A Homotopy Continuation Method for Parameter Estimation in MRF Models and Image Restoration

P. K. Nanda*, U. B. Desai and P. G. Poonacha

Department of Electrical Engineering
Indian Institute of Technology - Bombay
Bombay 400 076
India
email: ubdesai@ee.iitb.ernet.in

Abstract

In this paper, we present an alternate approach to estimate the parameters of a Markov random field (MRF) model for images using the concepts of homotopy continuation method. We also develop a joint parameter estimation and image restoration scheme where we have used a fairly general model involving the line fields and tested on a real image. Simulation results using gray level images are presented.

1 Introduction

Markov random field (MRF) models for images have been successfully exploited in image restoration and early vision problems like edge detection, segmentation, interpolation of sparse depth data, and integration of early vision modules. The literature on this is very vast and we cite a few papers ([1] - [12]) to illustrate the breadth of applicability of MRF models.

In most applications, the performance of the algorithms using MRF models depends on the choice of the MRF model parameters. In a broad sense, two class of approaches are adopted for estimating the parameters; (i) based on the generalized cross validation scheme of Wahba [12] and (ii) based on maximizing the likelihood function or the pseudolikelihood function of Besag [8] (for example [7] , [9]- [11]). Most of the approaches assume the availability of a good initial image, which is often not the case in practice. Lakshmanan and Derin [9] and Younes [11] have considered the problem of simultaneous parameter estimation and segmentation.

In particular, we consider the joint solution to the MRF model parameter estimation and image restoration. Moreover, we also let the noise variance σ^2 to be an unknown. Though, at first, this appears to be an addition of one extra parameter, it does require significant modifications in the earlier approaches.

Our formulation of the MRF model parameter estimation exploits some of the recent advances in homotopy continuation methods for finding roots of a *smooth* function (see for example [13] - [15]). For the joint parameter estimation and restoration scheme, we use the homotopy continuation method for parameter estimation, and simulated annealing for image restoration.

*On leave from Regional Engineering College, Rourkela

[0]This work is supported by an MHRD (India) R & D project on "Computer Vision".

2 Problem Formulation

2.1 The MRF Model

The image considered is defined on an $(N \times N)$ rectangular lattice. The model considered is $Y_{ij} = X_{ij} + N_{ij}, \forall\, (i,j) \in (N \times N)$ which, using a lexicographical ordering will be $Y = X + N$, where, $Y = [Y_{ij}]$ = observed image random field, $X = [X_{ij}]$ = unknown image random field, and $N = [N_{i,j}]$ = noise random field. We make the following assumptions. (a) $N_{i,j}$ is a white Gaussian sequence with zero mean and variance σ^2 (σ unknown). (b) $N_{i,j}$ is statistically independent of $X_{k,l}$, for all (i,j) and (k,l) belonging to $N \times N$. (c) $x_{i,j}$ takes any gray level value from the set $G = (1, \ldots, N_G)$, (typically $N_G = 256$).

It is known [1] that X is a MRF with respect to neighbourhood system η if and only if $P(X = x)$ is Gibbs distributed with respect to η. This is expressed as

$$P(X = x \mid \phi) = \frac{1}{Z} e^{-U(x,\phi)}$$

where $Z = \sum_x e^{-U(x,\phi)}$ is the partition function, ϕ represents the clique parameter vector, the exponent term $U(x,\phi)$ is called the energy function and is of the form

$U(x,\phi) = \sum_{C:(i,j)\in c} V_c(x,\phi)$, with $V_c(x,\phi)$ being referred to as the potential. In general $\theta = [\phi^T, \sigma^2]^T$.

2.2 Parameter Estimation and Restoration Problem

As mentioned in [9] a general approach for joint parameter estimation and say the restoration problem would be to solve the following problem :

$$(x^*, \theta^*) = arg \quad \max_{x,\theta} P(X = x, \theta \mid Y = y) \qquad (1)$$

This is an extremely difficult problem and no known algorithm exists. Thus we consider an alternate formulation, which of course would give suboptimal solution to (1). We consider a scheme where we alternate between parameter estimation and image restoration. Let at iteration k $\theta^k = [\phi^k, (\sigma^2)^k]^T$ be the estimate of the parameters, and x^k be the estimate of the image X. Now consider the following problems :

$$x^{k+1} = arg \quad \max_x P(X = x \mid Y = y, \theta^k) \qquad (2)$$

and

$$\theta^{k+1} = arg \quad \max_\theta P(X = x^{k+1} \mid Y = y, \theta) \qquad (3)$$

The first problem (2) can be solved as a maximum a posteriori (MAP) estimation problem using a Bayesian approach [1]. It is easily shown that problem (2) is equivalent to

$$x^{k+1} = arg \quad \max_x \left[e^{\frac{-\|y-x\|^2}{2\sigma^2}} e^{-U(x,\phi^k)} \right] \qquad (4)$$

which can be solved using for example the simulated annealing algorithm.

