# APPROXIMATIONS OF POSTERIOR DISTRIBUTIONS IN BLIND DECONVOLUTION USING VARIATIONAL METHODS

*Javier Mateos, Rafael Molina* *

Depto. de Ciencias de la Computación e I.A.
Universidad de Granada, Granada, Spain.

*Aggelos K. Katsaggelos*

Dept. of Electrical and Computer Engineering
Northwestern University, Evanston, IL, USA.

**ABSTRACT**

In this paper the blind deconvolution problem is formulated using the variational framework. With its use approximations of the involved probability distributions are developed resulting in two algorithms for the estimation of the posterior distributions of the hyperparameters, the blur, and the original image. The performance of the two proposed restoration algorithms is demonstrated experimentally.

## 1. INTRODUCTION

Blind deconvolution refers to a class of problems of the form

$$g = h * f + n, \tag{1}$$

where $f, g, h$, and $n$ represent respectively the unknown original image, the observed image, the unknown point spread function (PSF) and the observation noise, and the operator $(*)$ denotes 2-D convolution. The objective of blind deconvolution is to obtain estimates of $h$ and $f$ based on $g$ and prior knowledge about the unknown quantities and the noise.

There are two main approaches to the blind deconvolution problem [1, 2]. With the first one, the PSF is identified separately from the original image and later used in combination with one of the known image restoration algorithms, while in the second approach the estimation of the blur is incorporated into the restoration procedure. Within the second approach, most methods tackle the blind deconvolution problem by incorporating prior knowledge about the image and blur into the deconvolution process. Recently Miskin and MacKay [3], Adami [4] and Likas and Galatsanos [5] have tackled this problem using the variational approach to the approximation of probability distributions [6, 3]. This approach attempts to approximate posterior distributions using the Kullback-Leibler cross-entropy [7].

Using the variational framework to approximate probability distributions, we develop in this paper two estimates of the posterior distributions of the hyperparameters, blur, and restored image in blind deconvolution problems and re-examine the model proposed in [8] using the variational approach to distribution approximation.

The paper is organized as follows. In section 2 we describe the Bayesian modelling and inference for the blind deconvolution problem. Section 3 describes the variational approach to distribution approximation. Section 4 proposes two approximations of the posterior distribution of the image and blurring function as well as the hyperparameters, based on the variational approach. Finally, in section 5 experimental results are shown and section 6 concludes the paper.

*This work has been supported by the "Comisión Nacional de Ciencia y Tecnología" under contract TIC2003-00880.

## 2. BAYESIAN MODELLING AND INFERENCE

Our prior knowledge about the smoothness of the object luminosity distribution makes it possible to model the distribution of $f$ by a CAR [9], that is,

$$\mathrm{p}(f|\alpha_{\mathrm{im}}) \propto \alpha_{\mathrm{im}}^{N/2} \exp\{-\frac{1}{2}\alpha_{\mathrm{im}} \parallel Cf \parallel^2\}, \tag{2}$$

where $C$ denotes the Laplacian operator, $N$ is the size of the image column vector resulting from its lexicographic ordering by rows, and $\alpha_{\mathrm{im}}^{-1}$ is the variance of the Gaussian distribution.

We will use the same model for the PSF, that is

$$\mathrm{p}(h|\alpha_{\mathrm{bl}}) \propto \alpha_{\mathrm{bl}}^{M/2} \exp\{-\frac{1}{2}\alpha_{\mathrm{bl}} \parallel Ch \parallel^2\}, \tag{3}$$

where $M$ is the assumed known size of the support of the blur, $h$ is a column vector of size $N$, and $\alpha_{\mathrm{bl}}^{-1}$ is the variance of the Gaussian distribution.

A simplified but realistic degradation model for blind deconvolution is the one defined in Eq. (1) with the unknown blur approximated by a block circulant matrix and the Gaussian noise $n(x)$ having zero mean and variance $\beta^{-1}$.

