A BAYESIAN APPROACH TO BLIND DECONVOLUTION BASED ON DIRICHLET DISTRIBUTIONS

Rafael Molina[1] *Aggelos K. Katsaggelos*[2] *Javier Abad*[1] *Javier Mateos*[1] *

[1]Departamento de Ciencias de la Computación e I.A., Universidad de Granada, 18071 Granada, España.
[2]Department of Electrical and Computer Engineering, Northwestern University, Evanston, Illinois 60208-3118

ABSTRACT

This paper deals with the simultaneous identification of the blur and the restoration of a noisy and blurred image. We propose the use of Dirichlet distributions to model our prior knowledge about the blurring function together with smoothness constraints on the restored image to solve the blind deconvolution problem. We show that the use of Dirichlet distributions offers a lot of flexibility in incorporating vague or very precise knowledge about the blurring process into the blind deconvolution process. The proposed MAP estimator offers additional flexibility in modeling the original image. Experimental results demostrate the performance of the proposed algorithm.

1. INTRODUCTION

The blind deconvolution problem refers to a class of problems of the form

$$g(x) = d(x) * f(x) + n(x), \qquad x = (x_1, x_2) \in \Omega, \qquad (1)$$

where $\Omega \subset \mathcal{R}^2$ is the support of the image, and $f(x), g(x), d(x)$ and $n(x)$ represent respectively the original image, the observed image, blur operator (PSF) and observation Gaussian independent noise with variance σ_n^2. The operator $(*)$ in equation (1) denotes $2-D$ convolution given by

$$d(x) * f(x) = \sum_{s \in \mathcal{D}} d(s) f(x-s), \qquad (2)$$

where $\mathcal{D} \subset \mathcal{R}^2$ is the support of $d(s)$. We shall also assume that the support of the PSF is known and that the PSF is centered around zero. $k+1$ shall denote the size, number of pixels, of the support of the PSF.

In classical image restoration the blurring function is given, and the degradation process is inverted using one of the many known restoration algorithms. The various approaches that have appeared in the literature depend on the particular degradation and image models [1].

There are two main approaches to the blind deconvolution problem [2, 3]. With one of them, the PSF is identified separately from the original image, in order to use it later with one of the known image restoration algorithms while, with the other, the identification procedure is incorporated into the restoration procedure.

Recently You and Kaveh [7] have proposed a new method that incorporates smoothness constraints in both the image and blur models. This new method assumes piecewise smoothness of the PSF, that is modeled as a Gaussian process.

In this paper we propose a method based on the Bayesian approach to the blind deconvolution problem. We shall use a Dirichlet distribution as the blur model and show how to incorporate from very vague to very precise knowledge about the blurring process in the blind deconvolution problem.

The paper is organized as follows. In section 2. we describe the paradigm to be used to address the blind deconvolution problem, together with the image and noise models. The blur model is then described in section 3.. Section 4. describes the MAP estimator of the original image and the blurring function at the same time. Finally, in section 5. experimental results are shown and section 6. concludes the paper.

2. BAYESIAN PARADIGM

The Bayesian paradigm dictates that the inference about the true f and d should be based on $p(d, f \mid g)$ given by

$$\begin{aligned} p(d, f \mid g) &= p(g \mid f, d) p(f) p(d) / p(g) \\ &\propto p(g \mid f, d) p(f) p(d). \end{aligned} \qquad (3)$$

Maximization of equation (3) with respect to f and d yields

$$\hat{f}, \hat{d} = \arg\max_{f,d} p(f, d \mid g), \qquad (4)$$

the maximum-a-posteriori (MAP) estimator. Let us then proceed to study the prior and degradation models we propose to use for this problem. Although the theory we shall develop here can be applied to any quadratic prior on f, we shall study in full two particular priors [4].

Our prior knowledge about the smoothness of the object luminosity distribution makes it possible to model the distribution of f by a conditional autoregression [5] (CAR). Thus,

$$p(f|\alpha) \propto \exp\left\{-\frac{1}{2}\alpha\, f^t (I - \phi N) f\right\},$$

where $N_{ij} = 1$ if cells i and j are spatial neighbors (pixels at distance one), zero otherwise and ϕ is equal to 0.25. The

*The work of R. Molina, J. Mateos and J. Abad has been supported by the "Comisión Nacional de Ciencia y Tecnología" under contract PB93-1110.

term $f^t(I-\phi N)f$ represents in matrix notation the sum of squares of the values f_i minus ϕ times the sum of $f_i f_j$ for neighboring pixels i and j (we denote by f_i the i-th entry of the vector f).

