Probabilistic image processing based on the $Q$-Ising model by means of the mean-field method and loopy belief propagation

Kazuyuki Tanaka  
Graduate School of Information Sciences  
Tohoku University  
Sendai 980-8579, Japan  
kazu@statp.is.tohoku.ac.jp

D. M. Titterington  
Department of Statistics  
University of Glasgow  
Glasgow G12 8QQ, UK  
mike@stats.gla.ac.uk

Abstract

The framework is presented of Bayesian image restoration for multi-valued images by means of the $Q$-Ising model. Hyperparameters in the probabilistic model are determined so as to maximize the marginal likelihood. Practical algorithms are described based the conventional mean-field approximation and loopy belief propagation. We compare the results empirically with those provided by conventional filters and the new methods are found to be superior.

1 Introduction

It is well known that some ideas based on statistical-mechanical methods are very useful for probabilistic image processing based on Bayesian statistical methods [1]. In particular, the advanced mean-field methods from statistical mechanics provide effective practical algorithms for probabilistic image processing [2]. Moreover, some computer scientists have noted the similarity between advanced mean-field methods and a loopy belief propagation (LBP) approach, for constructing a probabilistic inference system, that has been investigated in the artificial intelligence literature [3]. Recently, LBP has also been applied to Bayesian image processing [4]. The LBP method is equivalent to a Bethe approximation, which is one extension of the conventional mean-field approximation (MFA). Tanaka, Inoue and Titterington [5] described the version of LBP for probabilistic image processing based on the $Q$-Ising model.

In the present paper, we investigate the performance of the MFA and LBP in Bayesian image restoration with the hyperparameters estimated by maximizing a marginal likelihood. The main purpose is to clarify how sensitive the realisations are to the values taken by the hyperparameters in the advanced mean-field method, and to what extent they are improved if LBP is used rather than the simple MFA. Moreover, we compare the results with those obtained with some conventional image processing filters.

2 Bayesian Image Analysis

We consider an image on a square lattice $\Omega \equiv \{i\}$ such that each pixel takes one of the gray-levels $Q = \{0, 1, 2, \ldots, Q - 1\}$, with 0 and $Q - 1$ corresponding to black and white, respectively. The intensities at pixel $i$ in the original image and the degraded image are regarded as random variables denoted by $F_i$ and $G_i$, respectively. Then the random fields of intensities in the original image and the degraded image are represented by $F \equiv \{F_i| i \in \Omega\}$ and $G \equiv \{G_i| i \in \Omega\}$, respectively. The actual original image and degraded image are denoted by $f = \{f_i| i \in \Omega\}$ and $g = \{g_i| i \in \Omega\}$, respectively.

The probability that the original image is $f$, $\Pr\{F = f\}$, is called the a priori probability of the image. Through the Bayes formula, the a posteriori probability $\Pr\{F = f| G = g\}$, that the original image is $f$ when the given degraded image is $g$, is expressed as

$$\Pr\{F = f| G = g\} = \frac{\Pr\{G = g| F = f\}\Pr\{F = f\}}{\sum_z \Pr\{G = g| F = z\}\Pr\{F = z\}},$$

(1)

where the summation $\sum_z$ is taken over all possible configurations of images $z = \{z_i| i \in \Omega\}$. The probability $\Pr\{G = g| F = f\}$ is the conditional probability that the degraded image is $g$ when the original image is $f$ and describes the degradation process.

In the present paper, it is assumed that the degraded image $g$ is generated from the original image $f$ by means of the $Q$-Ising model. Hyperparameters in the advanced mean-field method, and to what extent they are improved if LBP is used rather than the simple MFA. Moreover, we compare the results with those obtained with some conventional image processing filters.
f by changing the intensity of each pixel to another intensity with the same probability p, independently of the other pixels; the probability that the intensity is unchanged is therefore 1 - (Q - 1)p, which constrains p to lie in the range 0 ≤ p ≤ 1/Q. The conditional probability associated with the degradation process when the original image is f is

$$
\Pr\{G = g|F = f, p\} = \frac{1}{Z_{\text{PR}}(\alpha)} \prod_{ij \in B} \exp\left(-\alpha(f_i - f_j)^2\right),
$$

(2)

where $\delta_{a,b}$ is the Kronecker delta. Moreover, the a priori probability that the original image is $f$ is assumed to be

$$
\Pr\{F = f|\alpha, p\} = \frac{1}{Z_{\text{PO}}(\alpha)} \prod_{ij \in B} \exp\left(-\alpha(f_i - f_j)^2\right),
$$

where $B$ is the set of all the nearest-neighbor pairs of pixels on the square lattice $\Omega$. By substituting Eqs. (2) and (3) into Eq. (1), we obtain

$$
\Pr\{F = f|G = g, \alpha, p\} = \frac{1}{Z_{\text{PO}}(\alpha,p)} \prod_{ij \in B} \exp\left(-\alpha(f_i - f_j)^2\right)
$$

$$
\times \left(\prod_{i \in \Omega} \left(p(1 - \delta_{f_i, g_i}) + (1 - (Q - 1)p)\delta_{f_i, g_i}\right)\right).
$$

(4)

where $Z_{\text{PO}}(\alpha,p)$ is a normalization constant.

