Bayesian Nonparametric Priors for Hidden Markov Random Fields: Application to Image Segmentation

Hongliang Lu, Julyan Arbel, Florence Forbes

To cite this version:


HAL Id: hal-01950666
https://hal.archives-ouvertes.fr/hal-01950666
Submitted on 11 Dec 2018
I. BACKGROUND AND MOTIVATION

Image segmentation in real-world applications is typically performed on noisy images. To achieve better segmentation performance, several extensions of the usual Bayesian nonparametric (BNP) mixture model with spatial regularization are therefore necessary.

II. BNP PRIORS

The Dirichlet process (DP) is one of the most commonly used BNP priors. It is a random process \( G \) defined over a probability space \( \mathcal{Y} \) and characterized by a concentration parameter \( \alpha \) and a base distribution \( G_0 \), such that for any finite partition \( \{A_1, \ldots, A_N\} \) of \( \mathcal{Y} \), the random vector \( (P(A_1), \ldots, P(A_N)) \) is Dirichlet distributed:

\[
(P(A_1), \ldots, P(A_N)) \sim \text{Dir}(\alpha G_0, \ldots, \alpha G_0)
\]

which is often denoted by \( G \sim \text{DP}(\alpha, G_0) \).

III. STICK-BREAKING CONSTRUCTION

The DP has almost surely discrete realizations. It can be built by the stick-breaking construction:

\[
G = \sum_{k=1}^{\infty} \pi_k \delta_{\tau_k} = \sum_{k=1}^{\infty} \left( \tau_k \prod_{i=1}^{k-1} (1-\tau_i) \right) \delta_{\tau_k}
\]

where \( \tau_k \sim \text{Beta}(1, \alpha) \) and \( \pi_k \sim \text{Beta}(\alpha, \alpha) \).

V. VARIATIONAL BAYES

In a Bayesian setting, we need to evaluate the intractable posterior distribution \( p(x, \tau, \alpha, \theta^y, \phi) \) which is denoted by a set of hyperparameters) which can be estimated by means of the mean-field approximation:

\[
q(x, \tau, \alpha, \theta^y, \phi) \simeq q_0(x)q_0(\tau)q_0(\alpha)q_0(\phi^*)
\]

Variational Bayes (VB) consists of alternating maximization of free energy

\[
F(q; \theta) = \mathbb{E}_{q_\tau q_\alpha q_\phi^*} \left[ \log \frac{p(x, \tau, \alpha, \theta^y, \phi)}{q_\tau q_\alpha q_{\phi^*}} \right]
\]

which implies [2]

- E-steps: VE-\( \alpha \), VE-\( \alpha^y \), VE-\( \tau \) and VE-\( \theta^* \)
- M-steps: \( \phi \) updating is straightforward except for \( \beta \).

Here, the \( M=\beta \) step leads to the estimation of \( \beta \):

\[
\hat{\beta} = \arg \max_\beta \mathbb{E}_{q_\tau} \left[ \log p(x|\tau, \beta) \right]
\]

which involves \( p(x|\tau, \beta) = K(\beta, \tau)^{-1} \exp \left( V(x, \tau, \beta) \right) \) with the normalization constant \( K(\beta, \tau) \) and the potential function

\[
V(x, \tau, \beta) = \sum_i \log \tau_i(\beta) + \beta \sum_j \delta_{z_j=i}
\]

To find the optimal value of \( \beta \), further approximations, such as the mean-field-like approximation [3] of \( q_\beta \) and replacing \( \tau \) with a fixed \( \hat{\tau} = \mathbb{E}_\tau(\tau) \), are required.

REFERENCES


IV. DP-POTTS MIXTURE MODEL

The usual DP mixture model assumes that a set of data points \( y = \{y_1, \ldots, y_N\} \) with \( y_i \in \mathbb{R}^D \) (e.g., pixels) can be generated through the following hierarchical representation:

\[
\begin{align*}
G & \sim \text{DP}(\alpha, G_0) \\
\theta_i | G & \sim G, \ i = 1, \ldots, N \\
y_i | \theta_i & \sim F(\theta_i), \ i = 1, \ldots, N
\end{align*}
\]

where \( \theta = \{\theta_1, \ldots, \theta_N\} \) denotes a set of model parameters. To take into account spatial constraints, we introduce a Potts model component using a set of assignment variables \( z = \{z_1, \ldots, z_N\} \) such that to favor spatial aggregation [1]:

\[
M(\theta) \propto \exp \left( \beta \sum_{i,j} \delta_{z(i)=z(j)} \right)
\]

with \( \beta \) being the regularization parameter. The DP mixture model is thus extended to become the DP-Potts mixture model:

\[
\begin{align*}
\mathcal{G} & \sim \text{DP}(\alpha, G_0) \\
\theta | \mathcal{G} & \sim \mathcal{M}(\theta) = \prod_i G_i(\theta_i) \\
y | \theta_i & \sim F(\theta_i), \ i = 1, \ldots, N
\end{align*}
\]

Accordingly, the stick-breaking construction of the DP-Potts mixture model can be summarized as follows:

\[
\begin{align*}
\theta_0 \sim G_0 \quad \text{and} \quad \theta_k & \sim \text{Beta}(1, \alpha), \ k = 1, 2, \ldots \\
x_1(\tau) & \sim \tau \prod_{i=1}^{k-1} (1-\tau_i), \ k = 1, 2, \ldots \\
y_i(\tau, \theta_0) & \sim F(\theta_0), \ i = 1, \ldots, N
\end{align*}
\]

VII CONCLUSIONS AND FUTURE WORK

A general DP-Potts mixture model and the associated VB algorithm were proposed. The model was successfully applied to image segmentation on different types of datasets. We also investigated the impact of \( \beta \) on the segmentation results and presented an estimation procedure for \( \beta \).

In the sequel, we plan to survey how \( \beta \) affects the inferred number of components. Other types of priors (Pitman-Yor process, Gibbs-type priors, etc.) and other variational approximations (truncation-free) will also be considered. On the other hand, it is crucial to study theoretical properties of BNP priors under structural constraints (temporal or spatial). Other applications may also be possible, such as discovery probability and community detection in graphs.

Figure 1: Illustration of challenges for unsupervised image segmentation: blur, noise, color/contrast imperfection, partial volume effect (large slice thickness), anatomic variability and complexity, number of segments.

Experiments were performed using superpixels on a subset (154 images) of the Berkeley segmentation data set (BSDS) [4]. Regarding the performance evaluation, the probabilistic rand index (PRI) was computed under different conditions:

<table>
<thead>
<tr>
<th>PRI</th>
<th>DP</th>
<th>DP-Potts</th>
<th>HMRF</th>
<th>MRF-PYP</th>
<th>Graph Cuts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (%)</td>
<td>77.23</td>
<td>77.98</td>
<td>75.50</td>
<td>76.49</td>
<td>76.10</td>
</tr>
<tr>
<td>Median (%)</td>
<td>78.05</td>
<td>78.56</td>
<td>78.69</td>
<td>78.08</td>
<td>78.59</td>
</tr>
</tbody>
</table>

We also compared our best results \( K = 50 \) and about 500 superpixels with those given in the literature:

Proposed models | Results given in [1]
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>PRI</td>
<td>DP</td>
</tr>
<tr>
<td>Mean (%)</td>
<td>77.23</td>
</tr>
<tr>
<td>Median (%)</td>
<td>78.05</td>
</tr>
</tbody>
</table>