We shall also use the simultaneous autoregressive model (SAR). This model is characterized by

$$f_i - \phi \sum_{j \text{ nhbr } i} f_j = \epsilon_i,$$

with ϵ_i independent and zero mean Gaussian with variance α^{-1}. Then, the corresponding distribution is

$$p(f|\alpha) \propto \exp\{-\frac{1}{2}\alpha\, f^t C^t C\, f\}. \tag{5}$$

A simplified but realistic degradation model for many applications is the one defined in equation (1) with Gaussian noise, that is,

$$g(i) = (Df)(i) + n(i) = \sum_j d(i-j) f(j) + n(i),$$

where D is the $p \times p$ matrix defining the systematic blur, assumed to be unknown and approximated by a block circulant matrix, $n(i)$ is the additive Gaussian noise with zero mean and variance σ_n^2 and $d(j)$ are the coefficients defining the blurring function.

Then, the probability of the observed image g if f were the 'true' image and D the blurring matrix is

$$p(g \mid f) \propto \exp\left[-\frac{1}{2\sigma_n^2}\|g - Df\|^2\right]. \tag{6}$$

3. BLUR MODEL

Having defined the image and degradation models, let us examine the blur model. A different way is proposed in this paper to impose a smoothness constraint on the blur. First, we give a few definitions and properties of the Dirichlet distribution which is proposed to be used as the blur prior (see, for instance, [6])

The k-variate analogue of the *beta distribution* is the distribution having the density function

$$\begin{aligned} p(d(1),\ldots,d(k)) \quad = \quad & \frac{\Gamma(\mu_0+\mu_1+\mu_2+\cdots+\mu_k)}{\Gamma(\mu_0)\Gamma(\mu_1)\cdots\Gamma(\mu_k)} \times \\ & (1-d(1)-\cdots-d(k))^{\mu_0-1} \times \\ & d(1)^{\mu_1-1}\cdots d(k)^{\mu_k-1}, \end{aligned} \tag{7}$$

at any point in the simplex

$$\begin{aligned} S_k \quad = \quad & (d(1),\ldots,d(k)) : d(i) \geq 0, \qquad i=1,\ldots,k, \\ & \sum_{i=1}^{k} d(i) \leq 1, \end{aligned} \tag{8}$$

and zero outside, where the μ_i are all real and positive, and $\Gamma(\cdot)$ denotes the gamma distribution. We shall refer to a distribution having the density function given in equation (7) as the *k-variate Dirichlet distribution* $D(\mu_1 \ldots \mu_k; \mu_0)$. It is important to note that we have selected $d(0)$ as the part

d(4)	d(3)	d(2)
d(5)	d(0)	d(1)
d(6)	d(7)	d(8)

Figure 1. Blurring coefficients.

of the blurring function defined by $1-\sum_{i=1}^{k} d(i)$ and that, obviously, $\sum_{i=0}^{k} d(i) = 1$ for this distribution.

The next step is to examine some of the properties of this distribution. We have the following expressions for the mean, $E[d(i)]$, the variance, $var[d(i)]$ an the covariances, $cov[d(i), d(j)]$, respectively

$$E[d(i)] = \frac{\mu_i}{\mu_0+\mu_1+\cdots+\mu_k} \qquad i=1,\ldots,k,$$

$$\begin{aligned} & var[d(i)] = \\ & \quad \frac{\mu_i(\mu_0+\mu_1+\cdots+\mu_k-\mu_i)}{(\mu_0+\mu_1+\cdots+\mu_k)^2(\mu_0+\mu_1+\cdots+\mu_k+1)}, \\ & \quad i=1,\ldots,k, \end{aligned}$$

and, for $i, j = 1, \ldots, k$,

$$\begin{aligned} & cov[d(i), d(j)] = \\ & \quad -\frac{\mu_i\mu_j}{(\mu_0+\mu_1+\cdots+\mu_k)^2(\mu_0+\mu_1+\cdots+\mu_k+1)}. \end{aligned}$$

The above expressions also hold for $d(0)$, since $d(0) = 1 - \sum_{i=1}^{k} d(i)$.