Then, the probability of the observed image $g$ if $f$ and $h$ were respectively the 'true' image and blur is equal to

$$\begin{aligned} \mathrm{p}(g \mid f, h, \beta) &\propto \beta^{N/2} \exp\left[-\frac{1}{2}\beta \|g - Hf\|^2\right] \\ &\propto \beta^{N/2} \exp\left[-\frac{1}{2}\beta \|g - Fh\|^2\right], \end{aligned}$$

where we have used $H$ to denote the $N \times N$ blurring matrix corresponding to the blurring vector $h$ and $F$ to denote the $N \times N$ convolution matrix corresponding to $f$.

An important problem arises when $\alpha_{\mathrm{im}}$, $\alpha_{\mathrm{bl}}$, and $\beta$ are unknown. To deal with the estimation of these hyperparameters the hierarchical Bayesian paradigm introduces a second stage (the first stage consists of the formulation of $\mathrm{p}(g|f,h,\beta)$, $\mathrm{p}(f|\alpha_{\mathrm{im}})$ and of $\mathrm{p}(h|\alpha_{\mathrm{bl}})$). In this stage the hyperprior $\mathrm{p}(\Omega)$ on the so called hyperparameter vector, $\Omega = (\alpha_{\mathrm{im}}, \alpha_{\mathrm{bl}}, \beta)$, is also formulated, resulting in $\mathrm{p}(\Theta) = \mathrm{p}(\alpha_{\mathrm{im}}, \alpha_{\mathrm{bl}}, \beta)\mathrm{p}(f|\alpha_{\mathrm{im}})\mathrm{p}(h|\alpha_{\mathrm{bl}})$, with $\Theta = (\alpha_{\mathrm{im}}, \alpha_{\mathrm{bl}}, \beta, f, h)$ the set of unknown in our blind deconvolution problem.

Let us assume that each of the hyperparameters has the following gamma distribution

$$\mathrm{p}(\omega) = \Gamma(\omega|a_\omega, b_\omega) = \frac{b^a}{\Gamma(a)}\omega^{a-1}\exp[-b\,\omega], \tag{4}$$

where $\omega > 0$ denotes a hyperparameter, i.e., $\omega \in \Omega$, $b > 0$ is the inverse scale parameter and $a > 0$ is the shape parameter. This distribution has the following properties

$$E[w] = \frac{a}{b} \qquad \text{and} \qquad Var[w] = \frac{a}{b^2}. \tag{5}$$

The Bayesian paradigm dictates that inference on $\Theta$ should be based on $\mathrm{p}(\Theta \mid g)$ given by

$$\mathrm{p}(\Theta \mid g) = \frac{\mathrm{p}(g \mid \Theta)\mathrm{p}(\Theta)}{\mathrm{p}(g)} \propto \mathrm{p}(g \mid \Theta)\mathrm{p}(\Theta). \tag{6}$$

## 3. VARIATIONAL APPROACH TO BAYESIAN INFERENCE

Since $\mathrm{p}(\Theta \mid g)$ cannot be easily calculated, following the variational approach to distribution approximation, we approximate now this posterior distribution $\mathrm{p}(\Theta \mid g)$ by the distribution

$$\mathrm{q}(\Theta) = \mathrm{q}(\Omega)\mathrm{q}(f)\mathrm{q}(h), \tag{7}$$

where $\mathrm{q}(f)$ and $\mathrm{q}(h)$ denote distributions on $f$ and $h$, respectively, and $\mathrm{q}(\Omega)$ is usually given by

$$\mathrm{q}(\Omega) = \mathrm{q}(\alpha_{\mathrm{im}})\mathrm{q}(\alpha_{\mathrm{bl}})\mathrm{q}(\beta). \tag{8}$$

Note that once $\mathrm{p}(\Theta \mid g)$ has been approximated by $\mathrm{q}(\Theta)$ we can use $\mathrm{q}(\Omega)$, $\mathrm{q}(f)$, and $\mathrm{q}(h)$ to select single values of the hyperparameters, blur, and image (for instance their mean or mode values), but more importantly we can also examine the quality of those estimates by studying their corresponding distributions.

