Simulated Tearing: an Algorithm for Discontinuity-Preserving Visual Surface Reconstruction

Mário A. T. Figueiredo and José M. N. Leitão

Centro de Análise e Processamento de Sinais,
Departamento de Engenharia Electrotécnica e de Computadores,
Instituto Superior Técnico, Lisboa, PORTUGAL

Abstract

We introduce a new algorithm for discontinuity-preserving visual surface reconstruction, inspired on the well known formulation of the problem as the fitting of a weak membrane to the observed data. The method works by slowly applying the data "force" to a weak membrane which is allowed to tear when the tension exceeds a certain threshold; accordingly, the algorithm is named "simulated tearing" (ST). Formally, ST is a deterministic continuation method, i.e. the problem to be solved is embeded in a family of problems, of which the first member has a simple solution. The proposed method is tested and compared with mean field annealing (MFA), using real and synthetic images. Conclusions are that ST is simpler, faster and slightly outperforms MFA. Furthermore, ST allows implementation based on integer arithmetic.

1 Introduction

Reconstruction of visual surfaces representing some property of a scene (brightness, shape, distance, motion, ...), from observed image data, is known to be an ill-posed inverse problem [2], [11], [17]. Both regularization and Bayesian approaches provide ways to achieve stability by incorporating *a priori* knowledge or constraints into the problem.

In earlier work, the solutions were restricted to some class of continuous and smooth functions [1], [2], [3], [17], this being clearly unreasonable in the presence of discontinuities which are key features of visual perception. The incorporation of discontinuity detection into the surface reconstruction process has been studied by several authors. Some fundamental references are [4], [5], [13], [16], and also [8], [9], [11], these last ones in a Bayesian framework. Independently of their theoretical background, these formulations include a fundamental concept which is that of *weak constraint* [4], and lead to the same type of hard optimization problems (non-convex, huge dimensionality) for which several classes of methods have been proposed:

- Monte Carlo type stochastic methods, such as simulated annealing [9] and the Metropolis algorithm [11], which are (theoretically) able to reach optimal solutions at the cost of an enormous computational load;
- deterministic continuation methods, not guaranteed to reach a global optimum but much faster than the ones in the previous class, such as *graduated non convexity* (GNC) [5], *mean field annealing* (MFA) [8], [18], or the one proposed in [10];
- nonlinear/anisotropic diffusion methods [14], [15], which are iterative deterministic algorithms that work by simulating a diffusion process.

This paper introduces a new fully deterministic continuation method for discontinuity-preserving visual surface reconstruction. The algorithm is inspired on the interpretation of the reconstruction process as the fitting of a *weak membrane* [5] to the observed data. In the class of continuation methods, to which the proposed technique belongs, the function to be minimized is embeded in a family of functions (usually depending on a control parameter) of which the first member is easy to minimize; this minimum is then tracked along the family of minima (by varying the control parameter) until the desired solution is reached.

To understand the idea behind our approach, suppose we want to adjust a weak membrane to some hard discontinuous surface, allowing it to tear only when needed to fit the discontinuities. What physical intuition suggests is that we slowly press the membrane against the hard surface, letting it tear little by little, instead of forcing it with a single stroke. To translate this idea into our problem, consider the visual surface to be reconstructed as a discrete weak membrane

1063-6919/93 $03.00 © 1993 IEEE

which is to be fitted to a 2D force field (the observed data); the membrane is weak in the sense that it tears if the tension between two neighbor points exceeds a certain threshold. Let the force field be applied progressively, starting at a very low intensity up to its full magnitude. At the first steps, when the force field is faint, the surface has no discontinuities; as the force field increases, the higher tension sites start to *tear*; when the full force field is reached, the solution is a piece-wise continuous surface, torn at the locations of the detected discontinuities.

