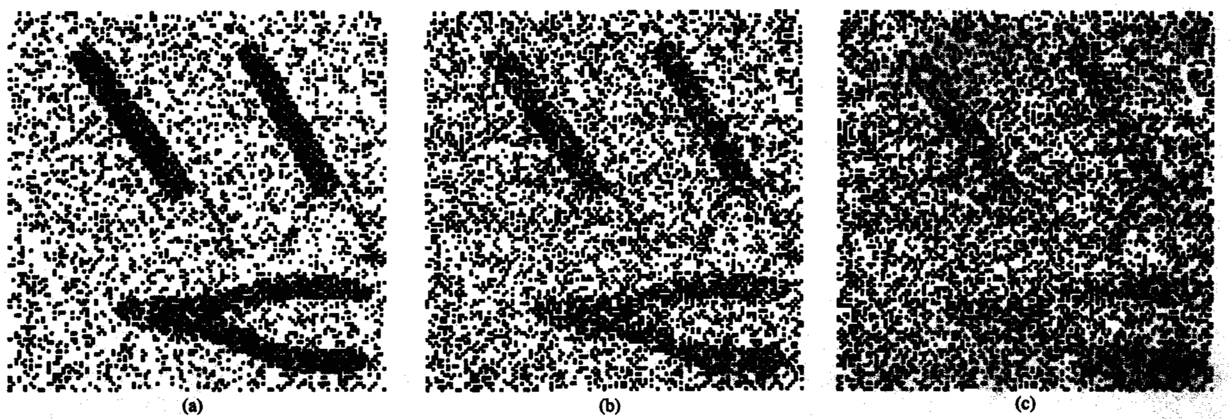

Figure 6

Image smoothing: noise-corrupted bitmaps. The image in Fig. 5 has been corrupted by adding random noise. Signal-to-noise ratios are 1.5, 0.87, and 0.41 in (a), (b), and (c), respectively.

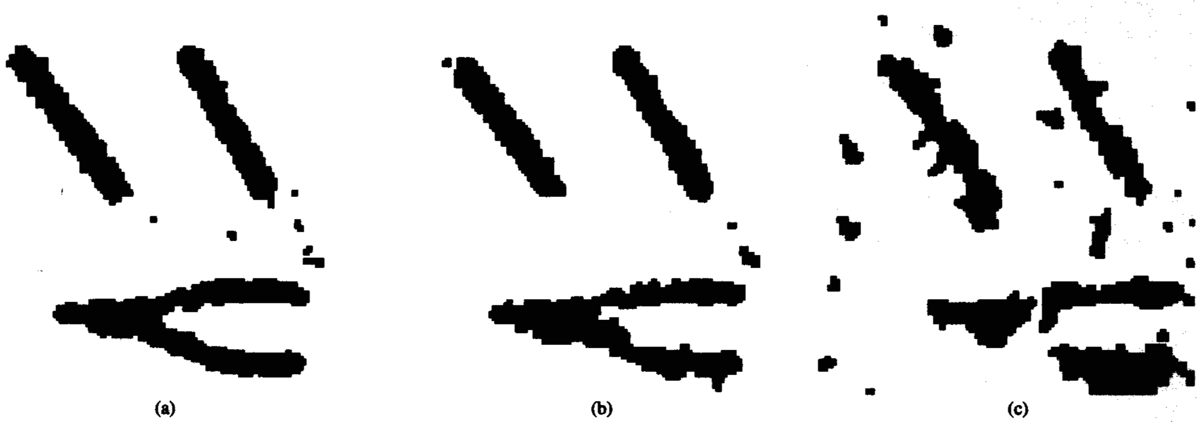

Figure 7

Results of simulated annealing on images of Fig. 6 obtained with $\lambda = 1$.

problem and the ground state problem for an Ising system imbedded in an external field. The image to be smoothed determines γ and thus plays the role of the external field.

The method has been tested using the original bitmap shown in **Figure 5** and corrupting it as in the problem of parameter estimation previously described. **Figure 6** shows the corrupted versions corresponding to $r = 0.2$, $r = 0.3$, and $r = 0.4$ (signal-to-noise ratios 1.5, 0.87, and 0.41). **Figures 7** and **8** show the results of simulated annealing, with different values of λ.

