

Approximate Optimization Algorithms in Markov Random Field Model Based on Statistical-Mechanical Techniques

著者	TANAKA Kazuyuki, MAEDA Junji			
journal or	Memoirs of the Muroran Institute of Technology			
publication title				
volume	50			
page range	9-15			
year	2000-11-30			
URL	http://hdl.handle.net/10258/130			

Approximate Optimization Algorithms in Markov Random Field Model Based on Statistical-Mechanical Techniques

Kazuyuki TANAKA* and Junji MAEDA**

(Accepted 31 August 2000)

An image restoration can be often formulated as an energy minimization problem. When an energy function is expressed by using the hamiltonian of a classical spin system only with finite range interactions, the probabilistic model, which is described in the form of Gibbs distribution for the energy function, can be regarded as a Markov random field (MRF) model. Some approximate optimization algorithms for the energy minimization problem were proposed in the standpoint of statistical-mechanics. In this paper, the approximate optimization algorithms are summarized and are applied to the image restoration for natural image.

Keywords: Soft Computing, Probabilistic Computing, Knowledge Information Processing, Statistical Method, Bayes Statistics, Gibbs Distribution, Markov Random Fields, Mean-Field Theory, Image Restoration

1 INTRODUCTION

Recently, many authors have shown an interest in image restoration using the Markov random field (MRF) model, in which the configuration of a lattice site is dependent only on the configurations of the nearest neighbours (1),(2),(3). The MRF model, can be regarded as one of classical spin systems with finite range interactions and non-uniform external fields on a finite square lattice. The MRF model can be regarded as a classical spin system in statistical mechanics. The authors have proposed a new method for systematically constructing the energy function⁽⁴⁾, which is based on constrained optimization. In the image restoration, the constraints are introduced as a priori information on the original image. By introducing a Lagrange multiplier for each of the constraints, the image restoration is reduced to an energy minimization problem^{(4),(5),(6)}. In Ref. (7), we described a classical spin system, which is applicable to gray-level image restoration.

In the search for an optimal solution of MRF model, many authors applied the iterated conditional modes (ICM) algorithm⁽⁸⁾. Though the ICM algorithm is simple algorithm and can erase noise when the noise is in an isolated site, it is difficult to avoid the local minimum and then it cannot erase noise when two or more successive sites are affected by the noise. to the search of the minimum-energy configuration. In order to avoid the local minimum, the cluster type Monte Carlo simulation⁽⁹⁾, and the cluster type mean field approximation (10) are also applied to the optimization of the MRF model. In the statistical mechanics, we have some important fluctuation effects to avoid the local minimum. One of them is thermal effect. In order to adopt the thermal fluctuation effect as annealing procedure, we introduce a temperature in the form of Gibbs distribution. Geiger and Girosi (11), and Zhang(12),(13) proposed a deterministic algorithm, which is based on the mean-field approximation. On the other hand, we have general methodology of construction of high-level effective-field approximation in statistical mechanics, which is called cluster variation method

^{*} Department of Computer and Mathematical Sciences, Graduate School of Information Sciences, Tohoku University

^{**}Department of Computer Science and Systems Engineering

(CVM)⁽¹⁴⁾. We gave a simple example of its application to the MRF model⁽¹⁵⁾. Moreover, in the standpoint of the CVM, we proposed the methodology in the extension of the ICM to the cluster type algorithm, which is called the cluster zero-temperature process (CZTP) and is obtained as the zero-temperature limit in the optimization algorithm of the CVM⁽¹⁶⁾.

In this paper, we clarify the mathematical structure of the ICM, the CZTP and MFA algorithms and compare them with each others in some numerical experiments for natural images. In Sec. 2, we explain the energy function and the probability distribution constructed by means of the mathematical framework of Ref.(7). In Sec. 3, we summarize some approximate optimization algorithms for the energy function. In Secs. 4 and 5, we give some numerical experiments and some concluding remarks, respectively.

2 ENERGY FUNCTION OF IMAGE RESTORATION

In this section, we give the energy function of image restoration for natural images with the aid of the mathematical framework in Ref.(7).

