HOGMep: Variational Bayes and higher-order graphical models applied to joint image recovery and segmentation
Aurélie Pirayre, Yuling Zheng, Laurent Duval, Jean-Christophe Pesquet

To cite this version:
Aurélie Pirayre, Yuling Zheng, Laurent Duval, Jean-Christophe Pesquet. HOGMep: Variational Bayes and higher-order graphical models applied to joint image recovery and segmentation. 2017 IEEE International Conference on Image Processing (ICIP), Sep 2017, Beijing, China. 10.1109/ICIP.2017.8296988 . hal-01862840

HAL Id: hal-01862840
https://hal-ifp.archives-ouvertes.fr/hal-01862840
Submitted on 27 Aug 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
HOGMep: VARIATIONAL BAYESIAN AND HIGHER-ORDER GRAPHICAL MODELS APPLIED TO JOINT IMAGE RECOVERY AND SEGMENTATION

Aurélie Pirayre¹, Yuling Zheng³, Laurent Duval¹, and Jean-Christophe Pesquet²

¹ IFP Energies nouvelles, Rueil-Malmaison, France ; ² CentraleSupélec, France ; ³ IBM Research, China

ABSTRACT

Variational Bayesian approaches have been successfully applied to image segmentation. They usually rely on a Potts model for the hidden label variables and a Gaussian assumption on pixel intensities within a given class. Such models may however be limited, especially in the case of multicomponent images. We overcome this limitation with HOGMep, a Bayesian formulation based on a higher-order graphical model (HOGM) on labels and a Multivariate Exponential Power (MEP) prior for intensities in a class. Then, we develop an efficient statistical estimation method to solve the associated problem. Its flexibility accommodates to a broad range of applications, demonstrated on multicomponent image segmentation and restoration.

Index Terms— variational Bayes, higher-order graphical models, multicomponent images, segmentation, restoration

1. INTRODUCTION

Variational Bayesian Approximation (VBA) has been widely used in various applications such as in graphical model learning [1], image processing [2, 3, 4, 5], source separation [6] or super-resolution [7]. Its popularity is mainly due to its ability to generate efficient approximations of the posterior distribution of the variables to be estimated. In image segmentation, VBA-based approaches traditionally uses a Potts (resp. Gaussian) model for hidden label variables (resp. pixel intensities corresponding to a given label) [8, 9, 10].

We propose a more generic Bayesian model, combining two main priors on the sought data. The first one concerns the latent label variables for which an HOGM is employed for clustering or classification [11, 12]. The second one adopts a MEP distribution [13, 14] for pixel intensities in a given class. To the best of our knowledge, these priors have not been jointly used for image recovery and segmentation tasks. The novel model is suitable for applications dealing with multicomponent or multivariate images (e.g. multi/hyperspectral data) [15, 16, 17]. One of the advantages of the proposed approach is that it allows us to estimate the associated hyperparameters by defining suitable priors. While alternatives could solve segmentation [18, 19, 20] or reconstruction [21, 22] problems, VBA benefits from low computational cost while providing high quality results.

Section 2 introduces the HOGMep Bayesian formulation for joint restoration and segmentation. The VBA methodology solving the related problem is described in Section 3. The proposed joint recovery and segmentation approach is favorably evaluated and compared within a deconvolution context in Section 4.

2. HOGMep BAYESIAN FORMULATION

2.1. Inverse problem model

We concentrate on the standard inverse problem consisting of recovering an unknown signal $x$ from a degraded one $y$. We consider a linear model with additive noise formulated as $y = Hx + n$, where, for $(M, N, B) \in (\mathbb{N}^*)^3$, $y \in \mathbb{R}^{M\times N\times B}$ is the observed data, $x \in \mathbb{R}^{M\times N\times B}$ corresponds to the unknown signal to be recovered, $H \in \mathbb{R}^{M \times N \times B}$ represents a linear degradation operator and $n$ is a noise (statistically independent of $x$). We are interested in $B$-component images where $x = [x_1, \ldots, x_B]^\top$ and, for every pixel $i \in \{1, \ldots, N\}$, $x_i = (x_{i,1}, \ldots, x_{i,B})^\top$. Assuming a zero-mean white Gaussian noise with inverse variance (or precision) $\gamma$, the likelihood $p(y|x, \gamma)$ is given by the normal distribution $\mathcal{N}(Hx, \gamma^{-1}I)$. In conjunction with the recovery task, our objective is to perform a classification of the components of $x$. For this purpose, $L$ being the number of expected classes, the label field is encoded by a vector of hidden variables $z \in \{1, \ldots, L\}^N$.