For comparison, **Figure 9** shows the results of applying the usual image processing algorithms to the same bitmap.

Image halftoning and antiferromagnetic systems

In the previous section we have shown how the process of smoothing a bitmap is formally equivalent to a ground state problem for a ferromagnetic Ising model imbedded in an external field. In this section we show that the same problem, but with a reversed sign for the spin-spin interaction (antiferromagnetic coupling), is equivalent to the problem of *halftoning*, i.e., of calculating a bi-level image (bitmap) whose average density mimics the one of a given continuous image.

Let γ_{ij} be the given continuous image with values in the interval $(-1, 1)$ and μ_{ij} the resulting bitmap with values -1,

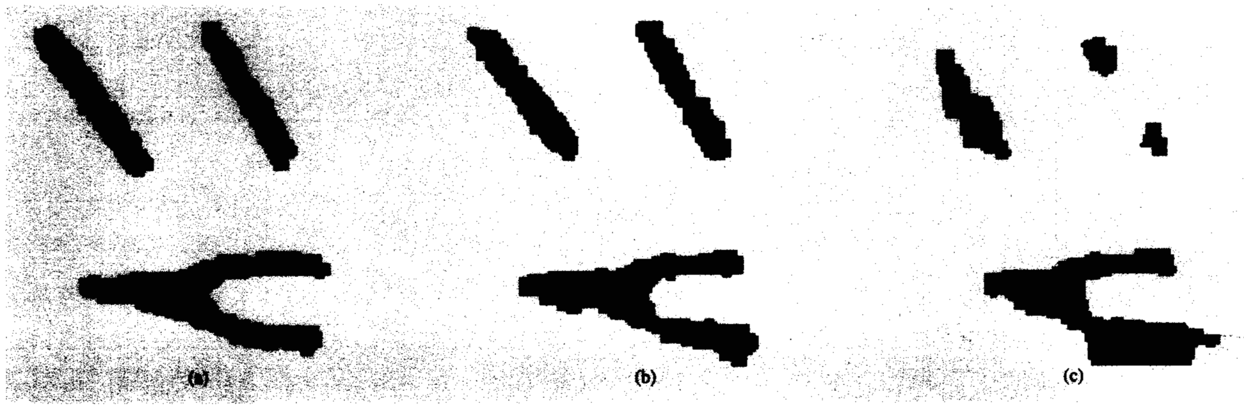

Figure 8

Results of simulated annealing on images of Fig. 6 obtained with $\lambda = 2$.

1. If one wishes to define the cost function for such a problem, one may try to minimize the difference between the continuous-level image and an average density of the bitmap, which can be suitably defined through the introduction of a filter V:

$$\rho_{ij} = \sum_k \sum_l V_{ijkl}\mu_{kl}\,. \tag{28}$$

The general properties of V_{ijkl}, in order to deal with a proper average, are the following:

$$\sum_k \sum_l V_{ijkl} = 1 \qquad V_{ijkl} \geq 0. \tag{29}$$

Moreover, if we ask for an average operation which is the same for the whole lattice, we have $V_{ijkl} = V_{i-k,j-l}$. In principle one can choose to extend the average over a large spatial extent, but it is more practical to restrict it to a small region around the spin μ_{ij}. With this choice V is a square matrix whose dimension is related to the spatial size of the average. The form of the cost function, choosing again the L^2 norm to measure the distance between the original image and the average introduced, is the following:

$$E = \sum_i \sum_j (\gamma_{ij} - \rho_{ij})^2. \tag{30}$$

Developing this expression, neglecting terms which do not depend on μ_{ij} and rescaling the energy by a factor $\frac{1}{2}$ (as in the previous paragraph), we obtain

$$E = -\frac{1}{2}\sum_{ij} \tilde{\gamma}_{ij}\mu_{ij} + \frac{1}{2}\sum_{i_1j_1i_2j_2} L_{i_1j_1i_2j_2}\mu_{i_1j_1}\mu_{i_2j_2}, \tag{31}$$