We consider a digital image with q grades on a $M \times N$ finite square lattice

$$\mathbf{L} \equiv \{(i,j) | i = 1, 2, \dots, M, j = 1, 2, \dots, N \},\$$

with the periodic boundary condition. We express the configurations of an original and a degraded images by $\mathbf{x} \equiv \{x_{i,j} | (i,j) \in \mathbf{L}\}$ and $\mathbf{y} \equiv \{y_{i,j} | (i,j) \in \mathbf{L}\}$, respectively. The variable $x_{i,j}$ on a pixel (i,j) takes a value from $\Lambda \equiv \{0,1,2,\cdots,q-1\}$. For the degradation process, we assume that a degraded image \mathbf{y} is obtained from the original image \mathbf{x} by changing the state of each pixel to another state by the same probability p, independently of the other pixel. The conditional probability for a degraded image \mathbf{y} when the original image is \mathbf{x} , $P_{yx}(\mathbf{y}|\mathbf{x})$, is given by

where

$$d(\mathbf{x}, \mathbf{y}) \equiv \sum_{(i,j)\in\mathbf{L}} (1 - \delta(x_{i,j}, y_{i,j})), \cdots (2)$$

and

$$T_p \equiv \frac{1}{\ln\left(\frac{1-qp+p}{p}\right)}.$$
 (3)

In the present paper, we treat only the case in which T_p is positive such that p < 1/q.

In Ref.(7), it is assumed that we know the following quantities for the true original image $\tilde{\mathbf{x}}$:

$$\sigma_{2}(\tilde{\mathbf{x}}) \equiv \sum_{(i,j)\in\mathbf{L}} \left(2 - \delta(\tilde{x}_{i,j}, \tilde{x}_{i+1,j}) - \delta(\tilde{x}_{i,j}, \tilde{x}_{i,j+1})\right), \cdots (4)$$

$$\sigma_{2,n}(\tilde{\mathbf{x}}) \equiv \sum_{(i,j)\in\mathbf{L}} \left(\delta(|\tilde{x}_{i,j} - \tilde{x}_{i+1,j}|, n) + \delta(|\tilde{x}_{i,j} - \tilde{x}_{i,j+1}|, n)\right),$$

$$(n = 1, 2, \dots, k-1). \cdots (5)$$

The image restoration in natural images is formulated as the following conditional optimization problem:

In order to ensure the constrained conditions $\sigma_2(\mathbf{x}) = \sigma_2(\tilde{\mathbf{x}})$ and $\sigma_{2,1}(\mathbf{x}) = \sigma_{2,1}(\tilde{\mathbf{x}})$, we introduce the Lagrange multipliers J_k and $J_k - J_n$ $(n = 1, 2, \dots, K - 1)$. The conditional optimization (6) can be reduced to the following energy minimization problem:

where

$$H(\mathbf{x}) \equiv d(\mathbf{x}, \mathbf{y}) + J_k \sigma_2(\mathbf{x}) + \sum_{n=1}^{k-1} (J_n - J_k) \sigma_{2,n}(\mathbf{x}).$$

Here the notation $\arg\min_x f(x)$ means any minimizer of a function f(x). The optimal parameters \hat{J}_n $(n=1,2,\cdots,k)$ should be determined so as to satisfy the following constraints:

$$\sigma_2(\mathbf{x}(\hat{J}_1, \hat{J}_2, \dots, \hat{J}_k)) = \sigma_2(\tilde{\mathbf{x}}), \dots \dots \dots (9)$$

$$\sigma_{2,n}(\mathbf{x}(\hat{J}_1, \hat{J}_2, \dots, \hat{J}_k)) = \sigma_{2,n}(\tilde{\mathbf{x}}),$$

$$(n = 1, 2, \dots, k - 1), \dots (10)$$

The restored image obtained by using Eqs. (7)-(10) is denoted by $\hat{\mathbf{x}} = \mathbf{x}(\hat{J}_1, \hat{J}_2, \dots, \hat{J}_k)$. The energy minimization problem (7) is equivalent to the following probability maximization problem:

where

$$\rho(\mathbf{x}) \equiv \frac{\exp\left(-\frac{1}{T}H(\mathbf{x})\right)}{\sum_{\mathbf{x}\in\Lambda^{MN}}\exp\left(-\frac{1}{T}H(\mathbf{x})\right)},$$

$$(T>0), \dots \dots (12)$$

and T is a temperature. The probabilistic model described by the probability distribution (12) with Eq.(8) can be regarded as the MRF model.

3 APPROXIMATE OPTIMIZATION ALGORITHMS

In this section, we give three approximate optimization algorithms for the energy minimization problem (7), such that ICM⁽⁸⁾, CZTP⁽¹⁶⁾ and MFA⁽¹¹⁾. Both the ICM and the CZTP are iterative algorithms at zero temperature and the MFA is one at a finite temperature. The approximate minimum-energy configuration obtained by means of an annealing procedure is closer to the true minimum-energy configuration than by means of the ICM.

