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

Hongliang Lu, Julyan Arbel, Florence Forbes

To cite this version:

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

Hongliang Lu, Jyluray Arbel, Florence Forbes
Univ. Grenoble Alpes, Inria, CNRS, Grenoble INP, LJK, 38000 Grenoble, France

I. BACKGROUND AND MOTIVATION

Approximate number of superpixels

Impact of superpixels on PRI

DP-Potts (K=50)

DP (K=50)

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 Υ and characterized by a concentration parameter α and a base distribution G0, such that for any finite partition \( \{A_1, ..., A_p\} \) of \( \Upsilon \), the random vector \((P(A_1), ..., P(A_p))\) is Dirichlet distributed:

\[
(P(A_1), ..., P(A_p)) \sim \text{Dir}(\alpha G_0(A_1), ..., \alpha G_0(A_p))
\]

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:

\[
G = \sum_{k=1}^{\infty} \pi_k \delta_{z_k} = \sum_{k=1}^{\infty} \left( \gamma_{1-k} \prod_{r=1}^{k} (1 - \gamma_{r}) \right) \delta_i
\]

where \( \theta_{k} \sim G_0 \) and \( \gamma_k \sim B(1, \alpha) \).

V. VARIATIONAL BAYES

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

\[
q(\mathbf{x}, \alpha, \mathbf{\theta}^* | \mathbf{y}, \phi) \simeq q(\mathbf{x}, \alpha | \mathbf{y}, \phi) q(\mathbf{\theta}^* | \mathbf{y}, \phi)
\]

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

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

which implies [2]

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

Here, M-\( \beta \) step leads to the estimation of \( \hat{\beta} \):

\[
\hat{\beta} = \arg \max_{\beta} \mathbb{E}_{\mathbf{q}_{\beta}} \log p(\mathbf{x} | \alpha, \beta)
\]

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

\[
V(\mathbf{x}, \beta, \mathbf{\tau}) = \sum_{i} \log \tau_{i} + \beta \sum_{k} (z_{k} = i)
\]

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

REFERENCES