MAP Image Restoration and Segmentation By Constrained Optimization

Stan Z. Li

Abstract

The combinatorial optimization problem of MAP estimation is converted to one of constrained real optimization and then solved by using the proposed augmented Lagrange-Hopfield (ALH) method. The ALH effectively overcomes instabilities that are inherent in the penalty method or the Lagrange multiplier method in constrained optimization. It produces good solutions with reasonable costs.

I. Introduction

The aim of image restoration is to recover a degraded image and that of image segmentation is to partition an image into regions of similar image properties. Efficient restoration and segmentation are very important for numerous image analysis applications. Both problems can be posed generally as one of image estimation where the underlying image or segmentation map is to be estimated from the degraded image. Due to various uncertainties, an optimal solution is sought. A popular optimality criterion is the maximum *a posteriori* (MAP) probability principle in which both the prior distribution of the true image class and the conditional (likelihood) distribution of the data are taken into account. Contextual constraints, *i.e.* constraints between pixels, are important in image analysis. Markov random fields (MRFs) or equivalently Gibbs distributions provide a convenient tool for modeling prior distributions of images which encode contextual constraints. Maximizing the posterior is equivalent to minimizing the energy function in the Gibbs distribution. The MAP principle and MRF together form the MAP-MRF framework [1], [2].

When the pixels of the image to be recovered take discrete values, as is the case dealt with in this paper, the minimization is combinatorial. Discrete optimization methods often used in statistical image analysis include

Stan Z. Li is with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore 639798. E-mail: szli@szli.eee.ntu.edu.sg . URL: http://markov.eee.ntu.edu.sg/~szli .

the iterative conditional modes (ICM) [3] and simulated annealing (SA) [4], [1]. Other discrete algorithms include the highest confidence first (HCF) [5]. The deterministic algorithms of ICM and HCF quickly converge to a local energy minimum but are dependent largely on the initial configuration. The stochastic SA with a slow enough schedule finds a global solution with probability approaching one but is well-known to be expensive.

A combinatorial optimization can often be converted into a constrained real optimization with equality and inequality constraints. The penalty and the Lagrange multiplier methods can be used for coping with equality constraints and the barrier method for coping with inequality constraints. However, the penalty method suffers from the ill-conditioning and the Lagrange method suffers from the zigzagging problem [6]. The augmented Lagrange (AL) method [7] combines both the Lagrange and the penalty methods and effectively overcomes the associated problems. In AL, the relative weight for the penalty terms need not be infinitely large, this not only overcoming the ill-conditioning problem but also beneficial for obtaining better quality solutions because the relative importance of the original objective function is more emphasized; on the other hand, its use of quadratic penalty terms "convexifies" and hence stabilizes the system, overcoming the zigzagging problem [6].

Mean field annealing (MFA) [8] provides still another continuous method. Assuming that the minima of the original energy and the corresponding mean field effective energy coincide, the MFA aims to approximate the global minimum of the original energy by tracking that of the effective energy with decreasing temperature. A recent analysis shows that the effective energy of MFA is identical to a combination of the original energy, a particular barrier term and a standard Lagrange term [9].

In this paper, we present another deterministic method, called the augmented Lagrange-Hopfield (ALH) method, for the combinatorial optimization in the MAP-MRF image restoration and segmentation. In solving the converted constrained real optimization, the ALH method uses the augmented Lagrangian multiplier method [7] to satisfy the equality constraints and the Hopfield network encoding [10] to impose the inequality constraints. The use of AL effectively overcomes instabilities inherent in the penalty method and the Lagrange multiplier method. The resulting algorithm solves a system of differential equations. Experimental results in both image restoration and segmentation are shown to compare the ALH method with the ICM, HCF and SA. The results show that the ALH outperforms ICM and HCF, and is comparable to SA in terms of the solution quality; it quickly yields a good solution after a dozen of iterations, a number similar to that required by ICM and HCF but much smaller than SA. A discussion on MFA results is also provided.

The rest of the paper is organized as follows: Section 2 describes the ALH method after introducing the MAP-MRF formulation. Section 3 presents the experimental comparisons. Conclusions are given in Section 4.