First we explain the iterative algorithm at zero temperature. The optimization problem (7) can be reduced to the following iterative equation:

$$x_{i,j}(J_1, J_2, \cdots, J_k)$$

$$= \arg\min_{x_{i,j} \in \Lambda} H\left(x_{i,j} \middle| x_{i',j'} = x_{i',j'}(J_1, J_2, \cdots, J_k), (i',j') \in \mathbf{L} \setminus (i,j)\right), \quad \cdots \quad (13)$$

From the iterative equation, we can construct the ICM algorithm⁽⁸⁾, which is the simplest algorithms for the optimization (7). Morita and the present author extend it to the cluster version ICM algorithm, which is constructed from the following iterative equation:

$$\mathbf{x_{c}}(J_{1}, J_{2}, \dots, J_{k})$$

$$= \arg \min_{\mathbf{x_{c}} \in \Lambda^{|c|}} H\left(\mathbf{x_{c}} \middle| x_{i',j'} = x_{i',j'}(J_{1}, J_{2}, \dots, J_{k}), (i',j') \in \mathbf{L} \setminus \mathbf{c}\right), \dots \dots (14)$$

where

$$\mathbf{x_c} \equiv \left\{ x_{i,j} \middle| (i,j) \in \mathbf{c} \right\},$$

$$\mathbf{x_c}(J_1, J_2, \dots, J_k) \equiv \left\{ x_{i,j}(J_1, J_2, \dots, J_k) \middle| (i,j) \in \mathbf{c} \right\}.$$

Morita and the present author called it CZTP⁽¹⁶⁾. Here **c** is a set of pixels, and the square CZTP can be constructed by setting $\mathbf{c} = \{(i, j), (i+1, j), (i+1, j+1), (i, j+1)\}$.

Second, we explain the MFA algorithm which is the most familiar deterministic annealing algorithm for the optimization (7). The annealing algorithm is able to avoid local minima. In order to adopt the MFA algorithm for the search of minimum-energy configuration $\mathbf{x}(J_1, J_2, \dots, J_k)$, we introduce the one-body marginal probability distributions:

$$\rho_{i,j}(n) \equiv \sum_{\mathbf{x} \in \Lambda^{MN}} \rho(\mathbf{x}) \delta(x_{i,j}, n),$$

$$(n \in \Lambda, (i, i) \in \mathbf{L}), \dots \dots (15)$$

In the mean-field approximation, the probability distribution $\rho(\mathbf{x})$ is approximately expressed as

$$\rho(\mathbf{x}) \simeq \prod_{(i,j)\in \mathbf{L}} \rho_{i,j}(x_{i,j}). \dots \dots \dots (16)$$

By substituting Eq.(16) into the free energy

$$\mathcal{F}[\rho] \equiv \sum_{\mathbf{x}} \rho(\mathbf{x}) \Big(H(\mathbf{x}) + T \ln(\rho(\mathbf{x})) \Big),$$
.....(17)

and by taking the first variation of the free energy $\mathcal{F}[\{\rho_{i,j}\}]$ with respect to $\rho_{i,j}(n)$, the deterministic meanfield equations for the set of one-body marginal distribution functions

$$\left\{ \rho_{i,j}(n) \middle| (i,j){\in}\mathbf{L},\; n{\in}\Lambda \right\}$$

are obtained as follows [3]:

$$x_{i,j}(J_1, J_2, \cdots, J_k) = \arg \max_{n \in \Lambda} \rho_{i,j}(n), \cdots (18)$$

where

$$\rho_{i,j}(n) = \frac{\exp\left(-\frac{1}{T}H_{i,j}(n)\right)}{\sum_{m \in \Lambda} \exp\left(-\frac{1}{T}H_{i,j}(m)\right)},$$
.....(19)

By solving Eqs.(18)-(21) at a sufficiently small positive value of T by using the annealing procedure, we obtain the approximate optimal solution of Eq.(7).

