BAYESIAN IMAGE RESTORATION USING A WAVELET-BASED SUBBAND DECOMPOSITION

Rafael Molina[a], Aggelos K. Katsaggelos[b] and Javier Abad[a]

a)Departamento de Ciencias de la Computación e Inteligencia Artificial,
Universidad de Granada. 18071 Granada, España.
b)Department of Electrical and Computer Engineering,
Northwestern University, Evanston, Illinois 60208-3118
e-mail: rms@decsai.ugr.es, aggk@ece.nwu.edu, abad@decsai.ugr.es

ABSTRACT

In this paper the subband decomposition of a single channel image restoration problem is examined. The decomposition will be carried out in the image model (prior model) in order to take into account the frequency activity of each band of the original image. The hyperparameters associated to each band together with the original image are rigurously estimated within the Bayesian framework. Finally the proposed method is tested and compared with other methods on real and synthetic images.

1. INTRODUCTION

A standard formulation of the image degradation model is given in lexicographic form by [1]

$$\mathbf{g} = \mathbf{D}\mathbf{f} + \mathbf{w}, \tag{1}$$

where the $p \times 1$ vectors $\mathbf{f}$, $\mathbf{g}$, and $\mathbf{w}$ represent respectively the original image, the available noisy and blurred image and the noise with independent elements of variance $\sigma^2_{\mathbf{w}} = \beta^{-1}$, and $\mathbf{D}$ represents the known blurring matrix. The images are assumed to be of size $n \times n$, with $p = n \times n$. The *restoration problem* calls for finding an estimate of $\mathbf{f}$ given $\mathbf{g}$, $\mathbf{D}$ and knowledge about $\mathbf{w}$ and possibly $\mathbf{f}$, (see Chapter 1 in [13]).

Smoothness constraints on the original image can be incorporated under the form of

$$p(\mathbf{f}|\alpha) \propto \alpha^{p/2} \exp\{-\frac{1}{2}\alpha \parallel \mathbf{C}\mathbf{f} \parallel^2\}, \tag{2}$$

where $\mathbf{C}$ is the Laplacian operator.

Then, following the Bayesian paradigm it is customary to select as the restoration of $\mathbf{f}$, the image $\mathbf{f}_{(\alpha,\beta)}$ defined by

$$\mathbf{f}_{(\alpha,\beta)} = \arg\{\min_{\mathbf{f}} [\,\alpha \parallel \mathbf{C}\mathbf{f} \parallel^2 + \beta \parallel \mathbf{g} - \mathbf{D}\mathbf{f} \parallel^2\,]\}$$

$$= \arg\{\max_{\mathbf{f}} p(\mathbf{f}|\alpha) p(\mathbf{g}|\mathbf{f},\beta)\}, \tag{3}$$

where from Eq. 1 we have

$$p(\mathbf{g}|\mathbf{f},\beta) \propto \beta^{p/2} \exp\{-\frac{1}{2}\beta \parallel \mathbf{g} - \mathbf{D}\mathbf{f} \parallel^2\}. \tag{4}$$

An important problem arises when α and/or β are unknown. Much interest has centered on the question of how these parameters should be estimated (see [8], [15]). However, it is widely accepted that the hyperparameter in the image model (α) should be adapted to the local image characteristics, although the computations involved have prevented the development of much research in that area (see however [11], [12]).

Since the pionner work of Galatsanos and Chin ([7]) an important amount of research has been carried out in the area of multichannel image restoration (see for example [9], [18], [10], [4],[16]), unfortunately no much work has been reported on the estimation of the parameters needed in multichannel image restoration (see however, [9], [21] , [22], [11], [12]). Although the computations involved in a multichannel problems are more time demanding, the paper [14] provided a general framework where all the calculations involved in a multichannel image restoration could be easily carried out in the frequency domain.

The application of multichannel techniques to single channel restoration problems images using a subband decomposition was proposed in [2] and [3] using the framework developed in [14] .

In this paper we examine the subband decomposition of the quadratic image model given in Eq. 2. Since by performing a subband decomposition we are extracting different frequency regions (channels) of an image, the process of associating a different image hyperparameter to each subband of the image model becomes equivalent to assigning different hyperparameters to different frequency bands in the image. Those hyperparameters will reflect then the activity of that band in the original image. We show how the estimation

This work has been supported by the "Comisión Nacional de Ciencia y Tecnología" under contract TIC-989.