2.2. Prior model on the unknown data and hyperpriors

HOGM for label variables [11, 12] deals with cliques of arbitrary size instead of being limited to pairwise interactions between variables. The chosen prior is:

$$p(z) \propto \exp \left( \sum_{s=1}^{S} \sum_{i \in \mathcal{V}_s} V_s(z_{i,1}, \ldots, z_{i,s}) \right),$$

where $S$ is the size of the maximal clique and, for every $s \in \{1, \ldots, S\}$, the function $V_s$ is a potential function of order $s$, and $\mathcal{V}_s$ is the set of cliques of size $s$. The model contains a prior weighting parameter $\lambda$, not explicitly written in (1).

In addition to a prior on hidden label variables, we assume that the conditional distribution of $x$ given label variables $z$ is a Multivariate Exponential Power (MEP) distribution denoted by $\mathcal{M}$. Such a prior is well suited to multicomponent images [23]. We denote by $z_i \in \{1, \ldots, L\}$ the label of the class to which the pixel indexed by $i$ belongs, and by $z = (z_1, \ldots, z_N)^\top$ the label vector for all pixels of the image. Furthermore, for every class labeled by $l \in \{1, \ldots, L\}$, $m_l$, $\Omega_l$, and $\beta_l$ denote the parameters of the MEP distribution associated with label value $l$. By assuming that $\beta_l \in (0, 1]$, the MEP distribution can be expressed as a Gaussian Scale Mixture (GSM) [24]:

$$\mathcal{M}(x_i; m_l, \Omega_l, \beta_l) = \int_{\mathbb{R}^+} \mathcal{N}(x_i; m_l, u_i^{-1}\Omega_l^{-1})p(u_l|\beta_l)du_i,$$

where $p(u_l|\beta_l)$ denotes the probability density function (pdf) of $u_l$ with shape parameter $\beta_l$ and $u = (u_1, \ldots, u_N)^\top$ gathers all the latent variables. When $\beta_l < 1$, this pdf can be expressed as a function...
of a positive alpha-stable distribution. When $\beta_i = 1$, it degenerates into a Dirac distribution [24]. As a result, we have

$$p(x_i|z_i, m, \Omega, \beta) = \prod_{i=1}^{N} p(x_i|z_i, m, \Omega, \beta),$$

where $p(x_i|z_i = l, m, \Omega, \beta) = M(x_i; m_l, \Omega_l, \beta_l)$. Our Bayesian model involves four types of hyper-parameters: inverse noise variance $\gamma$, mean variables $(m_l)_{1 \leq l \leq L}$, inverse covariance matrices $(\Omega_l)_{1 \leq l \leq L}$ and shape parameters $(\beta_l)_{1 \leq l \leq L}$. In this work, we assume that, for all the classes $l \in \{1, \ldots, L\}$, the shape parameters of the MEP distribution are identical, i.e. $\beta_1 = \cdots = \beta_L = \beta$. This single parameter is fixed in advance, according to prior knowledge. If $\mathcal{G}$ and $\mathcal{W}$ denote Gamma and Wishart distributions, to jointly estimate all the remaining parameters, we define a hyperprior on each of them:

$$p(\gamma) = \mathcal{G}(\bar{a}, \bar{b}), \quad p(m_l) = \mathcal{N}(\bar{\mu}, \Lambda), \quad p(\Omega_l) = \mathcal{W}(\Gamma, \nu),$$

Figure 1 summarizes the dependency relationships between variables involved in this hierarchical graphical model.

### 2.3. Joint posterior distribution

Using Bayes’ rule, the joint posterior distribution reads

$$p(x, u, z, \gamma, m, \Omega|y, \beta)$$

$$= p(y|x, \gamma) \prod_{i=1}^{N} \left( p(x_i|z_i, u_i, m, \Omega)p(u_i|\beta) \right) p(z)p(\gamma)$$

$$\times \prod_{l=1}^{L} p(m_l)p(\Omega_l).$$

Its intricate form is due to the dependence between the unknown variables. VBA can circumvent this issue [8, 9, 10], and is indeed shown in the following section to provide an elegant solution to the Bayesian inference.