4 NUMERICAL EXPERIMENTS

In this section, we give some numerical experiments for the original image $\tilde{\mathbf{x}} = \{\tilde{x}_{i,j}\}$ given in Fig.1. The degradation process is subject to the probability given in Eq.(1). Here, we set (q-1)p=0.1, 0.3 and 0.5 where q=8. The degraded images \mathbf{y} obtained from the original image $\tilde{\mathbf{x}}$ in Fig.1 are given in Fig.2. The quantities

$$\bar{\sigma}_{2,n}(\tilde{\mathbf{x}}) \equiv \frac{1}{2MN} \sigma_{2,n}(\tilde{\mathbf{x}}) \ (n=0,1,\cdots,q-1),$$

are shown in Table 1 where $\sigma_{2,0}(\tilde{\mathbf{x}}) \equiv 1 - \sigma_2(\tilde{\mathbf{x}})$. We see that the quantities $\sigma_{2,0}(\tilde{\mathbf{x}})$ and $\sigma_{2,1}(\tilde{\mathbf{x}})$ are especially important in images with 8 grades. In the energy function (8), we set k=2. In Table 2, we give the values of optimal parameters \hat{J}_1 and \hat{J}_2 , and of quantities

$$egin{aligned} ar{d}(\hat{\mathbf{x}}, \tilde{\mathbf{x}}) &\equiv & rac{1}{MN} d(\hat{\mathbf{x}}, \tilde{\mathbf{x}}), \\ ar{\sigma}_2(\hat{\mathbf{x}}) &\equiv & rac{1}{2MN} \sigma_2(\hat{\mathbf{x}}), \end{aligned}$$

and

$$\bar{\sigma}_{2,1}(\hat{\mathbf{x}}) \equiv \frac{1}{2MN}\sigma_{2,1}(\hat{\mathbf{x}}),$$

in the restored image $\hat{\mathbf{x}}$, which is obtained by using the ICM, the CZTP and the MFA. The obtained restored images $\hat{\mathbf{x}}$ are shown in Figs.3, 4 and 5. In the CZTP, we set

$$\mathbf{c} \equiv \left\{ (i,j), (i+1,j), (i+1,j+1), (i,j+1) \right\}.$$

5 CONCLUDING REMARKS

In this paper, we summarize some approximate optimization algorithms based on the statistical-mechanical techniques for the energy minimization problem formulated for the image restoration in Ref.(7). In some numerical experiments, we show that the algorithms are applicable to the image restoration of natural image. We remark that the ICM algorithm can erase when the noise is in an isolated pixel, but it cannot erase when two or more successive pixels are affected by the noise. On the other hand, the MFA and the CZTP algorithms can deal with this problem. However the MFA need a large memory. If we apply it to the image restoration of gray-level image with 256 grades, we have to treat a $256 \times 256 \times 256$ dimensions for the one-body marginal probability distribution $\rho_{i,j}(n)$. The ICM and CZTP algorithms do not need so many memory because most of the data stored are in integer. The present author are still studying the problem of finding a computer algorithm giving the same results in much shorter time and in much smaller memory.

ACKNOWLEDGEMENT

The author is thankful to Professor T. Morita of Department of Computer Science, Nihon University in Japan, and Professor T. Horiguchi of Graduate School of Information Sciences, Tohoku University in Japan for some valuable advises on the manuscript. This work was partly supported by the Grants-In-Aid (No.10480079 and No.10780159) for Scientific Research from the Ministry of Education, Science, Sports and Culture of Japan. The author was a visiting researcher, visiting Professor D. M. Titterington of the Department of Statistics, University of Glasgow in UK, as an abroad-researcher of the Ministry of Education, Science, Sports and Culture of Japan, during the period from March, 1997 to February, 1998. The author is grateful to Professor D. M. Titterington, hospitality and some comments on this work.