The criterion we use to find $\mathrm{q}(\Theta)$ is the minimization of the Kullback-Leibler divergence [7]:

$$C_{KL}(\mathrm{q}(\Theta) \parallel \mathrm{p}(\Theta|g)) = \int_\Theta \mathrm{q}(\Theta)\log\left(\frac{\mathrm{q}(\Theta)}{\mathrm{p}(\Theta, g)}\right)d\Theta + \text{const.} \tag{9}$$

We note that this quantity is always non negative and takes the value zero only when $\mathrm{q}(\Theta) = \mathrm{p}(\Theta|g)$.

For $\theta \in \Theta$, let us denote by $\Theta_\theta$ the subset of $\Theta$ with $\theta$ removed; for instance, if $\theta = f$, $\Theta_f = (\Omega, h)$. Then we can write

$$\begin{aligned} C_{KL}(\mathrm{q}(\theta)\mathrm{q}(\Theta_\theta) \parallel \mathrm{p}(\Theta|g)) &= \text{const}+ \\ \int_\theta \mathrm{q}(\theta)\left(\int_{\Theta_\theta}\mathrm{q}(\Theta_\theta)\log\left(\frac{\mathrm{q}(\theta)\mathrm{q}(\Theta_\theta)}{\mathrm{p}(\theta, \Theta_\theta, g)}\right)d\Theta_\theta\right)d\theta. \end{aligned} \tag{10}$$

Given $\mathrm{q}(\Theta_\theta)$, in order to obtain

$$\hat{\mathrm{q}}(\theta) = \arg\min_{\mathrm{q}(\theta)} C_{KL}\left(\mathrm{q}(\theta)\mathrm{q}(\Theta_\theta)\right) \parallel \mathrm{p}(\Theta|g)), \tag{11}$$

we differentiate Eq. (10) with respect to $\mathrm{q}(\theta)$ to obtain, see Eq. (2.28) of [3],

$$\hat{\mathrm{q}}(\theta) = \text{const} \times \exp\left(E\left[\,\log \mathrm{p}(\Theta)\mathrm{p}(g|\Theta)\,\right]_{\mathrm{q}(\Theta_\theta)}\right), \tag{12}$$

where

$$E\left[\log \mathrm{p}(\Theta)\mathrm{p}(g|\Theta)\right]_{\mathrm{q}(\Theta_\theta)} = \int \log \mathrm{p}(\Theta)\mathrm{p}(g \mid \Theta)\mathrm{q}(\Theta_\theta)d\Theta_\theta. \tag{13}$$

The above equations suggest the following iterative procedure to find $\mathrm{q}(\Theta)$.

**Algorithm 1**

*Given $q^1(h)$ and $q^1(\Omega)$, current estimates of the distributions $q(h)$, $q(\Omega)$,*

*For $k = 1, 2, \ldots$ until a stopping criterion is met:*

1. *Find* $q^k(f) = \arg\min_{q(f)} C_{KL}(q^k(\Omega)q(f)q^k(h) \parallel p(\Theta|g))$,
2. *Find* $q^{k+1}(h) = \arg\min_{q(h)} C_{KL}(q^k(\Omega)q^k(f)q(h) \parallel p(\Theta|g))$,
3. *Find* $q^{k+1}(\Omega) = \arg\min_{q(\Omega)} C_{KL}(q(\Omega)q^k(f)q^{k+1}(h) \parallel p(\Theta, h|g))$

## 4. VARIATIONAL APPROXIMATIONS OF THE POSTERIOR DISTRIBUTION FOR BLIND DECONVOLUTION PROBLEMS

In the previous section we have studied how the variational approach could be used in general to tackle the blind deconvolution problem. We now see how the method performs in practice.