Now in the parameter estimation problem (3) we can express the conditional probability as

$$P(X = x^{k+1} \mid Y = y, \theta) = \frac{\frac{1}{2\pi^{N^2/2}} e^{\frac{-\|y-x\|^2}{2\sigma^2}} \frac{1}{Z} e^{-U(x,\phi)}}{P(Y = y \mid \theta)}$$

Unfortunately, $P(Y = y \mid \theta)$ is no longer constant. It can be shown that

$$P(Y = y \mid \theta) = \sum_\xi \frac{1}{(2\pi)^{N^2/2}} e^{\frac{-\|y-\xi\|^2}{2\sigma^2}} \frac{1}{Z} e^{-U(\xi,\phi)}$$

which implies

$$P(X = x^{k+1} \mid Y = y, \theta) = \frac{e^{\frac{-\|y-x^{k+1}\|^2}{2\sigma^2}} e^{-U(x^{k+1},\phi)}}{\sum_\xi e^{\frac{-\|y-\xi\|^2}{2\sigma^2}} e^{-U(\xi,\phi)}} \qquad (5)$$

In (5) the summation is over all possible realizations of X. Thus, from a computational standpoint, handling (5) would be practically impossible. One can view (5) as a likelihood function to be maximized for estimating θ. To overcome the computational problem, we approximate (5) using the pseudolikelihood function.

$$\begin{aligned}\prod_{i,j} P(X_{i,j} = x_{i,j}^{k+1} \mid X_{m,n} = x_{m,n}^{k+1},\ (m,n) \in \eta_{i,j}, Y = y, \theta) \\ \triangleq \hat{P}(X = x^{k+1} \mid Y = y, \theta) \\ \approx P(X = x^{k+1} \mid Y = y, \theta)\end{aligned} \qquad (6)$$

This is essentially the pseudolikelihood function of Besag [8], except we are approximating the posterior probability distribution instead of the a priori probability distribution. It can be shown that

$$\hat{P}(X = x^{k+1} \mid Y = y, \theta) = \prod_{i,j} \left[\frac{e^{\frac{-(y_{i,j} - x_{i,j}^{k+1})^2}{2\sigma^2}} e^{-\sum_{C:(i,j)\in c} V_c(x_{i,j}^{k+1},\phi)}}{\sum_{x_{i,j}^{k+1} \in G} \left\{ e^{\frac{-(y_{i,j} - x_{i,j}^{k+1})^2}{2\sigma^2}} e^{-\sum_{C:(i,j)\in c} V_c(x_{i,j}^{k+1},\phi)} \right\}} \right] \qquad (7)$$

Due to space limitation, we have omitted the derivation of (7), nevertheless it can be found in [16]. In (7) the summation is over the all possible G gray levels of the pixel $x_{i,j}^{k+1}$. Now the parameter estimation problem is recast as

$$\theta^{k+1} = arg \quad \max_\theta \hat{P}(X = x^{k+1} \mid Y = y, \theta) \qquad (8)$$

2.3 Homotopy Continuation Method

It is evident from Section 2.2 that the parameter estimation problem has been reduced to maximization of (7) with respect to θ. Towards this end let

$$f(\theta) = \frac{\partial}{\partial\theta} \left\{ \log [\hat{P}(X = x^{k+1} \mid Y = y, \theta)] \right\} \qquad (9)$$

Now the homotopy method is employed to solve $f(\theta) = 0$. To have an arbitrary starting point for the path, we have considered the fixed point homotopy map given by

$$h(\theta, \lambda, q) = \lambda f(\theta) + (1 - \lambda)(\theta - q) \qquad (10)$$

where $0 \le \lambda \le 1$ and q is an arbitrary starting point. Here the predictor-corrector method is employed to track the path defined by the homotopy in (10). The procedure can be briefly mentioned as follows: $\lambda^{k+1} = \lambda^k + \Delta\lambda$

$$\hat{\theta}_0^k = \theta^k - \Delta\lambda J_\theta^{-1}[h(\theta^k, \lambda^{k+1}, \theta^{k-1})] \frac{\partial h}{\partial\lambda}[h(\theta^k, \lambda^{k+1}, \theta^{k-1})] \qquad (11)$$

$$\hat{\theta}^k_{i+1} = \hat{\theta}^k_i - J^{-1}_{\hat{\theta}}[h(\hat{\theta}^k_i, \lambda^{k+1}, \theta^{k-1})]h(\hat{\theta}^k_i, \lambda^{k+1}, \theta^{k-1}) \quad (12)$$