2 Problem formulation

Let discrete surfaces, defined on a $M \times M$ lattice $\mathcal{L} = \{(i,j),\ i,j = 1,...,M\}$, be described by 2D arrays of "heights" $\mathbf{x} = \{x_{i,j},\ (i,j) \in \mathcal{L}\}$. Within a Bayesian setting, the *maximum a posteriori* (MAP) estimate of a discrete surface $\mathbf{x}$, with a first order piecewise Gauss Markov random field prior, given the noisy observed data field $\mathbf{y} = \{y_{i,j},\ (i,j) \in \mathcal{L}\}$ (see [8] or [9], for further details), is

$$\left(\widehat{\mathbf{x}}, \widehat{\mathbf{h}}, \widehat{\mathbf{v}}\right)_{\text{MAP}} = \arg \min_{\mathbf{x},\mathbf{h},\mathbf{v}} E(\mathbf{x},\mathbf{h},\mathbf{v}) \qquad (1)$$

where

$$\begin{aligned} E(\mathbf{x},\mathbf{h},\mathbf{v}) = & \frac{1}{2\sigma^2}\sum_{i,j}(x_{i,j}-y_{i,j})^2 + \alpha\sum_{i,j}(h_{i,j}+v_{i,j}) \\ & + \mu\sum_{i,j}(x_{i,j}-x_{i,j-1})^2(1-v_{i,j}) \\ & + \mu\sum_{i,j}(x_{i,j}-x_{i-1,j})^2(1-h_{i,j}) \end{aligned} \qquad (2)$$

In (2), σ^2 is the (additive, white, and Gaussian) noise variance, μ and α are model parameters, and $\mathbf{h}$ and $\mathbf{v}$ (with $h_{i,j}, v_{i,j} \in \{0,1\}$, for $(i,j) \in \mathcal{L}$) are the, respectively, horizontal and vertical *line processes*, introduced to weaken the continuity constraint. Notice that when some line process site, say $h_{i,j}$, is *on* ($h_{i,j} = 1$), the continuity constraint $(x_{i,j} - x_{i-1,j})^2$ is turned *off*. There is, however, a cost α associated with turning *on* one line process element. This competition between the continuity term and the line creation process gives rise to simultaneous surface-reconstruction/discontinuity-detection. In (2), we assume that the penalty for line creation is the same for the horizontal and the vertical processes and uniform over $\mathcal{L}$, and that there is no interaction among line processes sites. The off-lattice terms (those for which $(i,j) \notin \mathcal{L}$) are taken as zero, i.e. free boundary conditions are assumed. *Energy* function (2), which can also be derived under a regularization framework, is known as the *weak membrane* model [5]. Due to the presence of the line processes $\mathbf{h}$ and $\mathbf{v}$, $E(\mathbf{x},\mathbf{h},\mathbf{v})$ is not convex and its minimization is not trivial. Note that although the model has three parameters, α, σ, and μ, the estimate (1) only depends on two: α/μ (the line creation threshold) and $\mu\sigma^2$ (which can be seen as a *noise to signal ratio*).

Before proceeding, it is convenient to cast the problem into vector notation. Let $\mathbf{x}$, $\mathbf{y}$, $\mathbf{h}$, and $\mathbf{v}$ now stand for vectors obtained by lexicographically ordering the field elements, i.e. $\mathbf{x} = [x_{1,1}, x_{1,2}, ..., x_{1,M}, x_{2,1}, ..., x_{M,M}]^T$, and similarly for $\mathbf{y}$, $\mathbf{h}$, and $\mathbf{v}$. Let us also define the vector $\mathbf{1} = [1,1,1,...,1]^T$, and denote a $n \times n$ identity matrix by $\mathbf{I}_n$. Then, the energy $E(\mathbf{x},\mathbf{h},\mathbf{v})$ can be written as

$$\begin{aligned} E(\mathbf{x},\mathbf{h},\mathbf{v}) = & \ \mathbf{x}^T\left(\mu\mathbf{A}(\mathbf{h},\mathbf{v}) + \frac{\mathbf{I}_{M^2}}{2\sigma^2}\right)\mathbf{x} \\ & + \alpha\mathbf{1}^T(\mathbf{v}+\mathbf{h}) - \frac{1}{\sigma^2}\mathbf{x}^T\mathbf{y}, \end{aligned} \qquad (3)$$

where additive terms, independent of $\mathbf{x}$, $\mathbf{h}$, and $\mathbf{v}$, thus irrelevant to the optimization problem, were omitted. In (3), $\mathbf{A}(\mathbf{h},\mathbf{v})$ is a $M^2 \times M^2$ highly sparse and structured matrix (block tridiagonal), function of $\mathbf{h}$ and $\mathbf{v}$,

$$\mathbf{A}(\mathbf{h},\mathbf{v}) = \begin{bmatrix} \mathbf{B}_1 & \mathbf{C}_1 & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{C}_1 & \mathbf{B}_2 & \mathbf{C}_2 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_2 & \mathbf{B}_3 & \mathbf{C}_3 & \cdots & \mathbf{0} \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & \mathbf{C}_{M-1} \\ \mathbf{0} & \cdots & \cdots & \cdots & \mathbf{C}_{M-1} & \mathbf{B}_M \end{bmatrix},$$