CASE I: Let us first assume that $\Omega = (\alpha_{\mathrm{im}}, \alpha_{\mathrm{bl}}, \beta)$ is known and proceed to estimate $\mathrm{q}(f)$ and $\mathrm{q}(h)$. Let us now assume that $\mathrm{q}^k_{BR}(h) = \mathcal{N}(h|\ E^k_{BR}(h), cov^k_{BR}(h))$, where we are using the subscript $BR$ to denote that the distributions on $h$ (an $f$) are random and $E[\cdot]$ and $cov(\cdot)$ denote respectively the mean and covariance of the normal distribution. Then, from Eq. (12) we have that

$$\mathrm{q}^k_{BR}(f) = \mathcal{N}\left(f \mid E^k_{BR}(f), cov^k_{BR}(f)\right), \tag{14}$$

In order to calculate the mean and covariance of the normal distribution we note that they are given by

$$\frac{\partial 2\log \mathrm{q}^k_{BR}(f)}{\partial f} = 0, \quad -\frac{\partial^2 2\log \mathrm{q}^k_{BR}(f)}{\partial f^2} = [cov^k_{BR}(f)]^{-1}, \tag{15}$$

and so we obtain

$$\begin{aligned} E^k_{BR}(f) = &\left(\alpha_{\mathrm{im}}C^tC + \beta E^k_{BR}(H)^t E^k_{BR}(H) + \beta cov^k_{BR}(h)\right)^{-1}\times \\ &\beta E^k_{BR}(H)^t g, \end{aligned} \tag{16}$$

$$cov^k_{BR}(f) = \left(\alpha_{\mathrm{im}}C^tC + \beta E^k_{BR}(H)^t E^k_{BR}(H) + \beta cov^k_{BR}(h)\right)^{-1}. \tag{17}$$

Once $\mathrm{q}^k(f)$ has been calculated, following the same steps we obtain that

$$\mathrm{q}^{k+1}_{BR}(h) = \mathcal{N}\left(h \mid E^{k+1}_{BR}(h), cov^{k+1}_{BR}(h)\right), \tag{18}$$

with

$$\begin{aligned} E^{k+1}_{BR}(h) = &\left(\alpha_{\mathrm{bl}}C^tC + \beta E^k_{BR}(F)^t E^k_{BR}(F) + \beta[cov^k_{BR}(f)]\right)^{-1}\times \\ &\beta E^k_{BR}(F)^t g, \end{aligned} \tag{19}$$

$$cov^{k+1}_{BR}(h) = \left(\alpha_{\mathrm{bl}}C^tC + \beta E^k_{BR}(F)^t E^k_{BR}(F) + \beta[cov^k_{BR}(f)]\right)^{-1}. \tag{20}$$

Although from Eq. (12) we know how to calculate the best forms of $\mathrm{q}(f)$ and $\mathrm{q}(h)$, in K-L divergence sense, we also examine the case of suboptimal choices of these distributions by assuming that $\mathrm{q}(f)$ and $\mathrm{q}(h)$ are degenerate distributions, that is, probability distributions assigning probability one to just one value.

Following the same steps as in the previous section and taking into account that the distributions on $f$ and $h$ are degenerate, we can calculate $E^k_{BD}(f)$ and $E^{k+1}_{BD}(h)$ (we use the subscript $BD$ to denote that the distributions of $f$ and $h$ are both deterministic) by simply setting $cov_{BR}(f)$ and $cov_{BR}(h)$ to zero in Eqs. (14)–(20).

It is very interesting to note that the iterative procedure resulting in this case is the same as the one proposed in [8] to jointly find the image and blurring functions in blind deconvolution problems.