II. The Augmented Lagrange-Hopfield Method

A. MAP-MRF Image Restoration and Segmentation

The underlying image signal is denoted as $f = \{f_i \mid i \in \mathcal{S}\}$ where $\mathcal{S} = \{1, \ldots, m\}$ indexes the set of sites corresponding to image pixels. Each pixel takes on a discrete value f_i in the label set $\mathcal{L} = \{1, \ldots, M\}$; f is an underlying image to be restored or a segmented map to be computed. The spatial relationship of the sites, each of which is indexed by a single number in $\mathcal{S}$, is determined by a neighborhood system $\mathcal{N} = \{\mathcal{N}_i \mid i \in \mathcal{S}\}$ where $\mathcal{N}_i$ is the set of sites neighboring i. A single site or a set of neighboring sites form a clique denoted by c. In this paper, only up to pair-site cliques defined on the 8-neighborhood system are considered.

The underlying image f can consists of blob-like regions or a texture pattern. Different types are due to different ways that pixels interact with each other, *i.e.* due to different contextual interactions. Such contextual interactions can be modeled as MRFs or Gibbs distributions of the form $P(f) = Z^{-1} \times \mathrm{e}^{-\sum_{c \in \mathcal{C}} V_c(f)}$ where $V_c(f)$ is the potential function for clique c, $\mathcal{C}$ is the set of all cliques, and Z is the normalizing constant. Among various MRFs, the multi-level logistic (MLL) model is a simple and yet powerful mechanism for encoding a large class of spatial patterns such as textured or non-textured images. In MLL, the pair-site clique potentials take the form: $V_2(f_i, f_{i'}) = \beta_c$ if sites in clique $\{i, i'\} = c$ have the same label or $V_2(f_i, f_{i'}) = -\beta_c$ otherwise, where β_c is a parameter for type-c cliques; the single site potentials are defined by $V_1(f_i) = \alpha_I$ where α_I is the potential for the label $I = f_i$. When the true pixel values are contaminated by identical independently distributed (i.i.d.) Gaussian noise, the observed data, or the likelihood model, is $d_i = f_i + e_i$ where $e_i \sim N(0, \sigma^2)$ is the zero-mean Gaussian distribution with standard deviation σ. With these prior and likelihood models, the energy in the posterior distribution $P(f \mid d) \propto \mathrm{e}^{-E(f)}$ is

$$E(f) = \sum_{i \in \mathcal{S}} (f_i - d_i)^2 / [2\sigma^2] + \sum_{\{i\} \in \mathcal{C}} V_1(f_i) + \sum_{\{i,i'\} \in \mathcal{C}} V_2(f_i, f_{i'}) \tag{1}$$

In this paper, we assume that the MRF and noise parameters α, β and σ are known. The MAP estimate for the restoration or segmentation is defined as $f^* = \arg\min_{f \in \mathcal{L}^m} E(f)$. The minimization of the energy function $E(f)$ is in the discrete space $\mathcal{L}^m$ and hence combinatorial.

B. MAP-MRF Estimation as Constrained Optimization

The original combinatorial optimization is converted into a constrained optimization in a real space, using the notion of continuous relaxation labeling can be used for this. Let real value $p_i(I) \in [0, 1]$ represent the strength with which label I is assigned to i, the M-element vector $p_i = [p_i(I) \mid I \in \mathcal{L}]$ represent the state of the assignment for $i \in \mathcal{S}$, and the matrix $p = [p_i(I) \mid i \in \mathcal{S}, I \in \mathcal{L}]$ the state of the labeling assignment. The energy with the p variables is given by

$$E(p) \quad = \quad \sum_{i \in \mathcal{S}} \sum_{I \in \mathcal{L}} r_i(I)\ p_i(I) + \sum_{i \in \mathcal{S}} \sum_{I \in \mathcal{L}} \sum_{i' \in \mathcal{S}, i' \neq i} \sum_{I' \in \mathcal{L}} r_{i,i'}(I, I')\ p_i(I)\ p_{i'}(I') \tag{2}$$

where $I = f_i$ and $I' = f_{i'}$, $r_i(I) = V_1(I \mid d) = (I - d_i)^2/2\sigma + V_1(I)$ is the single-site clique potential function in the posterior distribution $P(f \mid d)$, and $r_{i,i'}(I, I') = V_2(I, I' \mid d) = V_2(I, I')$ is the pair-site clique potential function in $P(f \mid d)$.