In the maximum marginal likelihood estimation approach, the hyperparameters $\alpha$ and $p$ are determined so as to maximize the marginal likelihood

$$
\Pr\{G = g|\alpha, p\} = \sum_z \Pr\{G = g|F = z, p\} \Pr\{F = z|\alpha\}.
$$

We denote the maximizers of the marginal likelihood $\Pr\{G = g|\alpha, p\}$ by $\hat{\alpha}$ and $\hat{p}$. The marginal likelihood $\Pr\{G = g|\alpha, p\}$ is expressed as follows:

$$
\Pr\{G = g|\alpha, p\} = \frac{Z_{\text{PO}}(\alpha,p)}{Z_{\text{PR}}(\alpha)}.
$$

(5)

Given the estimates $\hat{\alpha}$ and $\hat{p}$, the restored image $\hat{f} = \{\hat{f}_i|i \in \Omega\}$ is determined so as to maximize the posterior marginal probability

$$
\Pr\{F_i = f_i|G = g, \hat{\alpha}, \hat{p}\} = \sum_z \delta_{f_i, z} \Pr\{F = z|G = g, \hat{\alpha}, \hat{p}\}
$$

at each pixel $i$.

### 3 MFA and LBP

In the above framework, we have to calculate the marginal probabilities $\Pr\{F_i = f_i|G = g, \alpha, p\} \ (i \in \Omega)$, $Z_{\text{PO}}(\alpha,p)$ and $Z_{\text{PR}}(\alpha)$. Since it is hard to calculate the marginal probabilities exactly, we apply the MFA and LBP to the above a priori and a posteriori probabilistic models.

We now summarize loopy belief propagation for a probability distribution $\rho(f)$ defined by

$$
\rho(f) = \frac{1}{Z} \left(\prod_{i \in \Omega} \psi_i(f_i)\right) \left(\prod_{ij \in B} \phi(f_i, f_j)\right).
$$

(6)

Here the functions $\phi(\xi, \xi') = \exp(-\alpha(\xi - \xi')^2)$ and $\psi_i(\xi)$ are always positive for any values of $\xi$ and $Z$ is a normalization constant. The detailed derivation is similar to that of the MFA and LBP (i.e., the Bethe approximation) in Ref. [2] and here we merely state the main idea and the results.

In the MFA, we introduce the average $m_i = \sum_z z \rho(z)$ and the marginal probability distribution $\rho_i(\xi) = \sum_z \delta_{\xi, z} \rho(z)$. We assume that probabilities of configurations in which the approximate equalities $(f_i - m_i)(f_j - m_j) < 0$ hold are much larger than those of other configurations. By substituting this assumption in Eq. (6) and by calculating the marginal probability $\rho_i(\xi)$, we obtain simultaneous deterministic equations for $\{\rho_i(\xi)|i \in \Omega\}$ and the approximate value of $Z$ as follows:

$$
\rho_i(\xi) = \frac{\psi_i(\xi) \prod_{j \in c_i} \prod_{\xi} \phi(\xi, \xi') \rho_j(\xi')}{\sum_{\xi} \sum_{\xi'} \psi_i(\xi') \prod_{j \in c_i} \prod_{\xi} \phi(\xi', \xi') \rho_j(\xi')},
$$

(7)

and

$$
\ln Z \approx -\sum_{i \in \Omega} \sum_{\xi} \rho_i(\xi) \ln \left(\frac{\rho_i(\xi)}{\psi_i(\xi)}\right)
$$

$$
+ \sum_{ij \in B} \sum_{\xi} \sum_{\xi'} \rho_i(\xi) \rho_j(\xi') \ln \phi(\xi, \xi').
$$

(8)

Here the notation $c_i \equiv \{j|i \in B\}$ represents the set of all the nearest-neighbor pixels of pixel i and the summations $\sum_{\xi}$ and $\sum_{\xi'}$ are taken over all possible states $Q$. Equations (7) have the form of fixed-point equations for the marginal probabilities $\{\rho_i(\xi)|i \in \Omega, \xi \in Q\}$, and in practice we use iterative methods to solve them.