The derivation of (11) is analogous to the derivation of (5a) of Stonick and Alexander [14] for our homotopy map (10). If for some M, $|\hat{\theta}^k_{M+1} - \hat{\theta}^k_M| \leq \gamma$ then we set $\hat{\theta}^k_M = \hat{\theta}^k = \theta^{k+1}$. For the fixed point homotopy map considered, (11) becomes

$$\begin{aligned} \hat{\theta}^k_0 \;=\; & \theta^k - \Delta\lambda \left\{ \frac{I}{1-(\lambda^k+\Delta\lambda)} - \right. \\ & \frac{I}{(1-(\lambda^k+\Delta\lambda))^2}\left[\frac{F^{-1}_\theta(\theta^k)}{(\lambda^k+\Delta\lambda)} + \right. \\ & \left.\left. \frac{I}{1-(\lambda^k+\Delta\lambda)}\right]^{-1}\right\} \\ & \{f(\theta^k) - (\theta^k - \theta^{k-1})\} \end{aligned} \quad (13)$$

Where I is the identity matrix.

3 Parameter Estimation and Restoration

Here we consider a 256 gray level image with the MRF model suggested in [1] with line fields. The a priori energy function for the model is $U(x,h,v,\phi) = \sum_{i,j} \alpha[(x_{i,j} - x_{i,j-1})^2(1-h_{i,j}) + (x_{i,j} - x_{i-1,j})^2(1-v_{i,j})] + \beta(h_{i,j} + v_{i,j})$ The corresponding posterior energy function is

$$U_p(x,h,v,\theta) = \sum_{i,j} \frac{(y_{i,j} - x_{i,j})^2}{2\sigma^2} + U(x,h,v,\phi) \quad (14)$$

Substituting (14) in (7) and (9) we obtain

$$f(\theta) = \sum_{i,j} \frac{\partial}{\partial\theta}\left[\{-U_p(x,h,v,\theta)\} - \log\left\{\sum_{x_{i,j}\in G} e^{-U_p(x,h,v,\theta)}\right\}\right] \quad (15)$$

The basic steps in the algorithm for simultaneously updating x^k and θ^k are

Algorithm 1

1. Input noisy image Y.
2. Initialize the parameter vector θ to θ^0.
3. Given θ^k estimate x^{k+1} by minimizing

$$\frac{\|y - x^k\|^2}{2\sigma^2} + U(x^k, h^k, v^k, \phi^k)$$

Here, we used simulated annealing algorithm for minimization.
4. Having determined x^{k+1}, estimate θ^{k+1} using the homotopy map (10) and the corresponding update equations (12) and (13) and with $f(\theta)$ given in (15). In the parameter estimation part the estimated image x^{k+1} and the noisy image Y are used as input.
5. If a stopping criterion is met, stop or else go to Step 3.

In our simulations the following stopping criterion was used. If $\frac{1}{N^2}\sum_{i,j}(x^{k+1}_{i,j} - x^k_{i,j})^2 \leq$ threshold, then stop.

4 Simulation and Results

In the problem of simultaneous parameter estimation and restoration the image is the "LISA" image shown in Fig 1. Figure 2 displays the noisy image with SNR=5dB. The SNR is defined by SNR= $10.0\log_{10}\left\{\frac{1}{N^2}\sum_{i,j}(x_{i,j} - m)^2/\sigma^2\right\}$ where σ=standard deviation of the zero mean white Gaussian noise and m=mean of the image. The noisy image is the input to Algorithm 1. The initial parameter vector was $\theta_0 = [0.1, 15.0, 7.0]$. The results of algorithm 1 for 3 iterations are depicted in Figures 3, 4, and 5, and the corresponding parameter values in Table 1. The SNR of the estimated image, i e of Figure 5, is 11.87dB, an improvement of 6.87dB.

