Hyper-Parameter in Hidden Markov Random Field

Johan Lim¹ · Donghyeon Yu² · Kyungsuk Pyun³

¹Department of Statistics, Seoul National University
²Department of Statistics, Seoul National University; ³Samsung Electronics Company

(Received April 2010; accepted November 2010)

Abstract

Hidden Markov random field (HMRF) is one of the most common model for image segmentation which is an important preprocessing in many imaging devices. The HMRF has unknown hyper-parameters on Markov random field to be estimated in segmenting testing images. However, in practice, due to computational complexity, it is often assumed to be a fixed constant. In this paper, we numerically show that the segmentation results very depending on the fixed hyper-parameter, and, if the parameter is misspecified, they further depend on the choice of the class-labelling algorithm. In contrast, the HMRF with estimated hyper-parameter provides consistent segmentation results regardless of the choice of class labelling and the estimation method. Thus, we recommend practitioners estimate the hyper-parameter even though it is computationally complex.

Keywords: Hidden Markov random field, hyper-parameter, image segmentation.

1. Introduction

Segmentation is a process which divides an image into several homogeneous regions, whereas classification matches an unknown image with a known image in the database. Despite their apparent difference, classifiers are often used for segmenting an image with the name of context-based classifier.

A context-based classifier partitions the whole image into many sub-blocks and classifies each sub-block into one of the classes in the database. In doing so, a difficulty arises in obtaining smooth boundaries, which is often a trait of the true segmentation. To obtain smooth boundaries, hidden Markov models (HMM) have often been proposed, in which the spatial coherence between sub-blocks is modelled by the Markovian random field (Geman and Geman, 1984; Besag, 1986; Li and Gray, 1999; Li et al., 1999). The Ising model or the Potts model are popularly used depending on the number of classes in the true segmentation.

Suppose the observed image has \( n \times n \) pixels and is denoted as \( X = \{ X_{ij}, 1 \leq i \leq n, 1 \leq j \leq n \} \). Let \( Z_{ij} \in \{0, 1\} \) be the true class assigned to the image \( X_{ij} \) and \( Z = \{ Z_{ij}, 1 \leq i \leq n, 1 \leq j \leq n \} \).

Johan Lim’s research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by Ministry of Education, Science and Technology(No. 2010-0011448).

¹Corresponding author: Associate Professor, Department of Statistics, College of Natural Sciences Seoul National University, Daehak-dong, Gwanak-gu, Seoul 151-747, Korea. E-mail: johanlim@snu.ac.kr
denotes the true hidden segmentation corresponding to $X$. A common assumption for the true segmentation $Z$ is a symmetric nearest neighborhood (SNN) Ising model with $\alpha$:

$$P(Z = Z) = \frac{1}{\Psi(\alpha)} \exp \left\{ \alpha \cdot \sum_{(i,j) \sim (k,l)} V(Z_{ij}, Z_{kl}) \right\}, \quad (1.1)$$

where $\Psi(\alpha)$ is a normalizing constant, and $V(Z_{ij}, Z_{kl}) = 1$ if $Z_{ij} = Z_{kl}$, and 0 otherwise. Given $Z$, the observed image $X$ is from conditionally independent distributions, in which $P(X_{ij} = x|Z_{ij} = 1) = f_1(x)$ and $P(X_{ij} = x|Z_{ij} = 0) = f_0(x)$. Here, the distributions $f_0(x)$ and $f_1(x)$ are estimated from training data set and assumed to be known in segmentation procedure.

A main goal of image segmentation is to find the true class label $Z$ from observed image $X$. The most common estimate of $Z$, given known hyper-parameter, in previous literature is the maximum a posteriori (MAP) estimate which is defined by

$$\hat{Z}_{\text{MAP}} = \arg \max_{Z} P(Z = Z|X = X).$$

In practice, the hyper-parameter $\alpha$ associated with the Markov random field is unknown a priori and training data to estimate it are rarely available. Thus, it must be estimated directly from the given image data, whose maximum likelihood estimator (MLE) is the solution to

$$\sum_{(i,j) \sim (k,l)} E \{ V(Z_{ij}, Z_{kl}) |X \} = \sum_{(i,j) \sim (k,l)} E \{ V(Z_{ij}, Z_{kl}) \}, \quad (1.2)$$

where the expectation in right-hand side (RHS) is analytically known, but that in left-hand side (LHS) is approximated using various methods and is hard to compute. Due to this reason, we assume that $\alpha$ is known and fixed it as a subjectively chosen constant (Won and Gray, 2004).

In this paper, we demonstrate that it is better to estimate hyper-parameter than to fix it in all three image segmentation method - Gibbs Sampling, simulated annealing (SA) and iterative conditional mode (ICM).

2. Alternating Procedure to Jointly Estimate Class Labels and Hyper-parameter

In this section, we assume the observed image $X$ is binary and $Z$ is from the bond percolation model for notational simplicity. Given $Z$, the observed image $X$ is from conditionally independent Bernoulli distributions, in which $P(X_{ij} = 1|Z_{ij} = 1) = p_1$ and $P(X_{ij} = 1|Z_{ij} = 0) = p_0$. Here, model parameters $p_0$ and $p_1$ are accurately estimated from training data and are assumed to be known. Further, we rewrite (1.1) as

$$P(Z = Z) = \exp \left\{ \log \left( \frac{q}{1-q} \right) \cdot \sum_{(i,j) \sim (k,l)} V(Z_{ij}, Z_{kl}) \right\},$$

where $q$ is the probability that any adjacent pixels are independently in equal state.