![Fig. 1: Dependency relationships between variables.](image)

### 3. VARIATIONAL BAYESIAN APPROACH

Let $\Theta = (\Theta_j)_{1 \leq j \leq J}$ be a vector containing all the variables $(x, u, z, \gamma, m, \Omega)$ to be estimated. As previously mentioned, we want to estimate the true posterior distribution $p(\Theta|y)$. The core idea of VBA is to approximate $p(\Theta|y)$ by a separable distribution, denoted by $q(\Theta)$, such that $q(\Theta) = \prod_{j=1}^{J} q(\Theta_j)$. The minimization of the Kullback-Leibler ($KL$) divergence between $q(\Theta)$ and $p(\Theta|y)$ allows us to determine the optimal separable approximation:

$$KL(q(\Theta)||p(\Theta|y)) = \int q(\Theta) \ln \frac{q(\Theta)}{p(\Theta|y)} d\Theta.$$  

The optimal approximate distribution can be computed [25]:

$$\forall j \in \{1, \ldots, J\}, \quad q(\Theta_j) \propto \exp \left( (\ln p(y, \Theta))_{q-\Theta_j} \right),$$

where $q_{-\Theta_j} = \prod_{j' \neq j} q(\Theta_{j'})$, and $(\cdot)$ denotes the expectation value with respect to a probability distribution $q$. Implicit relations between pdfs $(q(\Theta_j))_{1 \leq j \leq J}$ generally prevent analytical expressions for $q(\Theta)$. These distributions are determined in an iterative way, by updating one of the separable components $(q(\Theta_j))_{1 \leq j \leq J}$ while fixing the others. In this work, the approximation takes the form:

$$q(\Theta) = q(\gamma) \prod_{i=1}^{N} \left( q(x_i, z_i)q(u_i) \right) \prod_{l=1}^{L} \left( q(m_l)q(\Omega_l) \right),$$

with $q(x_i, z_i) = q(x_i|z_i)q(z_i)$. Hence, using (4), for every $i \in \{1, \ldots, N\}$ and $l \in \{1, \ldots, L\}$, the optimal solutions for $q(x_i|z_i)$, $q(z_i)$, $q(m_l)$, $q(\Omega_l)$ and $q(\gamma)$ are such that

$$q(x_i|z_i = l) = N(\eta_i, \Sigma_i), \quad q(z_i = l) = \pi_{i,l},$$

$$q(m_l) = N(\mu_l, \Lambda_l), \quad q(\Omega_l) = \mathcal{W}(\Gamma_l, \nu_l), \quad q(\gamma) = \mathcal{G}(a, b).$$

Except for $q(u_i)$, the optimization of these distributions can be performed iteratively, since they belong to known parametrized distribution families. The MEP prior distribution can be expressed as a GSM. Hence, the mean value of $u_i$ can be determined without an analytical expression for $q(u_i)$ [26, 27]. This allows us to derive an approximate distribution of $x_i$, which depends on the mean value of $u_i$. In the following, assuming that $k \in \mathbb{N}$ designates the iteration number, we describe how to update these distributions by deriving closed form expressions for their parameters.

#### 3.1. Determination of model pdf $q(x_i, z_i)$

As mentioned in (6), for $l \in \{1, \ldots, L\}$, $q^{k+1}(x_i|z_i = l)$ is a Gaussian distribution. Using (4), its covariance matrix $\Sigma_{i,l}$ and mean $\eta_{i,l}$ are given, at the iteration $k + 1$, by

$$\Sigma^{k+1} = \left( \hat{\gamma}^{k} H^{k}_{i} + \tilde{u}^{k}_{i} \tilde{\Omega}^{k}_{i} \right)^{-1}$$

$$\eta^{k+1} = \left( \hat{\gamma}^{k} H^{k}_{i} \right) \left( y - \sum_{j<i} H_{j} \hat{x}_{j}^{k} - \sum_{j>i} H_{j} \hat{x}_{j}^{k} \right)$$

$$+ \tilde{u}^{k}_{i} \tilde{\Omega}^{k}_{i} \mu_{i}^{k}.$$
where ψ is the digamma function and
\[
\hat{V}_t^k = V_t(l) + \sum_{s=1}^{S-1} \left( \sum_{z_i} V_{s+1}(l, z_i) \right) \Pi_{j \neq i} q^k(z_j).
\]

For every \(j \in \{1, \ldots, N\}\), we have \(q^k_{j+1} = \sum_{l=1}^L n_{j+l} q^k_{j,l}\).

Instead of computing the \(q(u_i)\) distribution, we focus on the expectation of \(u_i\) which can be expressed as follows, thanks to the integral form of GSM [26]:
\[
\hat{u}^{k+1} = \beta \left( \sum_{l=1}^L n_{j+l} \text{tr} \left[ A_k \hat{\Omega}_k \right] \right)^{\beta-1},
\]
where \(A_k = (n_{j+l} - \mu^k_l)(\mu^k_{j+l} - \mu^k_l) + \Xi^{k+1}_{i,l} + \Lambda^k_i\).