where $\mathbf{0}$ is a $M \times M$ block of zeros, and $\mathbf{B}_i$ and $\mathbf{C}_i$ (also highly sparse and structured blocks) are given by:

$$\mathbf{B}_i[m,n] = \begin{cases} (4 - v_{i,m} - v_{i,m+1} \\ -h_{i,m} - h_{i,m+1}), & n = m = 1,...,M \\ v_{i,m} - 1, & n = m-1 \\ v_{i,m+1} - 1, & n = m+1, \end{cases}$$

and

$$\mathbf{C}_i = \text{diag}\left\{(h_{i,1}-1), (h_{i,2}-1), \cdots, (h_{i,M}-1)\right\}.$$

This notation, which is a modification of the one proposed in [12] to include the line processes, can easily be generalized in the following aspects:

- Higher orders; e.g. in the second order model, first order differences $(x_{i,j} - x_{i,j-1})$ are replaced by second order differences $(2x_{i,j} - x_{i,j-1} - x_{i,j+1})$, matrix $\mathbf{A}(\mathbf{h},\mathbf{v})$ is block pentadiagonal and matrices $\mathbf{B}_i$ are pentadiagonal. The second order model is called (in [5]) the *weak plate*.

- An observation operator $\mathbf{H}$, modeling, for example, a blur effect, could also have been considered; in (3), matrix $\mathbf{I}_{M^2}$ would be replaced by $\mathbf{H}^T\mathbf{H}$, and $\mathbf{y}$ by $\mathbf{H}^T\mathbf{y}$. Vectors $\mathbf{y}$ and $\mathbf{x}$ can be of different dimensions (non-square $\mathbf{H}$).

- Nonhomogeneous penalty for line creation; this would imply the replacement of the term $\alpha \mathbf{1}^T(\mathbf{v}+\mathbf{h})$ by $\boldsymbol{\alpha}_v^T\mathbf{v}+\boldsymbol{\alpha}_h^T\mathbf{h}$, where $\boldsymbol{\alpha}_v$ and $\boldsymbol{\alpha}_h$ are penalty vectors.

- Rectangular (nonsquare) image lattices can also be straightforwardly adopted.

The algorithm proposed in the next section can easily be extended to encompass any of the above mentioned aspects.

3 Proposed algorithm

3.1 Preliminary facts

We propose a new continuation method, in which the function to be minimized, $E(\mathbf{x},\mathbf{h},\mathbf{v})$, is embeded in a family of functions $\{E_\phi(\mathbf{x},\mathbf{h},\mathbf{v})\ ,\ \phi \in [0;1]\}$ controlled by parameter ϕ, as follows:

$$E_\phi(\mathbf{x},\mathbf{h},\mathbf{v}) = \mathbf{x}^T\left(\mu\mathbf{A}(\mathbf{h},\mathbf{v})+\frac{\mathbf{I}_{M^2}}{2\sigma^2}\right)\mathbf{x} + \alpha\mathbf{1}^T(\mathbf{v}+\mathbf{h}) - \frac{\phi}{\sigma^2}\mathbf{x}^T\mathbf{y}. \quad (4)$$

Some facts are now invoked for future reference:

a) It is clear that $E_1(\mathbf{x},\mathbf{h},\mathbf{v}) = E(\mathbf{x},\mathbf{h},\mathbf{v})$.

b) For sufficiently small ϕ, the minimization of (4) leads to $\{h_{ij}=0,\ (i,j)\in\mathcal{L}\}$, $\{v_{ij}=0,\ (i,j)\in\mathcal{L}\}$, and to some continuous surface $\mathbf{x}$. Since we are dealing with digital images which are bounded below by 0 and above by some constant L, it is enough (although better criteria can be devised) to consider $\phi < L^{-1}\sqrt{\alpha/\mu}$.