where we defined

$$\tilde{\gamma}_{ij} = \sum_k \sum_l V_{ijkl}\gamma_{kl} \tag{32}$$

and

$$L_{ijkl} = \sum_n \sum_m V_{ijnm}V_{nmkl}\,, \tag{33}$$

which is the convolution of the filter V_{ijkl} with itself. We recognize that the expression in (31) is of the general type (11); in fact, $\tilde{\gamma}_{ij}$ plays the role of the external field, L_{ijkl} is a connection matrix, and the constant J has the value -1. This shows the equivalence between image halftoning and the ground state problem for an Ising system. Due to the sign of J and to the properties of V_{ijkl}, this term in the energy is to be regarded as an antiferromagnetic interaction, which tends to align two neighboring spins in opposite directions. In the context of image processing the two terms have a straightforward interpretation: While the magnetic field term has the function of keeping the bitmap as similar as possible to the original image, the antiferromagnetic part of the energy is intended to produce the diffusion effect proper of halftoning. We performed the annealing on a 256×256 image with 256 gray levels; the resulting bitmaps for different filters V are shown in **Figure 10**. We discuss later the details of the filters V used and the influence of the choice of the filter on the result.

From the statistical mechanics point of view, we are in the presence of a complex general interaction: an average magnetic field, strongly depending on the position, which would tend to align the spins according to its direction, and the internal field which might be in conflict with it. In other words, it is possible to observe for this system the frustration effects already mentioned in [2]: A system subjected to different physical constraints may show very peculiar behavior in the annealing procedure; in fact, due to this competition, low-energy states may be highly degenerate and a great many local minima very near to the ground state energy may appear.

This statistical model is interesting *per se*, so we devote part of this section to the study of its thermodynamical features. A discussion of the issues concerning image processing follows.

By applying the annealing procedure in the case $\gamma = 0$ (no magnetic field), we observed that no phase transition occurs going from high to low temperature regardless of the filter V used. This is quite a common feature in two-dimensional antiferromagnetic spin systems.

To investigate the critical behavior of the system, we first analyzed the simplified case of uniform magnetic fields (i.e., $\gamma_{ij} \equiv \gamma$, constant over the whole lattice) and then observed the implications of such analysis on a real image with varying gray level.