The following result can be used to simulate conditional distributions. If $(d(1),\ldots,d(k))$ is a vector random variable having the *k-variate Dirichlet distribution* $D(\mu_1,\ldots,\mu_k;\mu_0)$ then the marginal distribution of $(d(1),\ldots,d(k_1))$, $k_1 < k$, is the *k_1-variate Dirichlet distribution* $D(\mu_1,\ldots,\mu_{k_1};\mu_0+\mu_{k_1+1}+\cdots+\mu_k)$. More interesting properties of this distribution can be found in [6].

Let us now interpret the parameters involved in the Dirichlet distribution. It is clear that modifying the values of $\mu_i, i = 0,\ldots,k$, we are able to model our prior knowledge about the blurring process. For instance, $\mu_i = \text{const}, i = 0,\ldots,k$, could be used to model an out of focus and also a motion degradation. Furthermore is we write $\mu_i = \epsilon \times \text{const}, i = 0,\ldots,k$, the expected mean values are the same but the uncertainty of the model, the variance of each $d(i)$, changes continuously with ϵ. Furthermore if for two indices i and j, μ_i is greater than μ_j we are expecting $d(i)$ to be greater than $d(j)$ although by modifying the variances we include our uncertainty on the model.

It is also clear that the Dirichlet distribution can be used in situations where we know, for instance, that all pixels at the same distance from the origin have the same weight (see figure 1). Defining $\tilde{d}(0) = d(0)$, $\tilde{d}(1) = d(1) + d(3) + d(5) + d(7)$ and $\tilde{d}(2) = d(2) + d(4) + d(6) + d(8)$ we can use a Dirichlet distribution for $\tilde{d}(0), \tilde{d}(1)$ and $\tilde{d}(2)$ and then obtain the corresponding blurring values. It is also obvious how to model a blurring function with weights decaying with the distance from the origin.

(a)

(b)

Figure 2. (a) Degraded (a) and Restored (b) Images.

4. FINDING THE MAP ESTIMATE

If the blur operator $d(x)$ is exactly known, the Bayesian approach to image restoration seeks to minimize the following function

$$\frac{\alpha}{2}||Cf||^2 + \frac{\beta}{2}||g - Df||^2$$

subject to

$$0 \leq \min \leq \hat{f}(x) \leq \max < \infty. \tag{9}$$

with $\beta = 1/\sigma_n^2$.

Applying the Bayesian paradigm to the blind deconvolution problem, we try to maximize equation (3) with respect to f and d. In our problem, this is equivalent to minimizing

$$L(d,f) = \sum_{i=0}^{k}(\mu_i - 1)\ln(d(i)) + \frac{\alpha}{2}||Cf||^2 + \frac{\beta}{2}||g - Df||^2, \tag{10}$$

where $d(0) = 1 - d(1) - \cdots - d(k))$ subject to the conditions in equations (8) and (9).

Many optimization procedures can be employed to minimize equation (10). However, as in [7] there is a scale problem in $L(d,f)$ which, if not treated carefully, may undermine our minimization effort (see [7] for details).

4.1. Alternating Minimization

In this subsection, following [7], we seek to avoid the scale problem by following the idea of coordinate descent [1], and to minimize equation (10) by descending with respect to the PSF and the image separately and alternatively. The steps are the following

- Fix, $f(x)$, $x \in \Omega$, descend equation (10) with respect to $d(x)$, $x \in D$.
- Fix $d(x)$, $x \in D$, descend equation (10) with respect to $f(x)$, $x \in \Omega$.

Let

$$L(d, f|d) = L(d,f) \text{ with } d \text{ fixed} \tag{11}$$

and

$$L(d, f|f) = L(d,f) \text{ with } f \text{ fixed.} \tag{12}$$

Then the above idea can be developed into the following iterative algorithm

1. Initialization:

$$f^0(x) = g(x) \text{ and } d^0(x) = \text{random numbers}$$

2. nth iteration
   (a) Descend in d, $L(d, f^{n-1}|f^{n-1})$ to obtain d^n
   (b) Descend in f, $L(d^n, f|d^n)$ to obtain f^n
3. Stop when converged.

The convergence of the algorithm to one of the local minima of $L(d,f)$ can be established by noting the fact that both steps 2a and 2b always decrease $L(d,f)$.