TABLE 1

iteration no.	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma}$
1	0.1	15.0	7.0
2	0.003705	2.5334	21.4814
3	0.0267	9.611	10.2

Though not discussed here, we also implemented the algorithm for only parameter estimation (supervised learning). To test the parameter estimation algorithm, simulations were carried out using a binary image under the Ising model. The input to the algorithm is the original image X and the noisy image Y. Table 2 gives a small sample of the simulation result. The notation θ^0 having value 0.1 implies $\alpha^0 = \beta^0 = (\sigma^2)^0 = 0.1$.

TABLE 2

θ^0	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\sigma^2}$	α	β	σ^2
0.1	0.0169	0.5382	0.2362	1.0	0.5	0.25
0.1	-0.081	0.37	0.063	1.0	0.5	0.0625
0.2	0.098	0.571	0.808	1.0	0.5	1.0

References

[1] S. Geman and D. Geman ," Stocastic Relaxation,Gibbs distribution and Bayesian restoration of images," *IEEE Trans. on Pattern Anal.and Machine Intell.*, Vol-6, No. 6, pp 721-741, 1984.

[2] J. Marroquin, S. Mitter and T. Poggio," Probabilistic solutions of ill-posed problems in computational vision. ," *J. of Amer. Stat. Asso.* Vol-82, pp 76-79, 1987.

[3] H. Derin, H. Elliot, R. Gristi and D. Geman," Bayes smothing algorithm for segmentation of binary images modelled by MRF fields.," *IEEE Trans. Patt. Annal. Machine Intell.* Vol-6, pp 707-720, Nov 1984.

[4] F. R. Hansen and H. Elliot," Image segmentation using simple Markov random field models.," *Compu. Vision Graph. and Image Proce.* Vol-20, pp 101-132, 1982.

[5] S. Kapoor, P. Y. Mundkur and U. B. Desai." Depth Recovery using MRF.,"*Sadhana,* Vol-18, Part-1, pp 17-29, March 1993.

[6] M. R. Bhatt and U. B. Desai," Robust image rstoration algorithm using MRF," *Proc. of Int. Sympo. on Circuits and Systems,* pp 2473-2476, 1992.

[7] R. L. Kashyap and R. Chellapa,"Estimation and choice of neighbours in spatial -interaction models of images," *IEEE Trans. Infomation Theory*, Vol-29, pp 60-72, Jan 1983.

[8] J. Besag,"On the statistical analysis of dirty pictures ," *J. Roy. Stat. Soc. Series-B*, Vol-48, pp 259-302, 1986.

[9] S. Lakshmanan and H. Derin,"Simultaneous parameter estimation and segmentation of Gibbs random fields using simulated annealing ," *IEEE Trans. Patt. Annal. and Machine Intell.* , Vol 11, pp 799-813, 1989.

[10] G. R. Cross and A. K. Jain," Markov Random Field texture models, *IEEE Trans. Patt. Annal. and Machine Intell.*, Vol 5, pp 25-39, Jan 1983.

[11] L. Younes," Estimation and annealing for Gibbsian fields," *Ann. Inst. Henri Poincar'e.* 24(2), pp 269-294, 1988.

[12] G. Wahba," Practical approximate solutions to linear operator equations when the data are noisy,"*SIAM J. Numer. Anal.* Vol.14, 1977.

[13] L. T . Watson,"Numerical linear algebra aspects of globally convergent homotopy methods ," *SIAM Review*, Vol. 28, No-4, pp 529-545, Dec 1986.

[14] V. L. Stonick and S. T. Alexander,"A Relationship between recursive least square update and homotopy continuation methods," *IEEE Tran. Signal Processing*, Vol. 39, No. 2, pp 530-532, Feb. 1991.

[15] V. L. Stonick and S. T. Alexander,"Globally optimal rational approximation using homotopy continuation methods," *IEEE Tran. Signal Processing*, Vol. 40, No. 9, pp 2358-2361, Sept 1992.

[16] P. K. Nanda, U. B. Desai and P. G. Poonacha,"A homotopy continuation method for parameter estimation in MRF models and image restoration", Tech. Report SPANN-94.1, Deptt.of Electrical Engg. IIT-Bombay.

Figure 1

Figure 2

Figure 3

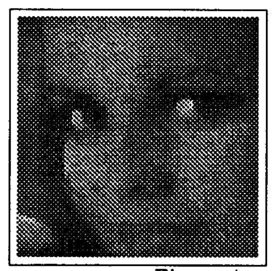

Figure 4

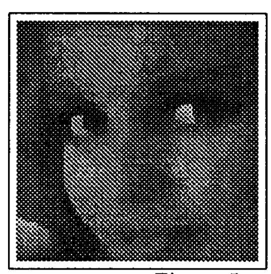

Figure 5