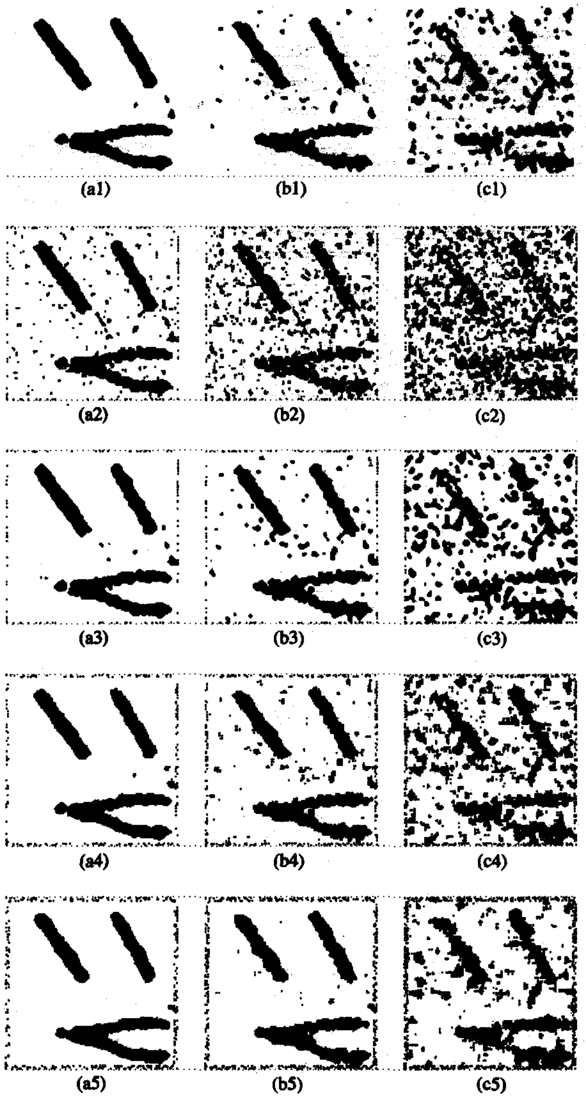

Figure 9

Image smoothing: application of standard algorithms. The bitmaps in Fig. 6 have been smoothed using the following standard algorithms: 3×3 low-pass filter iterated three times (a1, b1, c1); 3×3 median filter (a2, b2, c2); 3×3 median filter iterated three times (a3, b3, c3); 5×5 (a4, b4, c4), and 7×7 (a5, b5, c5) median filters. Comparison of these results with those of Figs. 7 and 8 shows that in the presence of low signal-to-noise ratios, simulated annealing yields comparatively better smoothing.

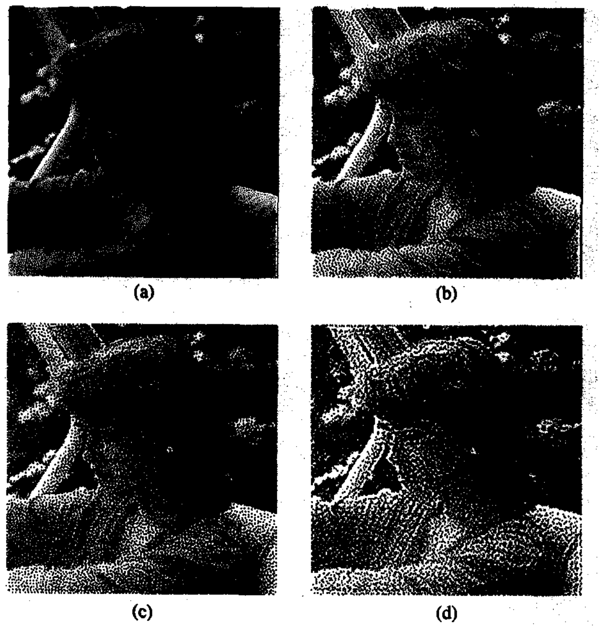

Figure 10

Image halftoning by simulated annealing. (a) Original 256×256, 8-bit/pel image; (b) result of halftoning by the standard MECCA algorithm [9]; (c) result of simulated annealing with a uniform 3×3 filter V; (d) result of simulated annealing with a uniform 5×5 filter V. Comparison of (c) and (d) shows the trade-off between spatial and tonal resolution.

A very useful parameter for understanding the response of the system to an applied magnetic field is the so-called order parameter (also called net magnetization), which is defined as

$$M \equiv \left| \sum_i \sum_j \frac{\mu_{ij}}{N^2} \right|. \quad (34)$$

If M equals 1, then the system is in a completely ordered state, i.e., all spins point to the same direction; conversely, in the absence of order, M would equal 0. We fixed γ and performed the annealing, bringing the system toward zero temperature; thus we reached approximately the ground state for the given γ and we measured its magnetization. Repeating this procedure for increasing values of γ, we found the occurrence of a phase transition driven by the external field: There is a critical value γ_c under which the magnetization grows linearly with the strength of the field; for $\gamma > \gamma_c$ there is an abrupt change in this response, and the magnetization jumps to the maximum value, i.e., $M = 1$ (**Figure 11**). To determine the location of this transition, we