With such a representation, the combinatorial minimization is reformulated as the following constrained minimization

$$\min_p \quad E(p) \tag{3}$$

$$\text{subject to} \quad C_i(p) = \sum_I p_i(I) - 1 = 0 \quad i \in \mathcal{S} \tag{4}$$

$$p_i(I) \geq 0 \qquad \forall i \in \mathcal{S}, \forall I \in \mathcal{L} \tag{5}$$

The final solution p^* is subject to additional ambiguity constraints: $p_i^*(I) \in \{0, 1\}$.

C. The ALH Method

The ALH method aims to solve the above constrained minimization problem. It uses the augmented Lagrange technique to satisfy the equality constraints of (4) and the Hopfield encoding to impose the inequality constraints of (5). In the Hopfield encoding, internal variables $u_i(I) \in (-\infty, +\infty)$ are introduced and related to $p_i(I)$ via the sigmoid function

$$p_i(I) = \psi_T(u_i(I)) = \frac{1}{1 + \mathrm{e}^{-u_i(I)/T}} \tag{6}$$

controlled by a temperature parameter $T > 0$. This treatment confines $p_i(I)$ to the range $(0, 1)$ so as to impose the inequality constraints. When $T \to 0^+$, $p_i(I)$ is forced to be 0 or 1 depending on whether $u_i(I)$ is positive or negative. Thereby, the ambiguity constraints $p_i^*(I) \in \{0, 1\}$ are imposed. With the free u variables, the energy function can be considered as $E(u) = E(p(u))$.

The sigmoid function of Hopfield is one of many ways for imposing the equality constraints. Yuille and Kosowshy [9] have used a barrier term of the form $T \sum_{i,I} p_i(I) \log p_i(I)$ for imposing $p_i(I) > 0$. However, care must be taken in the numerical implementation of a gradient descent algorithm involving the barrier term in order to avoid $p_i(I)$ from decreasing to non-positive values. A truncation operation $\min\{0, p_i(I)\}$ may be used to truncate negative $p_i(I)$ values to 0. But this inevitably causes oscillations and hence non-convergence. We find it easier and more stable to use the sigmoid function in that the u variables can be updated freely.

The augmented Lagrange (AL) function [7] takes the form

$$L_\beta(p, \gamma) = E(p) + \sum_{k \in \mathcal{S}} \gamma_k C_k(p) + \frac{\beta}{2} \sum_{k \in \mathcal{S}} [C_k(p)]^2 \tag{7}$$

where γ_k are the Lagrange multipliers and $\beta > 0$ is the weight for the quadratic penalty term. It is a function of the $M \times m$ variables of p and the m variables of γ. When $\gamma_i = 0$ for all i, the AL reduces to the penalty method $L_\beta(p) = E(p) + \frac{\beta}{2} \sum_k [C_k(p)]^2$. When $\beta = 0$, it reduces to the standard Lagrange $L(p, \gamma) = E(p) + \sum_k \gamma_k C_k(p)$. With the use of Lagrange multipliers, the weight β for the penalty terms need not be infinitely large and this overcomes the ill-conditioning problem in the penalty method. Smaller β is also beneficial for better quality solution because the original objective function E is more emphasized. The use of quadratic penalty terms "convexifies" the system and hence stabilizes it, and this overcomes the zigzagging problem in the standard Lagrange.

When (p^*, γ^*) is a saddle point for which $L_\beta(p^*, \gamma) \le L_\beta(p^*, \gamma^*) \le L_\beta(p, \gamma^*)$, then p^* is a local minimum of $E(p)$ satisfying (4) [11]. The dynamic equations for minimizing L_β are

$$\frac{\mathrm{d}p_i(I)}{\mathrm{d}t} = -\frac{\partial L_\beta(p,\gamma)}{\partial p_i(I)} = -\frac{\partial E(p)}{\partial p_i(I)} - \sum_k \gamma_k \frac{\partial C_k(p)}{\partial p_i(I)} - \beta \sum_k C_k(p) \frac{\partial C_k(p)}{\partial p_i(I)} \tag{8}$$

$$\frac{\mathrm{d}\gamma_i}{\mathrm{d}t} = +\frac{\partial L_\beta(p,\gamma)}{\partial \gamma_i} = +C_i(p) \tag{9}$$

where

$$\frac{\partial C_k(p)}{\partial p_i(I)} = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{otherwise} \end{cases} \tag{10}$$

for $C_i(p)$ defined in (4). The system performs energy descent on p but ascent on γ.