c) Given some fixed line configuration $\mathbf{h}_0$ and $\mathbf{v}_0$, the energy is quadratic with respect to $\mathbf{x}$,

$$E_\phi(\mathbf{x},\mathbf{h}_0,\mathbf{v}_0) \propto \mathbf{x}^T\left(\mu\mathbf{A}(\mathbf{h}_0,\mathbf{v}_0)+\frac{\mathbf{I}_{M^2}}{2\sigma^2}\right)\mathbf{x} - \frac{\phi}{\sigma^2}\mathbf{x}^T\mathbf{y}, \quad (5)$$

its minimization being trivial,

$$\widehat{\mathbf{x}}(\mathbf{h}_0,\mathbf{v}_0,\phi) = \frac{\phi}{2\sigma^2}\left(\mu\mathbf{A}(\mathbf{h}_0,\mathbf{v}_0)+\frac{\mathbf{I}_{M^2}}{2\sigma^2}\right)^{-1}\mathbf{y}.$$

The only difficulty is in the inversion of the very high dimension matrix involved, for which we have proposed neural structures [6], [7]. A recursive solution, based on the Cholesky decomposition of matrix $\mathbf{A}$, can also be considered [12].

3.2 Algorithm

The proposed method is supported on the above facts and works as follows:

Step 0 Set parameter ϕ to some small positive constant $\phi^{(0)}$. Perform the minimization of $E_{\phi^{(0)}}(\mathbf{x},\mathbf{h},\mathbf{v})$, knowing that $\phi^{(0)}$ is small enough to invoke fact b). The result of these first step is a continuous surface $\mathbf{x}^{(0)}$ and a zero line process configuration $\mathbf{h}^{(0)}=\mathbf{0}$, and $\mathbf{v}^{(0)}=\mathbf{0}$.

Step 1 Using the present line processes configuration $\mathbf{h}^{(t)}$ and $\mathbf{v}^{(t)}$, and parameter $\phi^{(t)}$, compute

$$\mathbf{x}^{(t)} = \frac{\phi^{(t)}}{2\sigma^2}\left(\mu\mathbf{A}(\mathbf{h}^{(t)},\mathbf{v}^{(t)})+\frac{\mathbf{I}_{M^2}}{2\sigma^2}\right)^{-1}\mathbf{y}, \quad (6)$$

by applying, for example, any of the techniques referred in fact c).

Step 2 Update the line processes according to the following rule:

$$\begin{cases} (x_{i,j}^{(t)} - x_{i,j-1}^{(t)})^2 > \frac{\alpha}{\mu} & \Rightarrow \quad v_{i,j}^{(t+1)} = 1 \\ (x_{i,j}^{(t)} - x_{i-1,j}^{(t)})^2 > \frac{\alpha}{\mu} & \Rightarrow \quad h_{i,j}^{(t+1)} = 1 \end{cases} \quad (7)$$

Step 3 Compute $\phi^{(t+1)} = \phi^{(t)} + \Delta\phi$. If $\phi^{(t+1)} > 1$ stop; if not, compute matrix $\mathbf{A}(\mathbf{h}^{(t+1)},\mathbf{v}^{(t+1)})$ and go back to Step 1.

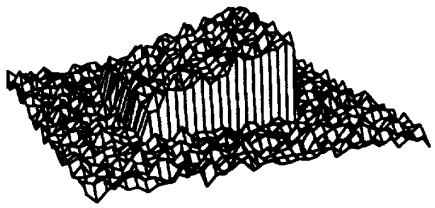

Figure 1: **Sloped flat surface over a constant background plus Gaussian white noise.**

The scheme just described is in fact a continuation method. It tracks the solution along a family of functions, of which the last member $E_1(\mathbf{x},\mathbf{h},\mathbf{v})$ is the original function to be minimized. The algorithm can easily be modified to take into account more general models, as was mentioned at the end of Section 2.

The physical motivation and interpretation presented in the introduction can now be related with the formal description just given. The 2D field $\mathbf{x}$ represents a piece-wise continuous membrane, torn vertically (horizontally) at the sites for which $v_{ij} = 1$ ($h_{ij} = 1$). The increasing force field applied to the surface is represented by $\phi\mathbf{y}$. When ϕ is small, the solution surface (see (6)) fits well the force field without being torn at any location, i.e. $(x_{i,j} - x_{i,j-1})^2 < \alpha/\mu$ and $(x_{i,j} - x_{i-1,j})^2 < \alpha/\mu$, for all sites (i, j). When the intensity of the force field ϕ raises, the tension at some sites starts to exceed the threshold α/μ and the surface is torn. Figures 1 and 2 illustrate the working of ST (gray dashed lines represent broken line sites).