In LBP, we have to introduce the pairwise marginal probabilities $\rho_{ij}(\xi, \xi') = \rho_{ij}(\xi', \xi) = \sum_{\delta_{\xi, z} \delta_{\xi', z}} \rho(z)$. The marginal probabilities $\rho_i(\xi)$ and $\rho_{ij}(\xi, \xi')$ are approximately expressed in the following forms:

$$
\rho_i(\xi) = \frac{1}{Z_i} \psi_i(\xi) \prod_{k \in c_i} \mu_{i \rightarrow k}(\xi),
$$

(9)

$$
\rho_{ij}(\xi, \xi') = \frac{1}{Z_{ij}} \psi_i(\xi) \phi(\xi, \xi') \psi_j(\xi')
$$

$$
\times \left(\prod_{k \in c_{i \setminus j}} \mu_{i \rightarrow k}(\xi)\right) \left(\prod_{l \in c_{j \setminus i}} \mu_{l \rightarrow j}(\xi')\right).
$$

(10)
For the marginal probabilities to be consistent, we must have the relationships \( \rho_i(\xi) = \sum_j \rho_{ij}(\xi, \zeta) \) (\( i \in \Omega, j \in c_i, \xi \in \mathbb{Q} \)). By substituting Eqs. (9) and (10) into these relationships, we obtain simultaneous fixed-point equations for the set of unknown quantities, \( \mu = \{\mu_{j-i}(\xi), \mu_{i-j}(\xi)\} | i, j \in B, \xi \in \mathbb{Q} \}, in the form

\[
\mu_{j-i}(\xi) = \frac{\sum \phi(\xi, \zeta) \psi_j(\zeta) \prod_{k \in c_j, \mu_{k-j}(\zeta)}}{\sum \phi(\zeta', \zeta) \psi_j(\zeta) \prod_{k \in c_j, \mu_{k-j}(\zeta)}} \cdot (11)
\]

In LBP, the approximate value of \( \ln Z \) for every probabilistic model in Eq. (6) is given by

\[
\ln Z \approx \sum_{i \in \Omega} \ln Z_i + \sum_{ij \in B} \left( \ln Z_{ij} - \ln Z_i - \ln Z_j \right) \cdot (12)
\]

In probabilistic inference, the quantity \( \mu_{i-j}(\xi) \) is called a message propagated from \( i \) to \( j \). Equations (11) have the form of fixed-point equations for the messages \( \mu_{i-j}(\xi) \). Again in practice we solve these equations by using iterative methods.

By setting \( \psi_i(\xi) = p(1-\delta_{f-G}) + (1-(Q-1)p)\delta_{f-G} \), we obtain the marginal probabilities \( \Pr\{F_i = \xi|G = g, \alpha, p\} \) and the normalization constant \( \mathcal{Z}_{P_{O}} \) as \( \rho_i(\xi) \) and \( \mathcal{Z} \), respectively. By setting \( \psi_i(\xi) = 1 \), we obtain the normalization constant \( \mathcal{Z}_{P_{R}} \) as \( \mathcal{Z} \). For various values of the hyperparameters \( \alpha \) and \( p \), we obtain the marginal probabilities \( \Pr\{F_i = f|G = g, \alpha, p\} \), \( \mathcal{Z}_{P_{O}}(\alpha, p) \) and \( \mathcal{Z}_{P_{R}}(\alpha) \) and search numerically for the optimal set of values, \((\hat{\alpha}, \hat{p})\).

4 Numerical Experiments

In this section, we report some numerical experiments. The optimal set of values of the hyperparameters, \((\hat{\alpha}, \hat{p})\), is determined by means of maximum marginal likelihood estimation combined with the MFA or LBP.

To evaluate restoration performance quantitatively, twenty original images \( f \) are generated from the \( a \ priori \) probability distribution (3) for \( Q = 4 \). We produce a degraded image \( g \) from each original image \( f \) by means of the degradation process (2) for \((Q-1)p = 0.3\). By applying the iterative algorithm of the MFA or LBP to each degraded image \( g \), we obtain estimates \( \hat{p} \) and \( \hat{\alpha} \) of the hyperparameters and the restored image \( \hat{f} \) for each degraded image \( g \). Examples of the original images in the numerical experiments are shown in Fig.1. From the twenty degraded images \( g \) and the corresponding restored images \( \hat{f} \), we calculate 95% confidence intervals for the hyperparameters, \( \hat{p} \) and \( \hat{\alpha} \), and the improvement in the signal to noise ratio, \( \Delta_{SNR} = 10 \log_{10} \left( \frac{\sum_{i \in \xi} (g_i - f_i)^2}{\sum_{i \in \xi} (f_i - \bar{f})^2} \right) \) in decibels (dB). These confidence intervals are given in Table 1. Clearly the true values of the hyperparameters, \( \alpha \) and \( p \), lie outside the 95% confidence intervals. The results might appear to throw doubt on the accuracy of the MFA and LBP so far as estimating the values of \( p \) and \( \alpha \) is concerned. The fact that the confidence intervals do not include the true values indicates a bias in estimating the hyperparameters using the MFA or LBP. However, the biases are clearly improved if LBP is used rather than the MFA. The biases would be very likely to be smaller again if we used an advanced mean-field method which is even more sophisticated than LBP.