REFERENCES

- R. Chellappa and A. Jain (eds), Markov Random Fields: Theory and Applications, Academic Press, New York, 1993.
- (2) H. Derin, H. Elliott, R. Cristi and D. Geman, "Bayes smoothing algorithms for segmentation of binary images modeled by Markov random fields," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.6, pp.707-720, 1984.
- (3) K. Tanaka, "Statistical-mechanical method in image restoration," *BUTSURI*, vol.54, no.1, pp.25-33, Jan. 1999 (in Japanese).
- (4) T. Morita and K. Tanaka, "Determination of parameters in an image recovery of statistical-mechanical means," *Physica A*, vol.223, pp.244-262, 1996.
- (5) K. Tanaka, M. Ichioka and T. Morita, "Statistical-mechanical iterative method in image restoration," *Trans. Inst. Electron. Inf. Commun. Eng. A*, vol.J80-A, pp.280-286, 1997 (in Japanese).
- (6) K. Tanaka, M. Ichioka and T. Morita, "Statistical-mechanical algorithm in MRF model based on variational principle," *Proceedings of 13th International Conference on Pattern Recognition*, vol.II, Track B: Pattern Recognition and Signal Analysis, Vienna, pp.381-388, 1996 (IEEE Computer Sciety Press).
- (7) K. Tanaka and T. Morita, "Statistical-mechanical iterative algorithm for image restoration of a gray level image," *Trans. Inst. Electron. Inf. Commun. Eng. A*, vol.J80-A, pp.1033-1037, 1997 (in Japanese).
- (8) J. Besag, "On the statistical analysis of dirty picture (with discussion)," *J. Royal Statistical Society Ser. B*, vol.48, pp.259-302, 1986.
- (9) W. Qian and D. M. Titterington, "Stochastic relaxations and EM algorithms for Markov random fields," J. Statist. Comput. Simul., vol.40, pp.55-69, 1991.
- (10) C.-h. Wu and P.C.Doerschk, "Cluster expansions for the deterministic computation of Bayesian estimators based on Markov random fields," *IEEE Trans. Pattern Anal. Mach. Intel.*, vol.17, pp.275, 1995.
- (11) D. Geiger and F. Girosi, "Parallel and deterministic algorithms from MRF's: surface reconstruction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol.13, pp.401-412, 1991.
- (12) J. Zhang, "The mean field theory in EM procedures for Markov random fields," *IEEE Trans. Signal Processing*, vol.40, pp.2570-2583, 1992.
- (13) J. Zhang, "The application of the Gibbs-Bogoliubov-Feynman inequality in mean field calculations for Markov random fields," *IEEE Trans. Image Processing*, vol.5, pp.1208-1214, 1996.
- (14) T. Morita, "General structure of the distribution functions for the Heisenberg model and the Ising model," *J. Math. Phys.*, vol.13, pp.115-123, 1972.
- (15) K. Tanaka and T. Morita, "Cluster variation method and image restoration problem," *Physics Letters A*, vol.203, pp.122-128, 1995.
- (16) T. Morita and K. Tanaka, "Cluster ZTP in the recovery of an image," *Pattern Recognition Letters*, vol.18, pp.1479-1493, 1997.

Table 1. Value of quantity $\bar{\sigma}_{2,n}(\tilde{\mathbf{x}})$ for the true original image $\tilde{\mathbf{x}}$ given in Fig. 1.

	n	$reve{\sigma}_{2,n}(ilde{\mathbf{x}})$	n	$ ilde{\sigma}_{2,n}(ilde{\mathbf{x}})$
	0	0.7224	4	0.0039
	1	0.2351	5	0.0031
-	2	0.0204	6	0.0033
	3	0.0067	7	0.0051



Fig. 1. Original image $\tilde{\mathbf{x}}$ (q = 8, M = N = 64).

Table 2. Value of optimal parameter \hat{J}_1 and \hat{J}_2 , and values of quantities $\bar{d}(\hat{\mathbf{x}}, \tilde{\mathbf{x}})$, $\bar{\sigma}_2(\hat{\mathbf{x}})$ and $\bar{\sigma}_{2,1}(\hat{\mathbf{x}})$ in the restored image $\hat{\mathbf{x}}$, which is obtained by using the ICM, the CZTP and the MFA. (a) (q-1)p=0.1. (b) (q-1)p=0.3. (c) (q-1)p=0.5.

(6	3)					
		\hat{J}_1	\hat{J}_2	$ar{d}(\mathbf{\hat{x}},\mathbf{ ilde{x}})$	$\bar{\sigma}_2(\hat{\mathbf{x}})$	$ar{\sigma}_{2,1}(\mathbf{\hat{x}})$
	ICM	0.2500	0.5015	0.0427	0.2782	0.2385
	CZTP	0.2500	0.5600	0.0420	0.2750	0.2395
l	MFA	0.2541	0.4511	0.0374	0.2772	0.2295
- {	$\hat{\mathbf{x}} = \tilde{\mathbf{x}}$			0	0.2764	0.2390

)		- ÷	1 1/2 2	I - (2)	[_ (<u>n</u>)
ICM	$\frac{J_1}{0.2533}$	$\frac{J_2}{0.8800}$	$d(\hat{\mathbf{x}}, \tilde{\mathbf{x}}) = 0.1211$	$\tilde{\sigma}_2(\hat{\mathbf{x}}) = 0.2769$	$\bar{\sigma}_{2,1}(\hat{\mathbf{x}}) = 0.2361$
CZTP	0.2533	0.6200	0.1211	0.2709	0.2356
MFA	0.2695	0.7631	0.1055	0.2734	0.2394
$\hat{\mathbf{x}} = \tilde{\mathbf{x}}$			0	0.2764	0.2390