### 3.2. Finding of hyperprior pdfs \(q(\mu_k), q(\Omega_i)\) and \(q(\gamma)\)

As shown by (6), \(q^{k+1}(\mu_k)\) is a Gaussian distribution with mean \(\mu^{k+1}_k\) and covariance matrix \(\Lambda^{k+1}_k\) expressed as
\[
\Lambda^{k+1}_k = \left( \bar{\Lambda}^{-1} + \bar{\Omega}_k \sum_{l=1}^N n_{j+l} \bar{u}_l \right)^{-1},
\]
\[
\mu^{k+1}_k = \bar{\mu} + \bar{\Omega}_k \sum_{l=1}^N n_{j+l} \bar{u}_l \cdot \bar{\gamma}^{k+1},
\]
Then, the mean value of \(q^{k+1}(\mu_k)\) is \(\hat{\mu}^{k+1} = \mu^{k+1}_k\).

The Wishart distribution \(q^{k+1}(\Omega_i)\) is parametrized by
\[
u_i = \bar{\nu}_i + \sum_{l=1}^N n_{j+l} \bar{u}_l \frac{1}{\bar{\gamma}^{k+1}},
\]
\[
\Gamma_i = \left( \sum_{l=1}^N n_{j+l} \bar{u}_l \cdot \bar{\gamma}^{k+1} \right)^{-1},
\]
where \(\bar{\Lambda}_k = (n_{j+l} - \mu^k_l)(\mu^k_{j+l} - \mu^k_l)^T + \Xi^{k+1}_{i,l} + \Lambda^k_i\).

The mean value of \(q^{k+1}(\Omega_i)\) is
\[
\hat{\Omega}_i = \bar{\nu}_i \cdot \hat{\Gamma}_i^{k+1}.
\]

The \(q^{k+1}(\gamma)\) distribution is a Gamma distribution with parameters \(a^{k+1} = \bar{a} + M/2\) and \(b^{k+1} = \bar{b} + \frac{1}{2} \|y - Hx^{k+1}\|^2 + \sum_{i=1}^N \sum_{l=1}^L n_{j+l} \text{tr} \left[ H_i^T H_i \cdot \Xi^{k+1}_{i,l} \right].\)

From standard properties of the Gamma distribution, the expectation of \(\gamma\) at the iteration \(k+1\) is equal to
\[
\hat{\gamma}^{k+1} = \frac{a^{k+1}}{b^{k+1}} = \frac{2\bar{a} + M}{2b^{k+1}}.
\]

### 3.3. Resulting algorithm

Altogether, our algorithm can be summed up as follows:

Set initial values: \(\eta^{0}_{i,l}, \Xi^{0}_{i,l}, \hat{\nu}^{0}_i, \bar{\nu}^{0}_i, \mu^{0}_l, \hat{\Gamma}^{0}_i, \nu^{0}_i, b^{0}_i, \) and set \(a^0 = \bar{a} + M/2\). Compute \(\hat{\nu}^{0}_i = \frac{1}{\nu^{0}_i} \hat{\Gamma}^{0}_i\), and set \(\eta^{0} = 1/b^{0}_i(\bar{b} + M/2)\).

For \(k = 0, 1, \ldots\)

1. Update parameters \(\Xi^{k+1}_{i,l} \) and \(\eta^{k+1}_{i,l}\) of \(q^{k+1}(x_i | z_i = l)\) using (7) and (8).
2. Update mean values \(\hat{\nu}^{k+1}_i\) of \(q^{k+1}(\nu_i)\) using (9).
3. Update parameters \(\Lambda^{k+1}_i\) and \(\mu^{k+1}_l\) of \(q^{k+1}(\mu_i)\) using (11) and (12).
4. Update parameters \(\bar{\nu}^{k+1}_i\) and \(\hat{\Gamma}^{k+1}_i\) of \(q^{k+1}(\Omega_i)\) using (13) and (14).
5. Compute \(\hat{\nu}^{k+1}_i\) from (15).

### 4. Simulation Results

The performance of HOGMep is assessed through segmentation and recovery tasks in both contexts. Experiments are performed on both synthetic and benchmark color images. Channels are degraded by distinct blur operators and additive Gaussian noises. A natural comparison is made with VB-MIG [8], in which our MEP prior is restricted to a Gaussian one. Restoration results are also compared to those obtained with the recent state-of-the-art variational approach 3MG [21, 28] and segmentation with the iterated conditional mode (ICM) algorithm [29]. For the ‘synth’ image with ground truth, results in SNR are given in Table 1a, and displayed in Figure 2 (restored channels) and 3 (color image). HOGMep consistently yields the best objective results, with crisper images.