In the ALH method, the updating of label assignment is performed on u, rather than on p variables. Equation (8) is replaced by

$$\frac{\mathrm{d}u_i(I)}{\mathrm{d}t} = -\frac{\partial L_\beta(p(u),\gamma)}{\partial p_i(I)} = -\left[q_i(I) + \gamma_i + \beta C_i(p)\right] \tag{11}$$

where $q_i(I) = \frac{\partial E}{\partial p_i(I)}$. Because $\frac{\partial p_i(I)}{\partial u_i(I)} = \frac{\mathrm{e}^{-u_i(I)/T}}{T(1+\mathrm{e}^{-u_i(I)/T})^2}$ is always positive, $\frac{\mathrm{d}p_i(I)}{\mathrm{d}t} = \frac{\partial p_i(I)}{\partial u_i(I)} \frac{\mathrm{d}u_i(I)}{\mathrm{d}t}$ has the same sign as $\frac{\mathrm{d}u_i(I)}{\mathrm{d}t}$.

Numerically, the ALH algorithm consists of the following three equations corresponding to (11), (6) and (9):

$$u_i^{(t+1)}(I) \leftarrow u_i^{(t)}(I) - \mu \left\{ q_i^{(t)}(I) + \gamma_i^{(t)} + \beta C_i(p^{(t)}) \right\} \tag{12}$$

$$p_i^{(t+1)}(I) \leftarrow \psi_T(u_i^{(t+1)}(I)) \tag{13}$$

$$\gamma_i^{(t+1)} \leftarrow \gamma_i^{(t)} + \beta C_i(p^{(t)}) \tag{14}$$

In the above, μ is the step size; during the update, β is increased to speed up convergence. In practice, T needs not be very low to impose the ambiguity constraints on the final solution p^*. The competitive mechanism in (4) will make the winner take all. We set $T = 10^5$ in our experiments and obtained good convergence. In our implementation, the updating is performed synchronously in parallel for all i and I.

III. Experimental Results

In the following, we present two experiments, one for MAP-MRF image restoration and the other for segmentation, to compare the performance of the following algorithms: (1) ALH of this paper, (2) ICM [3], (3) HCF [5] (the parallel version), (4) MFA [8], [9] (See Appendix for the algorithm), and (5) SA with the

Metropolis sampler [4] implemented based on a procedure given in [12]. The comparison is in terms of (i) the solution quality measured by the minimized energy values and (ii) the convergence rate measured by the number of iterations. In calculating the energy, $E(f)$ of (1) is used where for the continuous algorithms of ALH and MFA, f is obtained from p by a maximum selection (winner-take-all) operation.

For ALH, $\mu = 100$ and $T = 10^5$ are fixed, β is increased from 1 to 100 according to $\beta \leftarrow 1.01\beta$. The convergence criterion is $\|u^{(t)} - u^{(t-1)}\|_\infty < 0.0001$. The schedule for SA is $T^{(t+1)} \leftarrow 0.999T^{(t)}$ (with $T^{(0)} = 10^4$) and for MFA is $T^{(t+1)} \leftarrow 0.99T^{(t)}$ (with $T^{(0)} = 10^3$).

The first set of experiments is for MAP-MRF restoration performed on three synthetic images of $M = 4$ gray levels and Figures 1–3 show the results. In each figure, (a) is the true image with $M = 4$ gray levels, the label set $\mathcal{L} = \{1, 2, 3, 4\}$, and the pixel gray values also in $\{1, 2, 3, 4\}$. The clique potential parameters α_I and $(\beta_1, \cdots, \beta_4)$ for generating the three images are shown in Table I. (b) is the observed image in which every pixel takes a real value which is the true pixel value plus zero-mean i.i.d. Gaussian noise with standard deviation $\sigma = 1$. (c) is the maximum likelihood estimate which was used as the initial labeling. (d) to (h) are the solutions found by the compared algorithms.