CASE II: Let us now assume that the hyperparameter vector $\Omega = (\alpha_{\rm bl}, \alpha_{\rm im}, \beta)$ is unknown and try to estimate it. Then, in order to find $q^{k+1}(\omega)$, $\omega \in \Omega$, in step 3 of Algorithm 1 we have $\mathrm{q}^{k+1}(\omega) = \Gamma(\omega | a^{k+1}_\omega, b^{k+1}_\omega)$, where the parameters $a^{k+1}_\omega$ and $b^{k+1}_\omega$ defining $\mathrm{q}^{k+1}(\omega)$ are given by

$$a^{k+1}_{\alpha_{\rm im}} = a_{\alpha_{\rm im}} + \frac{N}{2} \quad (21)$$

$$b^{k+1}_{\alpha_{\rm im}} = b_{\alpha_{\rm im}} + \frac{1}{2} E\left[\| Cf \|^2\right]_{\mathrm{q}^k(f)} \quad (22)$$

$$a^{k+1}_{\alpha_{\rm bl}} = a_{\alpha_{\rm bl}} + \frac{M}{2} \quad (23)$$

$$b^{k+1}_{\alpha_{\rm bl}} = b_{\alpha_{\rm bl}} + \frac{1}{2} E\left[\| Ch \|^2\right]_{\mathrm{q}^{k+1}(h)} \quad (24)$$

$$a^{k+1}_{\beta} = a_{\beta} + \frac{N}{2} \quad (25)$$

$$b^{k+1}_{\beta} = b_{\beta} + \frac{1}{2} E\left[\| g - Hf \|^2\right]_{\mathrm{q}^k(f)\mathrm{q}^{k+1}(h)} \quad (26)$$

The process of re-estimating the image and blur once $q^{k+1}(\omega)$, $\omega \in \Omega$, has been found coincides with the previously described one by Eqs. (16), (17), (19), and (20) using now as known hyperparameters the mean values of the distributions $\mathrm{q}^{k+1}(\omega)$. These mean values are given by (see Eq. (5)),

$$E[\alpha_{\rm im}]^{-1}_{q^{k+1}(\alpha_{\rm im})} = \gamma_{\alpha_{\rm im}} \frac{1}{\overline{\alpha}_{\rm im}} + (1-\gamma_{\alpha_{\rm im}}) \frac{E\left[\| Cf \|^2\right]_{\mathrm{q}^k(f)}}{N} \quad (27)$$

$$E[\alpha_{\rm bl}]^{-1}_{q^{k+1}(\alpha_{\rm bl})} = \gamma_{\alpha_{\rm bl}} \frac{1}{\overline{\alpha}_{\rm bl}} + (1-\gamma_{\alpha_{\rm bl}}) \frac{E\left[\| Ch \|^2\right]_{\mathrm{q}^{k+1}(h)}}{M} \quad (28)$$

$$E[\beta]^{-1}_{q^{k+1}(\beta)} = \gamma_{\beta} \frac{1}{\overline{\beta}} + (1-\gamma_{\beta}) \frac{E\left[\| g - Hf \|^2\right]_{\mathrm{q}^k(f)\mathrm{q}^{k+1}(h)}}{N} \quad (29)$$

where $\overline{\omega}$, $\omega \in \Omega$, is the mean of the prior distribution of $\omega$, that is, $\overline{\omega} = a_\omega / b_\omega$, see Eqs. (4) and (5), and

$$\gamma_{\alpha_{\rm im}} = \frac{a_{\alpha_{\rm im}}}{a_{\alpha_{\rm im}} + \frac{N}{2}}, \;\; \gamma_{\alpha_{\rm bl}} = \frac{a_{\alpha_{\rm bl}}}{a_{\alpha_{\rm bl}} + \frac{M}{2}}, \;\; \gamma_{\beta} = \frac{a_{\beta}}{a_{\beta} + \frac{N}{2}}. \quad (30)$$

We note now that $\gamma_{\alpha_{\rm im}}$, $\gamma_{\alpha_{\rm bl}}$, and $\gamma_\beta$ can be understood as normalized confidence parameters taking values in the interval $[0, 1)$. That is, when they are zero no confidence is put on the given parameters $\overline{\alpha}_{\rm im}, \overline{\alpha}_{\rm bl}$ and $\overline{\beta}$, while when the corresponding normalized confidence parameter is one it fully enforces the prior knowledge of the mean (no estimation of the hyperparameters is performed).