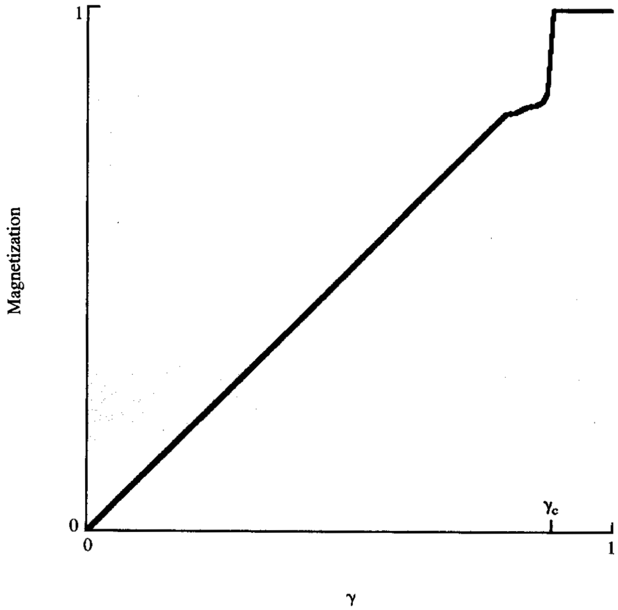

Figure 11

Net magnetization versus external field at $T=0$. This plot has been obtained by simulated annealing using a position-independent γ and a uniform 3×3 filter V.

present the following argument. Let us consider a magnetic field γ very near to 1, i.e., $\gamma = 1 - 2\epsilon$, with ϵ very small and positive (due to the symmetry of the problem, we could have chosen $\gamma = -1 + 2\epsilon$ with identical results). We can argue on intuitive grounds that if ϵ is chosen very near to 0, the state with all $\mu_{ij} = 1$ is the most favorable energetically; i.e., it is the ground state of the system. To confirm this picture, we compare the energy E_0 of this configuration with that of the state with a single spin reversed, E_1 (this is often indicated as an elementary excitation of the system); to be definite, we say that $\mu_{ij} = -1$ if $i = j = 0$ and $\mu_{ij} = 1$ otherwise, but translational symmetry guarantees the generality of the argument. Using Eq. (30) to compute the two energies, we find

$$\Delta E \equiv E_0 - E_1 = 2\epsilon - \sum_i \sum_j V_{ij}^2 . \tag{35}$$

Thus $\Delta E \leq 0$ if $\epsilon \leq \frac{1}{2} \sum_i \sum_j V_{ij}^2$; defining $\epsilon_c = \frac{1}{2} \sum_i \sum_j V_{ij}^2$, we have $\gamma_c = 1 - 2\epsilon_c$. In fact, when $\gamma > \gamma_c$ the state with all $\mu_{ij} = 1$ is the stable phase of the system, and $M = 1$; when γ is lowered under γ_c, this configuration becomes energetically unstable with respect to an elementary excitation, so we must conclude that the ground state will have a more complex structure, due to the increasing relevance of antiferromagnetic interaction. We notice that the critical field is near to 1 as $\sum_i \sum_j V_{ij}^2$ gets smaller. For the case in Fig. 11, where V is a uniform 3×3 filter, one obtains $\gamma_c = 0.89$, which agrees with the result in the figure, obtained by simulating annealing.

Let us see what this analysis implies for image processing. A uniform region of the original image with a gray level very near to black (white) corresponds to a subset of the spin lattice under the influence of a field very near to 1 (-1). The location of the phase transition (and its dependence on the specific filter used) tells us to which extent intermediate levels may be controlled: Thus a gray region very near to black ($\gamma > \gamma_c$) will be converted by the annealing procedure into a region of the bitmap completely black (with all $\mu_{ij} = 1$); in other words, the expression $\sum_i \sum_j V_{ij}^2$ is a measure of the tonal resolution of the filter.

For a given spatial extent of the filter (that is to say, for a given dimension of the matrix V), the best filter in terms of tonal resolution is the one in which the average is equally weighted for all the neighboring pixels, i.e., a matrix V with all the elements equal. In this case we have $\epsilon_c = 1/2m^2$ (where m is the size of V), from which we can deduce that the tonal resolution increases quadratically with the spatial extent of the average (or linearly with the number of pixels involved in the average). Of course, one is forced to pay a price in terms of spatial resolution of the image. The loss of spatial resolution due to the averaging can be measured by

$$\sigma^2 \equiv \sum_i \sum_j V_{ij}(i^2 + j^2); \tag{36}$$

we obtain for this class of filters

$$\sigma^2 = 2m \frac{(m + 1)}{3}. \tag{37}$$

Thus, spatial resolution decreases when tonal resolution increases and vice versa. For large m, spatial resolution is inversely proportional to tonal level resolution. Difficulties deriving from this trade-off could be overcome by using a modified filter (with larger m) only in regions for which $|\gamma| > \gamma_c$.