In this model, the estimating equation (1.2) becomes

$$n^2 \cdot q = \sum_{(i,j) \sim (k,l)} E \{ I(Z_{ij} = Z_{kl}) |X \}, \quad (2.1)$$
Table 3.1. Misclassification error rate. The numbers in parenthesis are the estimates of $\alpha$.

<table>
<thead>
<tr>
<th>Method</th>
<th>$\alpha = 0.60$</th>
<th>$\alpha = 0.75$</th>
<th>Besag’s method</th>
<th>MFA</th>
</tr>
</thead>
<tbody>
<tr>
<td>ICM</td>
<td>0.2284</td>
<td>0.0530</td>
<td>0.0514 (0.9298)</td>
<td>0.0496 (0.9226)</td>
</tr>
<tr>
<td>SA</td>
<td>0.1054</td>
<td>0.0421</td>
<td>0.0482 (0.9344)</td>
<td>0.0482 (0.9265)</td>
</tr>
<tr>
<td>Gibbs</td>
<td>0.3084</td>
<td>0.0840</td>
<td>0.0502 (0.9211)</td>
<td>0.0500 (0.9134)</td>
</tr>
</tbody>
</table>

As stated before, evaluating $E(\cdot|\mathcal{X})$ is computationally hard, and various approximations for it have been proposed in previous literature. Besag (1986) uses $E(I(Z_{ij} = Z_{kl})|\mathcal{X}) \approx I(\hat{Z}_{ij} = \hat{Z}_{kl})$, where $\hat{Z}_{ij}$s are the current MAP estimate of $Z_{ij}$. Second, the mean field approximation (MFA), which is originally designed to approximate derivatives of the partition function, uses $E(I(Z_{ij} = Z_{kl})|\mathcal{X}) \approx E[I(Z_{ij} = Z_{kl})|\hat{Z}_{-(ij),kl}|, \mathcal{X}]$ (Zhou et al., 1997).


In summary, overall procedure is the alternation between (i) get the MAP class label given current estimate $\hat{\alpha}$ and (ii) update the estimate of $\alpha$ using current MAP of $Z$. However, due to computational complexity, many authors assume hyper-parameter is a known constant.

3. Examples

3.1. Simulated image

We implement an experiment to show the importance of the hyper-parameter in segmenting an image. The true class labels $Z$ we use is the check-board image presented in Figure 3.1 (a). The observed image $X$ is generated using symmetric channel $P(X_{ij} = 1|Z_{ij} = 1) = 0.7$ and $P(X_{ij} = 1|Z_{ij} = 0) = 0.3$. We employ three methods (ICM, SA, and Gibbs sampling) to find MAP class labels. We use two methods, Besag’s mode approximation and mean field approximation (MFA) to estimate hyper-parameter $\alpha$. We also run the cases with fixed $\alpha$ as 0.6 and 0.75. The misclassification rate and parameter estimates for each case we consider are reported in Table 3.1. The visual segmentation results are reported in Figure 3.1.

The segmentation with an estimated hyper-parameter shows better performance in both misclassification error rate and visual segmentation. The results further shows that SA performs well in wider range of fixed parameter $\alpha$. However, unlike other two class labelling algorithms, the SA requires further specification on tempering schedule, for which, in our simulation, a linear tempering is chosen by trial and error. Thus, it is not very surprise that the SA performs better than Gibbs sampling and ICM when $\alpha$ is mis-specified. In both Besag’s method and MFA, the hyper-parameter was chosen as approximately 0.92 regardless of class labelling algorithm. In summary, we find that the hyper-parameters in HMRF is quite important to get better segmentation results.

3.2. Aerial image

In this section, we consider the segmentation of a real noisy binary image.
The database we use is composed of six aerial photographs of different locations in the San Francisco Bay Area. Each photograph is $1024 \times 1024$ and is a gray-level image with 8 bits per pixel. In each photo, white represents the manmade class, whereas black represents the natural class. It is paired with its hand-labelled classified image. These photographs have also been studied by Aiyer et al. (2005), Li et al. (2000), Pyun et al. (2007) and Pyun et al. (2009). In the analysis, we
To design computer-based system to measure the spatial correlation in land use, we first segment the images into manmade and natural area using a block-based classifier. In this paper, we use Gauss mixture based (GMM) classifier, which divides the image into blocks and is modeled separately using Gauss mixture model. In the analysis, we use 4 × 4 sub-blocks (256 × 256 sub-blocks in total) and estimating the Gauss mixture model for manmade and nature regions using 5 training images. The block processing steps are same with those in Aiyer et al. (2005) and Pyun et al. (2007) and their details are omitted here. Using the estimate GMMs, we classify each sub-block into the class having
the highest likelihood. The testing image and segmented image by GMM are plotted in the top two panels in Figure 3.2.

We apply the simulated annealing algorithm, which shows the best performance in the simulated data, with several choices of hyper-parameter $\alpha$. In this example, the segmented result for $\alpha = 0.75$ is similar to that with estimated $\tau$ using MFA, and is not reported here. Again, the segmentation results depend on the choice of $\alpha$. However, both estimates based on the mode approximation and MFA provide similarly reliable segmentations although the former is slightly smoother than the latter.

4. Conclusion

In this paper, we numerically show that the segmentation results very depending on the pre-specified hyper-parameter in HMRF based image segmentation. On the other hand, the HMRF with estimated hyper-parameter provides segmentation results which are robust to the choice of class-labelling algorithm as well as the estimation method. In consequence, we recommend practitioners estimate the hyper-parameter although it is rather computationally expensive.

Acknowledgments

We are grateful to the associate editor and two reviewers for their constructive comments.

References