The only remaining task is the calculation of $E\left[\| Cf \|^2\right]_{\mathrm{q}^k(f)}$, $E\left[\| Ch \|^2\right]_{\mathrm{q}^{k+1}(h)}$, and $E\left[\| g - Hf \|^2\right]_{\mathrm{q}^k(f)\mathrm{q}^{k+1}(h)}$ in Eqs. (27)–(29). For the case $h$ and $f$ are random we have

$$E\left[\| Cf \|^2\right]_{\mathrm{q}^k_{BR}(f)} = \| CE^k_{BR}(f) \|^2 + \mathrm{trace}(C^t C cov^k_{BR}(f))$$


**Fig. 1**. (a) Original "Lena" image, (b) Degraded image, (c) Restoration by MLE method. Restorations for (d) BR approximation with $\gamma_\omega = 0.0$, $\omega \in \Omega$, (e) BD approximation for $\gamma_\beta = 1.0$, $\gamma_{\alpha_{\rm im}} = 1.0$ and $\gamma_{\alpha_{\rm bl}} = 0.8$, and (f) BR approximation for $\gamma_\beta = 1.0$, $\gamma_{\alpha_{\rm im}} = 0.5$ and $\gamma_{\alpha_{\rm bl}} = 1.0$.

$$E\left[\| Ch \|^2\right]_{\mathrm{q}^{k+1}_{BR}(h)} = \| CE^{k+1}_{BR}(h) \|^2 + \mathrm{trace}(C^t C cov^k_{BR}(h))$$

$$\begin{aligned} E\left[\| g - Hf \|^2\right]_{\mathrm{q}^k_{BR}(f)\mathrm{q}^{k+1}_{BR}(h)} &= \| g - E^{k+1}_{BR}(h) E^k_{BR}(f) \|^2 \\ &+ \mathrm{trace}(cov^k_{BR}(f) cov^{k+1}_{BR}(h)) \\ &+ \mathrm{trace}(E^k_{BR}(F)^t E^k_{BR}(F) cov^{k+1}_{BR}(h)) \\ &+ \mathrm{trace}(E^{k+1}_{BR}(H)^t E^{k+1}_{BR}(H) cov^k_{BR}(f)) \end{aligned}$$

where $E^k_{BR}(f)$, $cov^k_{BR}(f)$, $E^{k+1}_{BR}(h)$ and $cov^{k+1}_{BR}(h)$ have been defined in Eqs. (16), (17), (19), and (20), respectively.

For the case $q(f)$ and $q(h)$ are degenerate distributions, the covariances in the above equations are set equal to zero.

## 5. EXPERIMENTAL RESULTS

A number of experiments have been performed with the proposed methods using several images and PSFs. The results of some of them are presented here. Since we have developed two different approximations of the conditional distributions of $f$ and $h$, namely BR and BD, and the hyperparameters, $\Omega$, given the observed image, $g$, we will present results for both of them on a synthetically degraded image.

For the experiments, the "Lena" image (depicted in Fig. 1(a)) was blurred with a Gaussian shaped PSF with variance $9$. Gaussian noise of variance $\beta^{-1} = 16$ was added to obtain the degraded image depicted in Fig. 1(b) with a resulting signal-to-noise ratio (SNR) of 20dB that was used as input image for the proposed methods.

For comparison we calculated the restoration obtained by the *Maximum Likelihood Estimator* (MLE) described in [10] to simultaneously estimate the parameters and restoration assuming that the blur is known, resulting in the restored image presented in Fig. 1(c). The estimated value for the parameters were $E[\beta] = 1/15.93$ and $E[\alpha_{\rm im}] = 1/684.9$ resulting in an improvement in SNR (ISNR) of $2.48$dB.