Since the descents with respect to the PSF and the image are separate, they can use different methods to descend the corresponding function. In our examples we have used steepest descent for d in step 2a and exact deconvolution methods based on Fourier analysis for step 2b.

5. EXPERIMENTAL RESULTS

In this section we apply the method using the typical cameraman image. The original image was first blurred with the PSF shown in figure 4(a) and then noise of $30dB$ was added to obtain the image shown in figure 2(a). Then we minimized the function. The function

$$\gamma \ln(d(0)) + 4\gamma \ln(\overline{d}(1)) + \frac{\alpha}{2}||Cf||^2 + \frac{\beta}{2}||g - Df||^2$$

is then minimized, where $\overline{d}(1) = d(1) + d(3) + d(5) + d(7)$, $d(0) + \overline{d}(1) = 1$ and we also assumed that $d(2) = d(4) = d(6) = d(8) = 0$ in the iterative algorithm.

Step 2a in the iterative algorithm was implemented using a gradient descent method and step 2b using Fourier analysis. The parameters used were $\gamma/(\sum_i g(i)) = 0.002$, $\beta = 1/9$ and $\alpha = 1/300$.

The estimated PSF is shown in figure 4(b) and the restored image is shown in figure 2(b).

(a) (b)

Figure 3. (a) Degraded (a) and Restored (b) Images.

0	0.125	0
0.125	0.5	0.125
0	0.125	0

(a)

0	0.1265	0
0.1265	0.4940	0.1265
0	0.1265	0

(b)

Figure 4. (a) Original (a) and Estimated (b) Blurs.

0.05	0.125	0.05
0.125	0.3	0.125
0.05	0.125	0.05

(a)

0.0816	0.0939	0.0816
0.0939	0.2980	0.0939
0.0816	0.0939	0.0816

(b)

Figure 5. (a) Original (a) and Estimated (b) Blurs.

We also run the method for the same image blurred with the PSF shown in figure 5(a). Then noise of $30dB$ was added to obtain the image shown in figure 3(a). The function

$$\gamma \ln(d(0)) + 4\gamma \ln(\overline{d}(1)) 4\gamma \ln(\overline{d}(2)) + \frac{\alpha}{2}||Cf||^2 + \frac{\beta}{2}||g - Df||^2$$

was minimized, where $\overline{d}(1) = d(1)+d(3)+d(5)+d(7)$, $\overline{d}(2) = d(2) + d(4) + d(6) + d(8) = 0$ with $d(0) + \overline{d}(1) + \overline{d}(2) = 1$. The parameters used were $\gamma/(\sum_i g(i)) = 0.0025$, $\beta = 1/9$ and $\alpha = 1/600$. The estimated PSF is shown in figure 5(b) and the restored image is shown in figure 3(b).

6. CONCLUSIONS

In this paper we have proposed the use of Dirichlet distributions to model our prior knowledge about the blurring function together with smoothness constraints on the restored image to solve the blind deconvolution problem.

We have shown that the use of Dirichlet distributions offers a lot of flexibility in incorporating vague or very precise knowledge about the blurring process into the blind deconvolution process and also that the proposed MAP estimator offers additional flexibility in modeling the original image.

We are currently working on improving the iterative scheme for the PSF estimation and also on the estimation of the α, β and γ following the evidence and MAP approaches to hyperparameter estimation.

REFERENCES

[1] A.K. Katsaggelos, ed., *Digital Image Restoration*, New York, Springer–Verlag, 1991.

[2] D. Kundur and D. Hatzinakos, "Blind Image Deconvolution", *IEEE Signal Processing Magazine*, vol. 13, No. 3, pp. 43-64, 1996.

[3] D. Kundur and D. Hatzinakos, "Blind Image Deconvolution Revisited", *IEEE Signal Processing Magazine*, vol. 13, No. 6, pp. 61-63, 1996.

[4] R. Molina and B.D. Ripley, "Using Spatial Models as Priors in Astronomical Image Analysis", *J. Appl. Statist.*, vol. 16, pp. 193-206, 1989.

[5] B.D. Ripley, *Spatial Statistics*, John Wiley, New York, pp. 88-90, 1981.

[6] S.S. Wilks, *Mathematical Statistics*, John Wiley and Sons, 1962.

[7] 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).