Table II shows the minimized energy values, the error rates, and the required iteration numbers. It can be seen from the table that objectively, ALH performs the best out of the three deterministic algorithms in terms of both the minimized energy values. Overall, the solution quality is ranked as "ICM < HCF < MFA < ALH < SA", which is in agreement with a subjective evaluation of the results. We also implemented a parallel ICM using codings but the results were not as good as the serial ICM.

The second experiment compares the algorithms in performing MAP-MRF segmentation on the Lenna image of size 256×240 into a tri-level segmentation map. The results are illustrated in Figure 4. The input image (a) is the original Lenna image corrupted by the i.i.d. Gaussian noise with standard deviation 10. The observation model is a Gaussian distribution of standard deviation 10 with mean values 40, 125 and 200 for the three-level segmentation. An isometric MRF prior is used with the four β parameters being $(-1, -1, -1, -1)$. (b) is the maximum likelihood estimate which was used as the initial segmentation. (c)–(f) are the segmentation results of ALH, ICM, MFA and SA, respectively. Table III illustrates the minimized energy, *i.e.* maximized posterior probability, values and the iteration numbers numerically. The solution quality is ranked as "HCF < ICM < SA < ALH < MFA".

According to the iteration numbers, the ALH method takes much fewer iterations than SA, about 0.2% – 4.4%, to converge. Although taking more iterations than the other deterministic algorithms, it is not so slow as it might seem. That the ALH converged after thousands of iterations was due to the stringent convergence criterion $\|u^{(t)} - u^{(t-1)}\|_\infty < 0.0001$. However, looking at the energy evolution curve in Figure 5 (SA is not included there because it is far from the convergence within a thousand of iterations), we see that the ALH reached a solution better than ICM and HCF after only a dozen of iterations and becomes nearly convergent after a hundred of iterations. The MFA needs five hundred of iterations to converge.

IV. Conclusions

The advantages of the augmented Lagrange-Hopfield (ALH) method are due to the use of the augmented Lagrange method. The ALH method not only overcomes the ill-conditioning problem of the penalty method and the zigzagging problem of the standard Lagrangian but also improves convergence. With the augmented Lagrangian, the weights for the penalty terms can be much smaller than those required by the penalty method and the relative weight for the original objective can be increased. This helps yielding lower minimized energy values. The ALH method as a general method has also been used for solving other combinatorial optimization problems. For example in solving traveling salesman problem, the ALH finds remarkably better solutions than Hopfield type neural networks [**?**].

Appendix

The MFA algorithm [8] approximates the global minimum of $E(f)$ by evaluating the partition function at the global saddle point of the effective energy. The effective energy is derived in [8] as

$$E_{eff}(p,q) = E(p) + T\left[\sum_i \sum_I p_i(I) q_i(I) + \sum_i \ln \sum_I \mathrm{e}^{q_i(I)}\right] \tag{15}$$

The saddle points are among the roots of the following mean field theory equations

$$\nabla_p E_{eff}(p,q) = 0, \qquad \nabla_q E_{eff}(p,q) = 0 \tag{16}$$

where ∇_x is the gradient with respect to x. Adding to the effective energy a term $-\frac{\beta}{2}\sum_i \sum_I p_i^2(I)$ has a constructive balancing effect in solving the mean field equations [8] and when β is high enough, the minima of $E(p) - \frac{\beta}{2}\sum_i \sum_I p_i^2(I)$ always lies at the vertices of the hypercube [9]. With the additional term, we have the following the mean field theory equations

$$q_i(I) = -\frac{1}{T}\left[\frac{\partial E(p)}{\partial p_i(I)} - \beta p_i(I)\right], \qquad p_i(I) = \frac{\mathrm{e}^{q_i(I)}}{\sum_J \mathrm{e}^{q_i(J)}} \tag{17}$$

Schemes for selecting β values are provided in [8], [9], which require knowing the eigenvalues of the Hessian of $E(p)$. We set $\beta = 20$ by trial and error; a smaller value would not essure the convergence for some cases. Our MFA results were obtained with a synchronous updating procedure.