;)					
	\hat{J}_1	\hat{J}_2	$ ilde{d}(\mathbf{\hat{x}},\mathbf{ ilde{x}})$	$\bar{\sigma}_2(\mathbf{\hat{x}})$	$\bar{\sigma}_{2,1}(\hat{\mathbf{x}})$
ICM	0.4611	1.4611	0.2793	0.2763	0.2416
CZTP	0.2599	0.8899	0.2327	0.2632	0.2354
MFA	0.2785	0.7701	0.2210	0.2752	0.2394
$\mathbf{\hat{x}} = \mathbf{\tilde{x}}$			0	0.2764	0.2390

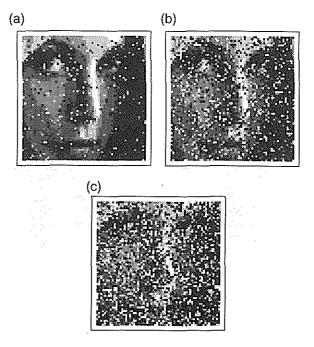


Fig. 2. Degraded images (q=8, M=N=64). (a) $(q-1)\tilde{p}=0.1$ $(\bar{d}(\mathbf{\tilde{x}},\mathbf{y})\simeq 0.0947)$. (b) $(q-1)\tilde{p}=0.3$ $(\bar{d}(\mathbf{\tilde{x}},\mathbf{y})\simeq 0.2937)$. (c) $(q-1)\tilde{p}=0.5$ $(\bar{d}(\mathbf{\tilde{x}},\mathbf{y})\simeq 0.4968)$.

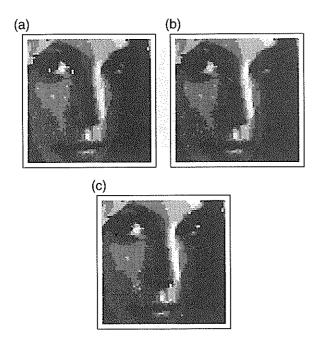


Fig. 3. Restored images $\hat{\mathbf{x}}$ obtained from the degraded image \mathbf{y} given in Fig. 2a. (a) ICM. (b) CZTP. (c) MFA.

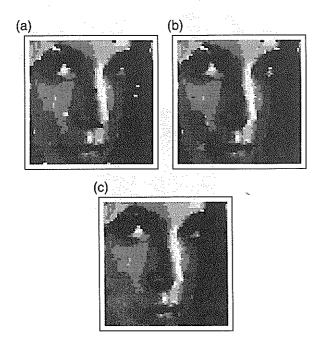


Fig. 4. Restored images $\hat{\mathbf{x}}$ obtained from the degraded image \mathbf{y} given in Fig. 2b. (a) ICM. (b) CZTP. (c) MFA.

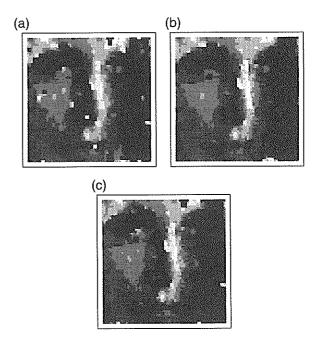


Fig. 5. Restored images $\hat{\mathbf{x}}$ obtained from the degraded image \mathbf{y} given in Fig. 2c. (a) ICM. (b) CZTP. (c) MFA.

統計力学的手法にもとづくマルコフ確率場モデルにおける近似最適化アルゴリズム

田中 和之*, 前田 純治**

概要

濃淡画像の画像修復をエネルギー最小化問題として定式化し、そのエネルギー関数からギブス分布として確率モデルを定義した場合、この確率モデルは統計力学においては古典スピン系と見なすことができ、画像処理においてはマルコフ確率場モデルと呼ばれている。本論文では統計力学的手法をもとにしたいくつかのエネルギー最小化アルゴリズムすなわち最適化アルゴリズムについて概説し、画質の改善にどのように結びつくかについて理論的に比較検討する。

キーワード:ソフトコンピューティング、確率コンピューティング、知識情報処理、統計 的手法、ベイズ統計、ギブス分布、マルコフ確率場、平均場理論、画像修復

*東北大学大学院情報科学研究科情報基礎科学専攻、**情報工学科