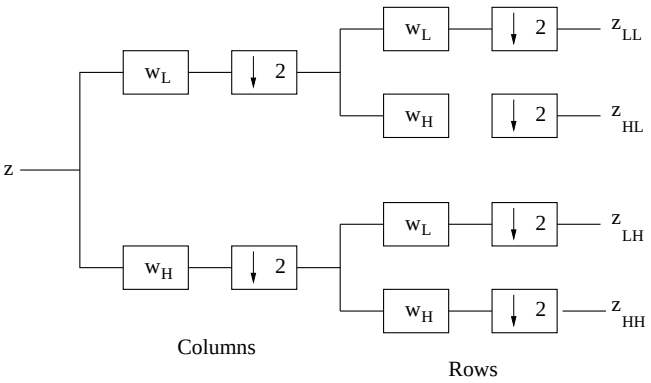


Figure 1: Four-channel 2-D decomposition

of those parameters can be carried out within the Bayesian image restoration paradigm and the computation efficiently implemented.

The rest of the paper is organized as follows. In section 2 the image and noise models are defined in order to apply the Bayesian paradigm. For those image and noise models, the estimation of the hyperparameters and the original image is performed in section 3. Finally in section 4 experimental results are shown and section 5 concludes the paper. The paper also contains an appendix where it is shown how to calculate all the involved computations.

2. IMAGE AND NOISE MODELS

We will only discuss the image model and use the noise model defined by Eq. 4.

A simple way to incorporate the smoothness of the object luminosity is to model the distribution of $\mathbf{f}$ by Eq. 2. It is important to note that this model is a simultaneous autorregression (SAR) ([17]) and that is characterized by

$$(\mathbf{Cf})_i = \epsilon_i, \tag{5}$$

where ϵ_i are independent $\mathcal{N}(0, \alpha^{-1})$.

A carefull examination of Eq. 5 shows that this expression is not true for real images. The spectrum of $\mathbf{Cf}$ is not normally flat and the energy in each frequency is not the same $(1/\alpha)$. Obviously the image model is just a simple approximation.

Let us now consider $\mathbf{z} = \mathbf{Cf}$ and perform a multichannel descomposition of that image. Let $\mathbf{w}_l$ and $\mathbf{w}_h$ be $1-D$ quadrature mirror filters (QMF) based on the orthonormal wavelet bases with compact support [5], so that one set of coefficients may be used to defined the other [20]. Then, the subband decomposition of $\mathbf{z}$ can be calculated as described in Fig. 1.

We note that

$$\mathbf{I} = \mathbf{W}_{ll}^t\mathbf{W}_{ll} + \mathbf{W}_{hl}^t\mathbf{W}_{hl} + \mathbf{W}_{lh}^t\mathbf{W}_{lh} + \mathbf{W}_{hh}^t\mathbf{W}_{hh}, \tag{6}$$

where $\mathbf{W}_{uv}$ with $u, v \in l, h$ are the $[(n/2) \times (n/2)] \times [n \times n]$ matrices used to obtain the bands $\mathbf{z}_{uv}$ (see Fig. 1) and ^{t} denotes transpose. It is important to observe that now $\mathbf{W}_{uv}^t\mathbf{W}_{uv}\mathbf{z}$ contains information on some part of the spectrum of $\mathbf{z}$.

Let us consider the quadratic form defining the image model, we have

$$\begin{aligned} \alpha \parallel \mathbf{Cf} \parallel^2 &= \alpha\mathbf{f}\mathbf{C}^t\mathbf{Cf} \\ &= \mathbf{f}\mathbf{C}^t(\alpha\mathbf{W}_{ll}^t\mathbf{W}_{ll} + \alpha\mathbf{W}_{hl}^t\mathbf{W}_{hl} + \\ &\quad \alpha\mathbf{W}_{lh}^t\mathbf{W}_{lh} + \alpha\mathbf{W}_{hh}^t\mathbf{W}_{hh})\mathbf{Cf} \end{aligned} \tag{7}$$

Now, in order to adapt the image model, and so have a hyperparameter for each of the descomposed channels, we define the following image model

$$p(\mathbf{f}|\underline{\alpha}) \propto \frac{1}{Z_{prior}(\underline{\alpha})} \exp\{-\frac{1}{2} \sum_{u,v\in\{l,h\}} \alpha_{uv} \parallel \mathbf{W}_{uv}\mathbf{Cf} \parallel^2\} \tag{8}$$

where $\underline{\alpha}$ denotes the vector $(\alpha_{ll}, \alpha_{hl}, \alpha_{lh}, \alpha_{hh})$ and

$$Z_{prior}(\underline{\alpha}) = |\mathbf{P}(\underline{\alpha})|^{-1/2} \tag{9}$$

where

$$\mathbf{P}(\underline{\alpha}) = \sum_{u,v\in\{l,h\}} \alpha_{u,v}\mathbf{C}^t\mathbf{W}_{uv}^t\mathbf{W}_{uv}\mathbf{C} \tag{10}$$

Before estimating the hyperparameters and restore the original image using the Bayesian paradigm, let us examine how to extend the model to a multiscale decomposition.