<table>
<thead>
<tr>
<th>Channel</th>
<th>Initial</th>
<th>3MG</th>
<th>VB-MIG</th>
<th>HOGMep</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red</td>
<td>10.73</td>
<td>11.21</td>
<td>22.62</td>
<td>24.25</td>
</tr>
<tr>
<td>Green</td>
<td>12.42</td>
<td>15.17</td>
<td>16.41</td>
<td>17.14</td>
</tr>
<tr>
<td>Blue</td>
<td>3.92</td>
<td>2.60</td>
<td>18.39</td>
<td>19.75</td>
</tr>
<tr>
<td>Color</td>
<td>6.98</td>
<td>6.08</td>
<td>19.41</td>
<td>20.63</td>
</tr>
</tbody>
</table>

(a) Channel and color restoration results in terms of SNR. Best performers are in bold.

<table>
<thead>
<tr>
<th>Channel</th>
<th>Initial</th>
<th>3MG</th>
<th>VB-MIG</th>
<th>HOGMep</th>
</tr>
</thead>
<tbody>
<tr>
<td>VI</td>
<td>0.344</td>
<td>0.13</td>
<td>0.006</td>
<td></td>
</tr>
<tr>
<td>RI</td>
<td>92.72</td>
<td>98.24</td>
<td>99.98</td>
<td></td>
</tr>
<tr>
<td>% misc.</td>
<td>41.10</td>
<td>4.70</td>
<td>0.05</td>
<td></td>
</tr>
</tbody>
</table>

(b) Segmentation: variation of information [30], Rand index [31].

**Table 1**: Objective restoration and segmentation performance (‘synth’).

Segmentation performance, summarized in Table 1b, is numerically assessed through variation of information (VI, [30]) and Rand index (RI, [31]) measures. Segmented images and their binary difference to the original segmentation are displayed in Figure 4. HOGMep offers a very low rate of wrongly assigned pixels, and globally correctly labels all regions. For the ‘peppers’ image, Table 2 gathers objective SNR and SSIM performance, while restoration is depicted in Figure 5, and segmentation results in Figure 6. Through the obtained results, we conclude that HOGMep outperforms state-of-the-art methods in both recovery and segmentation problems.
Fig. 2: RGB channels (‘synth’): (a) original (b) blurred and noisy. Restored with (c) 3MG, (d) VB-MIG, (e) HOGMep.

Fig. 3: Restoration results for the color image (‘synth’).

Fig. 4: Segmentation results for ‘synth’ and binary difference to original ground truth (wrongly assigned pixels are in black).

<table>
<thead>
<tr>
<th>Channel</th>
<th>Initial</th>
<th>3MG</th>
<th>VB-MIG</th>
<th>HOGMep</th>
</tr>
</thead>
<tbody>
<tr>
<td>Red</td>
<td>SNR</td>
<td>24.30</td>
<td>28.99</td>
<td>19.94</td>
</tr>
<tr>
<td></td>
<td>SSIM</td>
<td>0.938</td>
<td>0.972</td>
<td>0.939</td>
</tr>
<tr>
<td>Green</td>
<td>SNR</td>
<td>18.36</td>
<td>22.82</td>
<td>21.13</td>
</tr>
<tr>
<td></td>
<td>SSIM</td>
<td>0.824</td>
<td>0.903</td>
<td>0.875</td>
</tr>
<tr>
<td></td>
<td>SSIM</td>
<td>0.688</td>
<td>0.748</td>
<td>0.779</td>
</tr>
<tr>
<td>Color</td>
<td>SNR</td>
<td>19.70</td>
<td>22.25</td>
<td>21.77</td>
</tr>
</tbody>
</table>

Table 2: Objective restoration performance (‘peppers’). Best performers are in bold.

Fig. 5: Restoration results for the color images (‘peppers’).

Fig. 6: Segmentation results for ‘peppers’ and the corresponding silhouette. Silhouette image: blackish pixels reflect silhouette indices close to –1 while whitish pixels correspond to scores near 1. Using this color code, brighter zones reflect satisfying clustering.

5. CONCLUSIONS

In this work, we have performed a joint image recovery and segmentation via a Bayesian approach. This formulation combines two main priors on the sought data: an HOGM for the latent label variables and a MEP distribution for the unknown signal within a given class. We have proposed an original VBA method for solving the related problem where most of the hyper-parameters are automatically estimated. The provided simulation results show that HOGMep delivers good results for both recovery and segmentation tasks.
6. REFERENCES