Acknowledgement This work was supported by NTU AcRP projects RG 43/95 and RG 51/97. The author would also like to thank the reviewers for their comments and suggestions.

References

[1] S. Geman and D. Geman, ““Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images””, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, pp. 721–741, November 1984.

[2] S. Z. Li, *Markov Random Field Modeling in Computer Vision*, Springer-Verlag, New York, 1995.

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

[4] S. Kirkpatrick, C. D. Gellatt, and M. P. Vecchi, ""Optimization by simulated annealing"", *Science*, vol. 220, pp. 671–680, 1983.

[5] P. B. Chou, P. R. Cooper, M. J. Swain, C. M. Brown, and L. E. Wixson, ""Probabilistic network inference for cooperative high and low level vision"", in R. Chellappa and Anil Jain, editors, *Markov Random Fields: Theory and Applications*, pp. 211–243, Boston, 1993. Academic Press.

[6] R. Fletcher, *Practical Methods of Optimization*, Wiley, 1987.

[7] M. J. D. Powell, ""A method of nonlinear constraints in minimization problems"", in R. Fletcher, editor, *Optimization*, London, 1969. Academic Press.

[8] C. Peterson and B. Soderberg, ""A new method for mapping optimization problems onto neural networks"", *International Journal of Neural Systems*, vol. 1, pp. 3–22, 1989.

[9] A. L. Yullie and J. J. Kosowsky, ""Statistical physics algorithms that converge"", *Neural Computation*, vol. 6, pp. 341–356, 1994.

[10] J. J. Hopfield, ""Neurons with graded response have collective computational properties like those of two state neurons"", *Proceedings of National Academic Science, USA*, vol. 81, pp. 3088–3092, 1984.

[11] B. S. Gottfried, *Introduction to Optimization Theory*, Prentice-Hall, 1973.

[12] William H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical recipes in C*, Cambridge University Press, 2 edition, 1988.

	σ	α_I	β_1	β_2	β_3	β_4
Image No.1	1	0	-1	-1	-1	1
Image No.2	1	0	-2	-2	1	1
Image No.3	1	0	1	1	1	1

TABLE I

THE MRF PARAMETERS (α AND β) AND NOISE PARAMETER σ FOR GENERATING THE THREE IMAGES.

	ALH	ICM	HCF	MFA	SA
Image No.1	-11049	-10003	-10049	-10824	-11988
Image No.2	-10191	-8675	-9650	-10073	-11396
Image No.3	-26974	-25881	-26629	-26716	-27526
Image No.1	0.268	0.349	0.367	0.291	0.164
Image No.2	0.360	0.438	0.414	0.381	0.306
Image No.3	0.212	0.327	0.273	0.181	0.125
Image No.1	3654	7	31	553	83034
Image No.2	1456	6	29	553	68804
Image No.3	2789	7	36	560	92721