2.1. Multiscale decomposition of the image model

The model we have just proposed can be extended to a 4^i channel decomposition. To work with 4^i channels and so asociate 4^i hyperparameters to the image model we use as prior model the one defined by the quadratic form:

$$\begin{aligned} \parallel \mathbf{A}(i)\mathbf{f} \parallel^2 &= \sum_{\substack{u_k,v_k\in\{l,h\},\\ k=1,\ldots,i}} \alpha_{u_1v_1\ldots u_iv_i} \times \\ &\quad \parallel \mathbf{W}_{u_iv_i}(i)\ldots\mathbf{W}_{u_1v_1}(1)\mathbf{Cf} \parallel^2, \end{aligned}$$

where $\mathbf{W}_{u_kv_k}(k)$ with $u_k, v_k \in l, h$ and $k = 1, \ldots, i$ are the $\left[(n/2^i) \times (n/2^i)\right] \times \left[n/2^{i-1} \times n/2^{i-1}\right]$ matrices used to subband decompose $[n/2^{i-1}] \times [n/2^{i-1}]$ images.

In this paper, for notation simplicity, we will only use a 4 channels decomposition, section 4 however contains test examples carried out at higher decomposition.

2.2. Using the same hyperparameters for several subbands

The image model we are proposing allows the use of the same hyperparameters for several subbands. If 4 channels

are used, the same hyperparameter could be assigned to all the detail subbands and that one be different from the one used for the ll band. The same ideas can be applied when 4^i channels are used.

Finally, it is worth mentioning that the method proposed in [19] is a particular case of the one we are proposing here and that it is obtained by using as image prior the one defined by the quadratic form $\| f \|^2$ and decomposing it in subbands.

Let us now examine how tho estimate the unknown parameters and the restoration in the coming section.

3. BAYESIAN ANALYSIS

The steps we follow in this paper to estimate the hyperparameters and the original image are

Step I: Estimation of the hyperparameters

$\underline{\hat{\alpha}} = (\hat{\alpha}_{ll}, \hat{\alpha}_{hl}, \hat{\alpha}_{lh}, \hat{\alpha}_{ll})$ and $\hat{\beta}$ are first selected as

$$\underline{\hat{\alpha}}, \hat{\beta} = \arg\max_{\underline{\alpha},\beta} \mathcal{L}_{\mathbf{g}}(\underline{\alpha}, \beta) = \arg\max_{\underline{\alpha},\beta} \log p(\mathbf{g}|\underline{\alpha}, \beta), \quad (11)$$

where

$$p(\mathbf{g}|\underline{\alpha}, \beta) = \int_{\mathbf{f}} p(\mathbf{f}|\underline{\alpha}) p(\mathbf{g}|\mathbf{f}, \beta) d\mathbf{f}$$

Step II: Estimation of the original image

Once the hyperparamenters have been estimated, the estimation of the original image, $\mathbf{f}_{(\hat{\underline{\alpha}},\hat{\beta})}$, is selected as the image satisfying

$$\mathbf{f}_{(\hat{\underline{\alpha}},\hat{\beta})} = \arg\min_{\mathbf{f}} \sum_{uv\in\{lh\}} \hat{\alpha}_{uv} \| \mathbf{W}_{uv}\mathbf{C}\mathbf{f} \|^2 + \hat{\beta} \| \mathbf{g} - \mathbf{D}\mathbf{f} \|^2 \quad (12)$$

Note that we are estimating the hyperparameter by maximum likelihood and that the estimation of $\mathbf{f}$ is performed by using the *maximum a posteriori* MAP. Furthermore, although steps I and II are separated, the iterative scheme to be proposed performs both estimations simultaneously.

The estimation process we are using could be performed within the so called hierarchical Bayesian approach (see [15]) by including hyperpriors on the unkonwn hypervector $\underline{\hat{\alpha}}$ and hyperparameters $\hat{\beta}$. However, the possibility of incorporating additional knowledge on them by means of gamma or other distributions will not be discussed here (see [15]).