The initial distributions in the BR algorithm were chosen as follows: The observed image was used as initial estimation for $E^0(f)$. A Gaussian distribution with mean, $E^1(h)$, a Gaussian shaped PSF with variance $0.009$ and covariance matrix equal to

**Table 1**. Some representative values obtained for the parameters and ISNR for the Lena image using $\overline{\beta} = 1/15$, $\overline{\alpha}_{\rm im} = 1/206$ $\overline{\alpha}_{\rm bl} = 4 \times 10^8$ and different values of $\gamma_\omega, \omega \in \Omega$.

BD approximation

| $\gamma_\beta$ | $\gamma_{\alpha_{\rm im}}$ | $\gamma_{\alpha_{\rm bl}}$ | $E[\beta]$ | $E[\alpha_{\rm im}]$ | $E[\alpha_{\rm bl}]$ | ISNR (dB) |
|---|---|---|---|---|---|---|
| 0.0 | 0.0 | 0.0 | 1/1949.5 | $6.1 \times 10^7$ | $1.6 \times 10^{22}$ | -8.36 |
| 0.0 | 1.0 | 0.8 | 1/15.2 | 1/206 | $3.3 \times 10^8$ | 2.09 |
| 1.0 | 1.0 | 0.8 | 1/15 | 1/206 | $3.3 \times 10^8$ | 2.11 |

BR approximation

| $\gamma_\beta$ | $\gamma_{\alpha_{\rm im}}$ | $\gamma_{\alpha_{\rm bl}}$ | $E[\beta]$ | $E[\alpha_{\rm im}]$ | $E[\alpha_{\rm bl}]$ | ISNR (dB) |
|---|---|---|---|---|---|---|
| 0.0 | 0.0 | 0.0 | 1/16.2 | 1/794 | $5.2 \times 10^8$ | 0.88 |
| 0.0 | 0.6 | 1.0 | 1/16.5 | 1/217 | $4 \times 10^8$ | 2.08 |
| 1.0 | 0.5 | 1.0 | 1/15 | 1/223 | $4 \times 10^8$ | 2.11 |

zero was chosen as initial distribution for the PSF, $q^1(h)$. To obtain a value for $\overline{\beta}$ and $\overline{\alpha}_{\rm im}$ we run the MLE method in [10] using a Gaussian shaped PSF with variance $4$ as 'true' PSF. Also, a value of $\overline{\alpha}_{\rm bl} = 4 \times 10^8$ was empirically chosen. We chose a set of eleven values ranging from $0$ to $1$ for the confidence parameters, $\gamma_\omega, \omega \in \Omega$.

The initial values $E^1[\beta]$, $E^1[\alpha_{\rm im}]$ and $E^1[\alpha_{\rm bl}]$ were then calculated using Eqs. (27)–(29), assuming a BD approximation and using $E^0(f)$ and $E^1(h)$ as described above. For all the experiments, the criterion $\parallel E^k(f) - E^{k-1}(f) \parallel^2 / \parallel E^{k-1}(f) \parallel^2 < 10^{-4}$ was used for terminating the iterative procedure.

Results are summarized in Table 1. This table shows that the BD approximation converges to the trivial solution, $E(f) = const$, $E(h) = const$, when the confidence parameters are close to zero while the BR approximation does not show this behavior.

When the solution is forced to be apart from the trivial solutions (using a confidence value on $\overline{\alpha}_{\rm im}$ and $\overline{\alpha}_{\rm bl}$ greater that zero) both methods give useful and very similar solutions in terms of ISNR and visual quality.

The noise parameter $E[\beta]$ is always accurately estimated regardless of the confidence on the parameters value (except when the trivial solution $E(f) = const$, $E(h) = const$ is obtained). The introduction of information on the expected value of this parameter does not produce a significant increase on the ISNR.