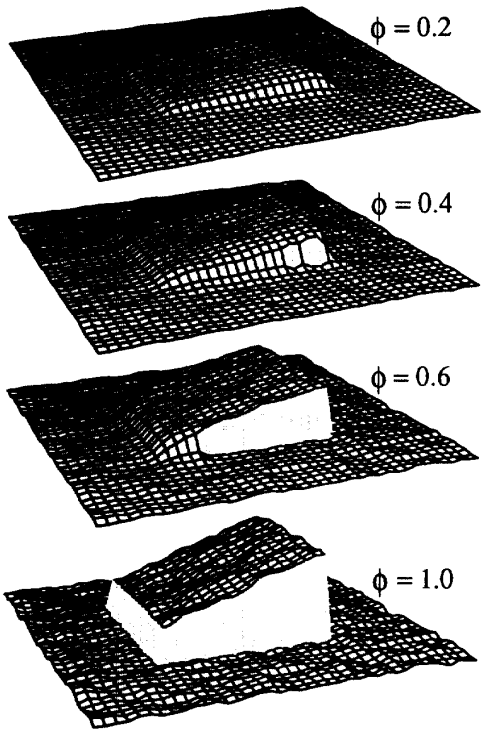

Figure 2: **Surface reconstruction from the data of Fig.1, for $\phi = 0.2$, $\phi = 0.4$, $\phi = 0.6$, and $\phi = 1.0$.**

A key aspect of the algorithm is the slow increase of ϕ (force field intensity), such that only a few locations have to be torn at each iteration. This is crucial with respect to the creation of double edges. Figures 3 and 4 illustrate this aspect for a 1D problem. Fig. 3(a) shows noisy data having a discontinuity around pixel 27. In Fig. 3(b) the data has been smoothed;

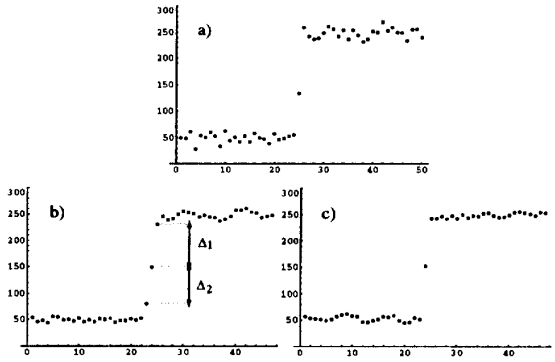

Figure 3: **a) Noisy data with discontinuity around pixel 27; (b) After smoothing, both Δ_1 and Δ_2 exceed the discontinuity threshold. (c) Final reconstruction with two discontinuities.**

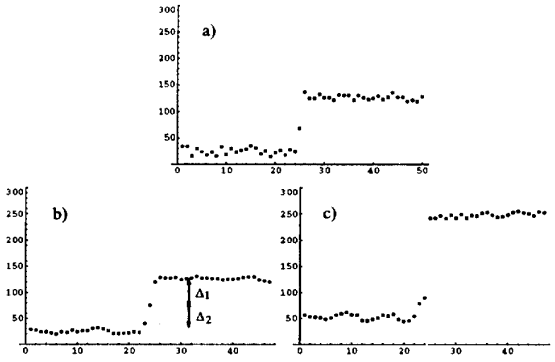

Figure 4: **(a) The same data of Fig.3(a) multiplied by 0.5; (b) After smoothing, only Δ_1 exceeds the discontinuity threshold. (c) Final reconstruction with a single discontinuity.**

let Δ_1 and Δ_2 denote the differences between 3 consecutive pixel intensities around the discontinuity. If the line creation threshold is smaller than both Δ_1 and Δ_2, two discontinuities will be signaled (a double edge) and the final reconstructed *weak string* (the 1D version of the weak membrane [5]) appears as shown in Fig. 3(c). If, however, we start with a weaker force field (Fig.4(a)) and increase its strength slowly, there will very probably be some point at which only one of the differences Δ_1 or Δ_2, say Δ_1, will exceed the discontinuity detection threshold (Fig. 4(b)). At this point, the line process will be turned on at that site, subsequent smoothing will decrease Δ_2, and the result will have only one discontinuity, as can be seen in Fig. 4(c).