Conclusions

It has been shown that seemingly unrelated problems in image processing, such as smoothing or halftoning a bitmap, are analogous to ground state problems for spin systems with different kinds of interaction: ferromagnetic and antiferromagnetic, respectively. Simulated annealing is a good tool to solve these problems, as well as other problems in image processing, such as the one of parameter estimation described in this paper. In the case of highly noise-corrupted bitmaps (both in parameter estimation and image smoothing problems), this technique gives better results than existing methods. As far as halftoning is concerned, results are comparable with those of standard algorithms. The main drawback of the annealing procedure is that computer requirements are larger and, in general, not easily predictable. On the other hand, the connection between image processing and Ising spin systems is important because the extensive results on Ising spin systems might provide useful hints when approaching the corresponding problems in image processing.

References

1. S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by Simulated Annealing," *Science* **220,** 671 (1983).
2. S. Kirkpatrick, "Optimization by Simulated Annealing: Quantitative Studies," *J. Statist. Phys.* **34,** 975 (1984).
3. N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, "Equation of State Calculations by Fast Computing Machines," *J. Chem. Phys.* **21,** 1087 (1953).
4. K. A. Narayanan, D. O'Learly, and A. Rosenfeld, "Image Smoothing and Segmentation by Cost Minimization," *IEEE Trans. Syst., Man, Cyber.* **SMC-12,** 91 (1982).
5. K. Huang, *Statistical Physics*, John Wiley & Sons, Inc., New York, 1963, pp. 329–373.
6. H. E. Stanley, *Introduction to Phase Transition and Critical Phenomena*, Oxford University Press, Oxford, England, 1971.
7. E. Ising, "A Contribution to the Theory of Ferromagnetism," *Z. Phys.* **31,** 253 (1925).
8. L. Onsager, "A Two-Dimensional Model with an Order-Disorder Transition," *Phys. Rev.* **65,** 117 (1944).
9. P. Stucki, "MECCA—A Multiple-Error Correction Computation Algorithm for Bi-Level Image Hardcopy Reproduction," *Research Report RZ-1060*, IBM Research Laboratory, Zurich, Switzerland, 1981.

Received February 11, 1985; revised June 10, 1985

Paolo Carnevali *IBM Italy, Scientific Center, Via Giorgione, 129, 00147 Rome, Italy.* Dr. Carnevali is currently working at the Rome Scientific Center in the parallel processing group. He obtained his degree in physics from the University of Rome in 1979 and subsequently spent two years at the IBM Thomas J. Watson Research Center in Yorktown Heights, New York, as part of a postdoctoral fellowship program, working on the dynamics of galaxies and clusters of galaxies. Since joining IBM Italy in 1982, he has worked in the fields of solid-state physics, parallel computing, and image processing. In 1983 Dr. Carnevali received an IBM Outstanding Technical Achievement Award for his work on the IBM 3838 Array Processor.

Lattanzio Coletti *IBM Italy, Scientific Center, Via Giorgione, 129, 00147 Rome, Italy.* Dr. Coletti joined the Rome Scientific Center in 1982 as a systems engineer. His current research is in image acquisition and processing, specifically in image acquisition devices connected to IBM computers and in restoration of blurred images. He was involved in the development of the IBM 7350 Image Processing System (previously known as HACIENDA). Dr. Coletti received his degree in electronics from Rome University in 1980.

Stefano Patarnello *IBM Italy, Scientific Center, Via Giorgione, 129, 00147 Rome, Italy.* Dr. Patarnello obtained his degree in physics from the University of Pisa in 1983. His main interest has been in statistical physics, with emphasis on the application of renormalization group techniques to the critical behavior of systems involving multivalued spins. Dr. Patarnello joined the IBM Rome Scientific Center in 1985, and at present is involved in large-scale numerical computations in fluid dynamics.