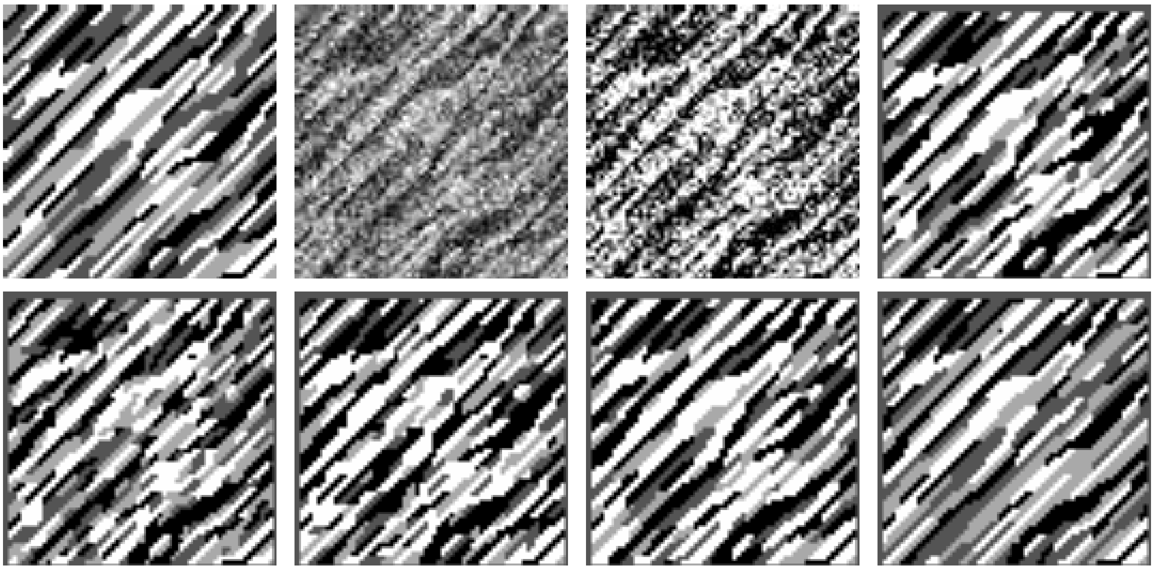
TABLE II

THE MINIMIZED ENERGY VALUES (TOP BLOCK), ERROR RATES (MIDDLE BLOCK), AND ITERATION NUMBERS (BOTTOM BLOCK) FOR IMAGE NOS 1–3.

	ALH	ICM	HCF	MFA	SA
Min. Energy	-180333	-171806	-176167	-180617	-173301
Iterations	1255	7	38	545	593916

TABLE III

NUMERICAL COMPARISON OF THE ALGORITHMS ON THE SEGMENTATION OF THE LENNA IMAGE.



(a)	(b)	(c)	(d)
(e)	(f)	(g)	(h)

Fig. 1. Restoration of image No.1. (a) Original image. (b) Observed noisy image. (c) Maximum likelihood estimate. (d) ALH solution. (e) ICM solution. (f) HCF solution. (g) MFA solution. (h) SA solution.

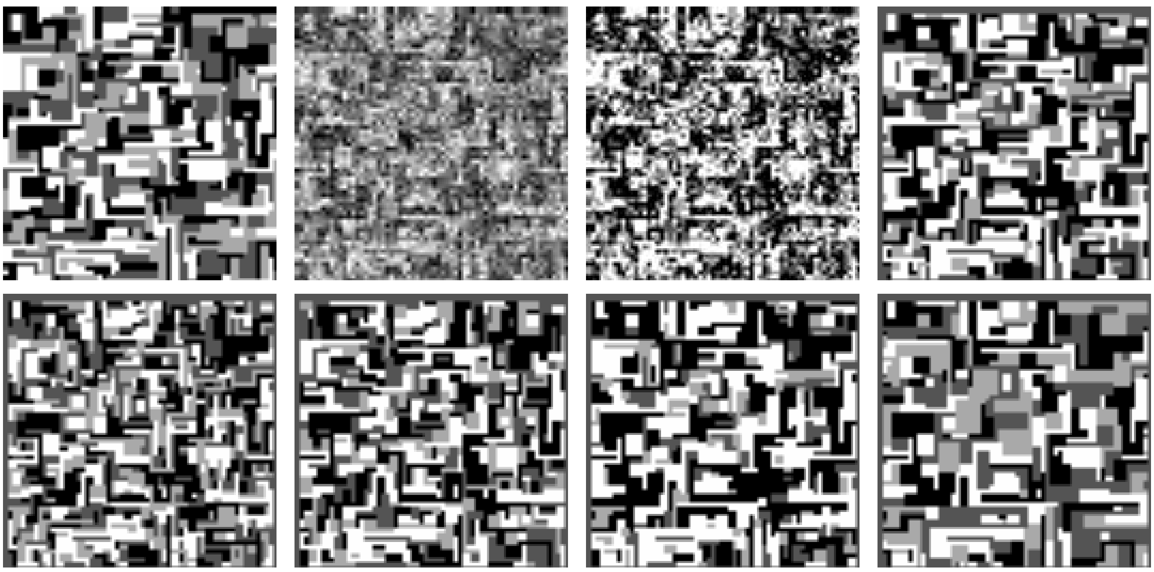


Fig. 2. Restoration of image No.2 (refer to Fig.1 for legend).