4 Examples

In this section some examples, comparing the ST algorithm with MFA are presented. The GNC algorithm could also have been considered; however, it can be seen as an approximation of MFA [8], which does not produce better results and has a similar computational load. In our tests, ST was about 2 times faster than MFA, for the same number of steps in the continuation process. This is mainly due to the more complex line process computations of MFA (which are sigmoidal functions, see [8]) when compared to the very simple threshold-based Step 2 of the ST algorithm. Expression (7) allows implementation based on integer arithmetic, as long as an adequate scaling is used. Since Step 1 of ST can also be implemented using only integer computations, without a significant degradation of the result [6], the complete algorithm can be implemented on integer arithmetic.

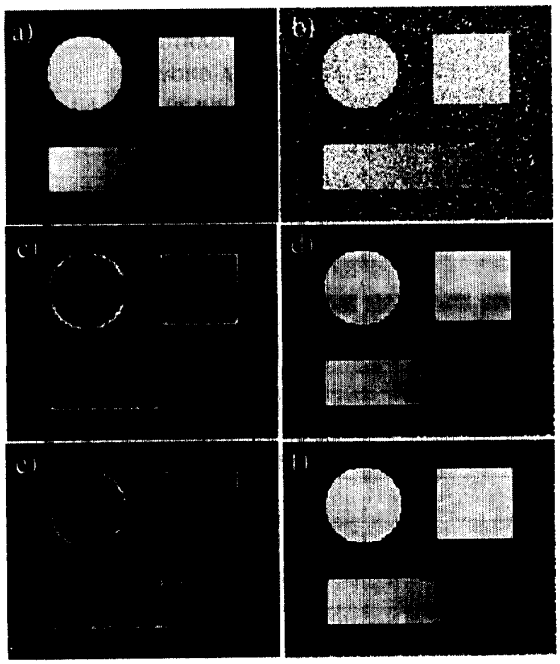

Figure 5: **Original image (a) and its (σ = 40) noise contaminated version (b). Edges (c) and reconstructed image (d) obtained with ST. Edges (e) and reconstructed image (f) obtained with MFA.**

A test with a synthetic image contaminated by noise (σ = 40) is exhibited in Fig. 5. The parameters used were $\sigma = 40.0$, $\mu = 0.1$, and $\alpha = 70$. For ST, we used $\phi^{(0)} = 0.1$ and $\Delta\phi = 0.1$; for MFA parameter β (see [8]) evolved as $\beta^{(t)} = 0.0002 \times K^t$, stopped at $\beta = 1$ (as suggested in [18]) with K such that also 10 iterations were performed. The lower (about 8%) energy was obtained by ST.

To give experimental support to the statement made at the end of the previous section, concerning the importance of the slow increase of the force field intensity, we studied the influence of the number of iterations on the final energy obtained. The same study was also performed for MFA. The results are summarized in Fig. 6. For both algorithms, the control parameters kept the same initial and final values ($\phi^{(0)} = 0.1$, $\phi^{(\text{last})} = 1$, for ST; $\beta^{(0)} = 0.0002$ up to $\beta = 1$, for MFA), only the variation rate being changed ($\Delta\phi$, for ST, and K for MFA). The final energy obtained does in fact decrease when the number of iterations is increased. For up to 4 iterations, MFA performed slightly better than ST. Above 5 iterations, ST outperformed MFA, with both methods reaching a plateau (with ST about 10% better than MFA) at about 15 $\sim$ 20 iterations.

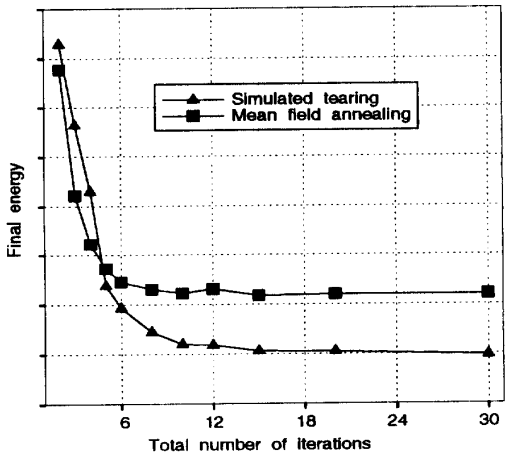

Figure 6: **Final energy reached by ST and MFA as a function of the total number of iterations performed .**

In Figure 7, ST is compared with MFA using a real image and equal model parameters ($\sigma = 1.0$, $\mu = 0.5$, $\alpha = 120$, 15 iterations). The results are visually similar, although ST reached a slightly lower (6%) final energy .