Note that the BD approximation is not able to estimate $E[\alpha_{\rm im}]$ and $E[\alpha_{\rm bl}]$ accurately if no information about the parameter value is introduced. However, the BR approximation provides a good solution when all confidence parameters $\gamma_\omega$, $\omega \in \Omega$, are set equal to zero, although better results are obtained if some information about the expected value of $E[\alpha_{\rm im}]$ and especially $E[\alpha_{\rm bl}]$ is provided. Note also that the obtained values for $E[\alpha_{\rm im}]$ when no information about the parameters value is provided is similar to the ones obtained by the MLE method for the known PSF case. In our experiments we found that there is almost no ISNR variation when moving the noise parameter confidence, $\gamma_\beta$, from 0 to 1, while increasing the confidence parameters on $\overline{\alpha}_{\rm im}$ and especially on $\overline{\alpha}_{\rm bl}$ also increases the ISNR.

Restorations for different values of the confidence parameters are depicted in Fig. 1(d)-(f). A cut of the central part of their corresponding PSF estimates is shown in Fig. 2. From the displayed images it is clear that all approximations give almost indistinguishable restorations, visually and from an ISNR point of view and quite close to the restoration obtained by MLE (see Fig. 1(e)-(f)). All the estimated PSFs for those cases approximate quite well the real PSF. On the other hand, when using $\gamma_\omega = 0$, $\omega \in \Omega$ with the

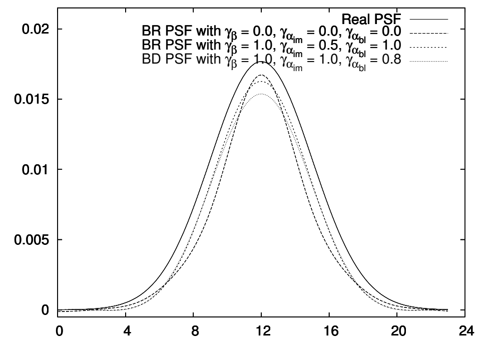


**Fig. 2**. A central cut of the PSFs for the Lena image.

BD approximation, the method converges to the trivial solution $E[f] \sim const$, $E[h] \sim const$. However, the BR approximation provides a valid solution (see fig. 1(d)) although it is noisier and not all the blur has been successfully removed.

## 6. CONCLUSIONS

Two new methods for the simultaneous estimation of the image, blur, and unknown hyperparameters in blind deconvolution problems have been proposed, based on the variational approach to distribution approximation. Using this approach we can approximate the posterior distribution of the image and blurring function, as well as, the unknown hyperparameters. The proposed methods have been validated experimentally.

## 7. REFERENCES


[1] D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Sig. Proc. Mag.*, vol. 13, no. 3, pp. 43 – 64, 1996.

[2] D. Kundur and D. Hatzinakos, "Blind image deconvolution revisited," *IEEE Sig. Proc. Mag.*, vol. 13, no. 6, pp. 61 – 63, 1996.

[3] J. Miskin, *Ensemble Learning for Independent Component Analysis*, Ph.D. thesis, Astrophysics Group, University of Cambridge, 2000.

[4] K.Z. Adami, "Variational methods in bayesian deconvolution," *PhyStat2003, eConf C030908*, 2003.

[5] A. C. Likas and N. P. Galatsanos, "A variational approach for bayesian blind image deconvolution," *IEEE Trans. on Signal Processing*, vol. 52, no. 8, pp. 2222 – 2233, 2004.

[6] M. I. Jordan, Z. Ghahramani, T. S. Jaakola, and L. K. Saul, *An Introduction to Variational Methods for Graphical Models*, pp. 105 – 162, MIT Press, 1998.

[7] S. Kullback, *Information Theory and Statistics*, Dover Publications, 1959.

[8] Y. L. You and M. Kaveh, "A regularization approach to joint blur and image restoration," *IEEE Trans. on Image Processing*, vol. 5, no. 3, pp. 416 – 428, 1996.

[9] B. D. Ripley, *Spatial Statistics*, pp. 88 – 90, John Wiley, 1981.

[10] R. Molina, "On the hierarchical Bayesian approach to image restoration. applications to astronomical images," *IEEE Trans. on Pat. Anal. and Mach. Intell.*, vol. 16, no. 11, pp. 1122 – 1128, 1994.