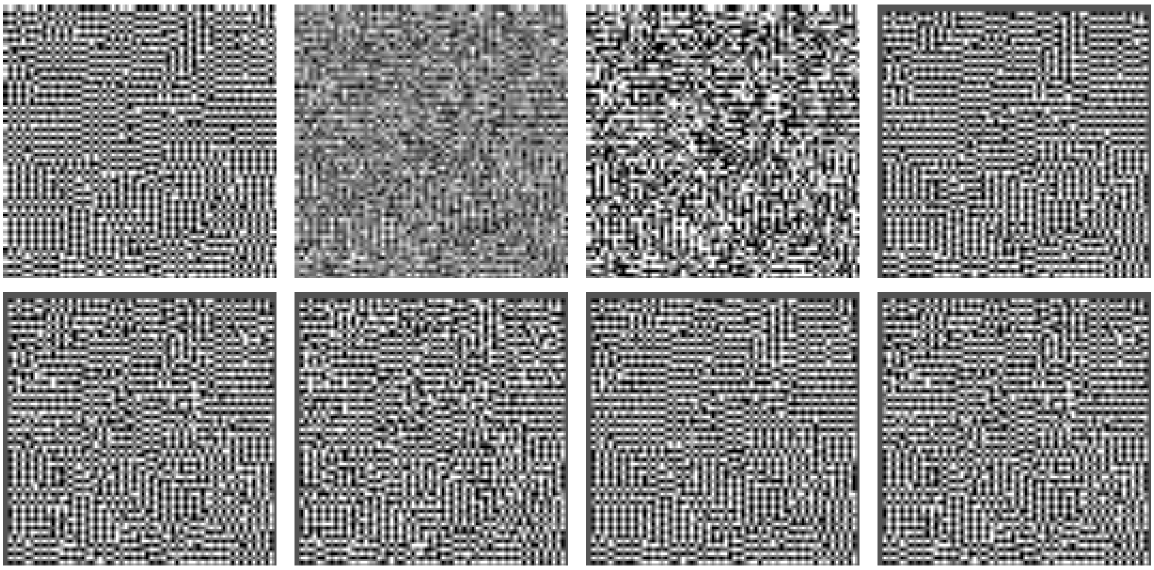


Fig. 3. Restoration of image No.3 (refer to Fig.1 for legend).

Fig. 4. Segmentation of the Lenna image. Top raw: Input image, maximum likelihood segmentation, ALH solution. Bottom raw: ICM solution, MFA solution, SA solution.

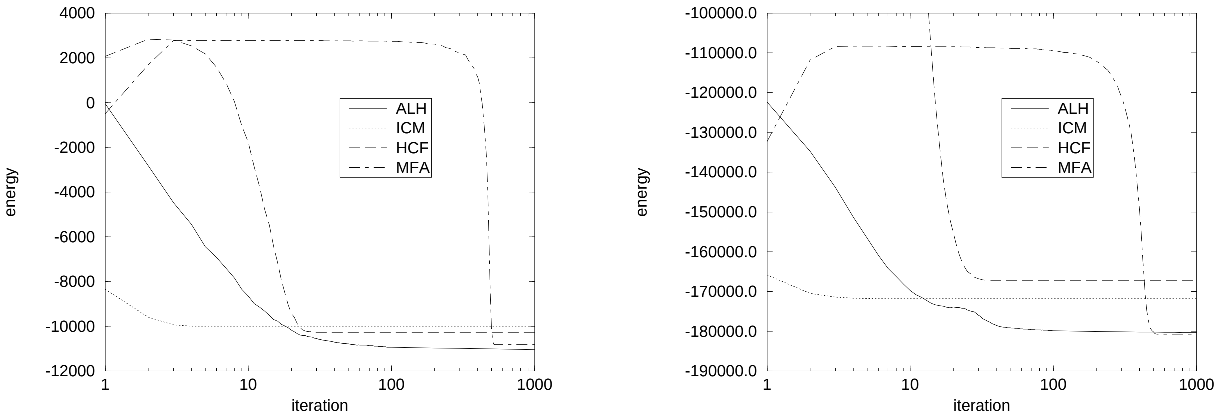


Fig. 5. The solution evolution curves for image no.1 (left) and the Lenna image.