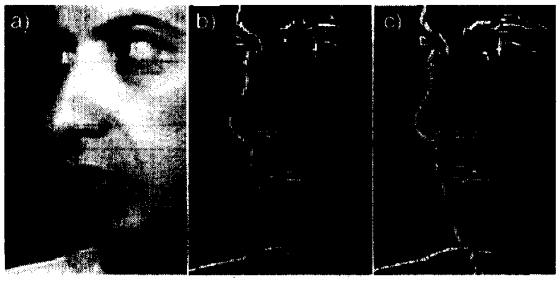

Figure 7: **(a) Original image. Edges obtained with the ST algorithm (b), and the MFA algorithm (c).**

5 Concluding remarks

We have presented a new algorithm for visual surface reconstruction capable of preserving discontinuities, which we called *simulated tearing* (ST). Formally a continuation method, ST works by slowly applying the data *force* to a *weak* surface which is allowed to tear when the tension exceeds a certain threshold. The algorithm was tested and compared with *mean field annealing* (MFA) (also a continuation method), using real and synthetic images. Although exhaustive tests (e.g. Monte Carlo) and a deeper theoretical analysis are still needed, we showed that ST and MFA provide visually similar results, with ST yielding slightly better energies. An important advantage of ST is that it is faster and simpler than MFA and it can be implemented using only integer arithmetic.

References

[1] H. C. Andrews and B. R. Hunt. *Digital Image Restoration.* Prentice Hall, New Jersey, 1977.

[2] M. Bertero, T. Poggio, and V. Torre. "Ill-posed problems in early vision". *Proceedings of the IEEE*, vol. 76(8):869–889, August 1988.

[3] J. Besag. "On the statistical analysis of dirty pictures". *Journal of the Royal Statistical Society B*, vol. 48, pp. 259–302, 1986.

[4] A. Blake. "The least disturbance principle and weak constraints". *Pattern Recognition Letters*, vol.1, pp. 393–399, 1983.

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

[6] M. Figueiredo and J. Leitão. "Image restoration using neural networks". In *Proc. of ICASSP'92*, vol.II, pp. 409–412, S. Francisco, 1992.

[7] M. Figueiredo and J. Leitão. "Image restoration using a standard Hopfield neural network". In *Proceedings of Image Processing Theory and Applications -IPTA'93*, S. Remo, June 1993.

[8] D. Geiger and F. Girosi. "Parallel and deterministic algorithms from MRF's: Surface reconstruction". *IEEE Trans. on Patt. Anal. and Mach. Intell.*, vol. PAMI-13, pp. 401–412, 1991.

[9] S. Geman and D. Geman. "Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images". *IEEE Trans. on Patt. Anal. and Mach. Intell.*, vol. PAMI-6, pp. 721–741, 1984.

[10] Y. Leclerc. "Constructing simple stable descriptions for image partitioning". *International Journal of Computer Vision*, vol. 3, pp. 73–102, 1989.

[11] J. Marroquin, S. Mitter, and T. Poggio. "Probabilistic solution of ill-posed problems in computacional vision". *Journal of the American Statistical Association*, vol. 82, pp. 76–89, 1987.

[12] J. Moura and N. Balram. "Recursive structure of noncausal Gauss-Markov random fields". *IEEE Trans. on Inf. Th.*, vol. IT-38, pp. 334–354, 1992.

[13] D. Mumford and J. Shah. "Boundary detection by minimizing functionals". In *Proc. of CVPR'85*, pp. 22–26, 1985.

[14] P. Perona and J. Malik. "Scale space and edge detection using anisotropic diffusion". In *Proc. of the IEEE Computer Soc. Workshop on Computer Vision*, pages 16–22, Miami, 1987.

[15] J. Shah. "Segmentation by nonlinear diffusion". In *Proc. of CVPR'91*, pp. 202–207, 1991.

[16] D. Terzopoulos. "Regularization of inverse visual problems involving discontinuities". *IEEE Trans. on Patt. Anal. and Mach. Intell.*, vol. PAMI-8, pp. 413–424, 1986.

[17] A. Tikhonov, A. Goncharsky, and V. Stepanov. "Inverse problems in image processing". In A. Tikhonov and A. Goncharsky, editors, *Ill-Posed Problems in the Natural Sciences*, pp. 220–232. Mir Publishers, Moscow, 1987.

[18] J. Zerubia and R. Chellappa. "Mean field approximation using compound Gauss-Markov random field for edge detection and image restoration". In *Proceedings of ICASSP'90*, Albuquerque, 1990.