Table 1: Approximate 95% confidence intervals for \( \hat{\alpha} \), \( (Q-1)\hat{p} \) and \( \Delta_{SNR} \) obtained by using the MFA and LBP for \( \alpha = 0.75 \) and \((Q-1)p = 0.3\). The twenty original images \( f \) are generated by Monte Carlo simulation from the \( a \ priori \) probabilities (3).

<table>
<thead>
<tr>
<th></th>
<th>MFA</th>
<th>LBP</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{\alpha} )</td>
<td>0.4609±0.0010</td>
<td>0.6206±0.0013</td>
</tr>
<tr>
<td>( (Q-1)\hat{p} )</td>
<td>0.2170±0.00071</td>
<td>0.2696±0.00075</td>
</tr>
<tr>
<td>( \Delta_{SNR} ) (dB)</td>
<td>2.7705±0.03492</td>
<td>4.6184±0.03245</td>
</tr>
</tbody>
</table>

We then performed numerical experiments based on the artificial image in Fig.2(a). A degraded image \( g \), produced from the original image \( f \) with \((Q-1)p = 0.3\), is shown in Fig.2(b). The image restorations created by means of the iterative algorithms based on the MFA and LBP for the \( Q \)-Ising model are shown in Figs.2(c) and (d). We give in Table 2 the estimates, \( \hat{p} \) and \( \hat{\alpha} \), of the hyperparameters, and the values of the improvement in the signal to noise ratio, \( \Delta_{SNR} \) (dB). Table 3 provides values of \( \Delta_{SNR} \) corresponding to image restorations obtained by the application of some conventional filters. In comparison, the MFA and especially LBP both give us good results.
Figure 2: Image restorations obtained by using the MFA and LBP. (a) Original image $f$ ($Q = 4$). (b) Degraded image $g$ which was produced for $(Q - 1)p = 0.3$ from the original image $f$. (c) Restored image $\hat{f}$ based on the MFA. (d) Restored image $\hat{f}$ based on LBP.

Table 2: The estimates $\hat{p}$ and $\hat{\alpha}$ of the hyperparameters and the value of $\Delta_{\text{SNR}}$ obtained by applying the MFA and LBP to the degraded image $g$ in Fig. 2(b).

<table>
<thead>
<tr>
<th></th>
<th>$\hat{p}$</th>
<th>$\hat{\alpha}$</th>
<th>$\Delta_{\text{SNR}}$ (dB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MFA</td>
<td>0.248</td>
<td>0.602</td>
<td>7.653</td>
</tr>
<tr>
<td>LBP</td>
<td>0.281</td>
<td>0.788</td>
<td>8.892</td>
</tr>
</tbody>
</table>

5 Concluding Remarks

We have investigated the performance of the MFA and LBP for probabilistic image restoration when the $Q$-Ising model is adopted as the a priori probability distribution. The hyperparameters are determined so as to maximize the marginal likelihood for each degraded image. The results obtained by LBP represent systematic improvements over those obtained by the MFA. Both the MFA and LBP both provide better results than those provided by some conventional filters.

Recently, some computer scientists have suggested that loopy belief propagation can be generalized by using a cluster variation method[6], and we have already applied generalized belief propagation to the problem of binary image restoration[7]. Further application of generalized belief propagation to gray-level image restoration will be the subject of future research.

Table 3: The values of $\Delta_{\text{SNR}}$ obtained in the image restorations of the conventional image processing filters for the degraded image $g$ in Fig. 2(b). The notation $(3 \times 3)$ and $(5 \times 5)$ denote the window sizes in the conventional filters.

<table>
<thead>
<tr>
<th>Filter</th>
<th>$(3 \times 3)$</th>
<th>$(5 \times 5)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lowpass filter</td>
<td>3.437</td>
<td>3.153</td>
</tr>
<tr>
<td>Wiener filter</td>
<td>1.037</td>
<td>2.177</td>
</tr>
<tr>
<td>Median filter</td>
<td>6.715</td>
<td>7.301</td>
</tr>
</tbody>
</table>

Acknowledgements

The authors are grateful to Professor J. Inoue of the Graduate School of Engineering, Hokkaido University and Professor T. Horiguchi of the Graduate School of Information Science, Tohoku University, for valuable discussions. This work was partly supported by the Grants-In-Aid (No.14084203) for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan.

References