Let us examine the estimation process in detail. Fixing $\underline{\alpha}$ and β and expanding the function

$$M(\mathbf{f}, \mathbf{g}|\underline{\alpha}, \beta) = \sum_{u,v\in\{l,h\}} \alpha_{uv} \| \mathbf{W}_{uv}\mathbf{C}\mathbf{f} \|^2 + \beta \| \mathbf{g} - \mathbf{D}\mathbf{f} \|^2 \quad (13)$$

around $\mathbf{f}_{(\underline{\alpha},\beta)}$, we have

$$\begin{aligned} M(\mathbf{f}, \mathbf{g}|\underline{\alpha}, \beta) &= M(\mathbf{f}_{(\underline{\alpha},\beta)}, \mathbf{g}|\underline{\alpha}, \beta) \\ &+ \frac{1}{2}(\mathbf{f} - \mathbf{f}_{(\underline{\alpha},\beta)})^t \mathbf{Q}(\underline{\alpha}, \beta)(\mathbf{f} - \mathbf{f}_{(\underline{\alpha},\beta)}) \end{aligned}$$

where

$$\mathbf{Q}(\underline{\alpha}, \beta) = \sum_{u,v\in\{l,h\}} \alpha_{u,v} \mathbf{C}^t \mathbf{W}_{u,v}^t \mathbf{W}_{u,v} \mathbf{C} + \mathbf{D}^t\mathbf{D}$$

and therefore

$$\begin{aligned} p(\mathbf{g}|\underline{\alpha}, \beta) &= \frac{\exp\{-M(\mathbf{f}_{(\underline{\alpha},\beta)}, \mathbf{g}|\underline{\alpha}, \beta)/2\}}{Z_{prior}(\underline{\alpha}) Z_{noise}(\beta)} \times \\ &\int_{\mathbf{f}} \exp\{-\frac{1}{2}(\mathbf{f} - \mathbf{f}_{(\underline{\alpha},\beta)})^t \mathbf{Q}(\underline{\alpha}, \beta)(\mathbf{f} - \mathbf{f}_{(\underline{\alpha},\beta)})\} d\mathbf{f} \\ &= \frac{\exp\{-M(\mathbf{f}_{(\underline{\alpha},\beta)}, \mathbf{g}|\underline{\alpha}, \beta)/2\} |\mathbf{Q}(\underline{\alpha}, \beta)|^{-1/2}}{Z_{prior}(\underline{\alpha}) Z_{noise}(\beta)} \quad (14) \end{aligned}$$

We then have

$$\begin{aligned} 2\mathcal{L}_{\mathbf{g}}(\underline{\alpha}, \beta) = &- \sum_{u,v\in\{l,h\}} \alpha_{uv} \| \mathbf{W}_{uv}\mathbf{C}\mathbf{f}_{(\underline{\alpha},\beta)} \|^2 \\ &- \beta \| \mathbf{g} - \mathbf{D}\mathbf{f}_{(\underline{\alpha},\beta)} \|^2 - \log |Q(\underline{\alpha}, \beta)| \\ &- 2\log Z_{prior}(\underline{\alpha}) - 2\log Z_{noise}(\beta) + const. \end{aligned}$$

We now differentiate $-2\mathcal{L}_{\mathbf{g}}(\underline{\alpha}, \beta)$ with respect to $\underline{\alpha}$ and β so as to find the conditions which are satisfied at the maxima. We have

$$\begin{aligned} &\| \mathbf{W}_{uv}\mathbf{C}\mathbf{f}_{(\underline{\alpha},\beta)} \|^2 + trace[\mathbf{Q}(\underline{\alpha}, \beta)^{-1} \mathbf{C}^t \mathbf{W}_{uv}^t \mathbf{W}_{uv} \mathbf{C}] = \\ &trace[\mathbf{P}(\underline{\alpha})^{-1} \mathbf{C}^t \mathbf{W}_{uv}^t \mathbf{W}_{uv} \mathbf{C}] \quad \text{for } u, v \in \{l, h\} \quad (15) \\ &\| \mathbf{g} - \mathbf{D}\mathbf{f}_{(\underline{\alpha},\beta)} \|^2 + trace[\mathbf{Q}(\underline{\alpha}, \beta)^{-1} \mathbf{D}^t\mathbf{D}] = p/\beta. \quad (16) \end{aligned}$$

Note that if the same hyperparameter is used for some subbands, the corresponding Eqs. 15 and 16 are also easy to calculate.

By multiplying Eqs. 15 by α_{uv} and 16 by β we have

$$\begin{aligned} &\alpha_{uv} \| \mathbf{W}_{uv}\mathbf{C}\mathbf{f}_{(\underline{\alpha},\beta)} \|^2 + \\ &\alpha_{uv} trace[\mathbf{Q}(\underline{\alpha}, \beta)^{-1} \mathbf{C}^t \mathbf{W}_{uv}^t \mathbf{W}_{uv} \mathbf{C}] = \\ &\alpha_{uv} trace[\mathbf{P}(\underline{\alpha})^{-1} \mathbf{C}^t \mathbf{W}_{uv}^t \mathbf{W}_{uv} \mathbf{C}] \text{ for } u, v \in \{l, h\} \quad (17) \\ &\beta \| \mathbf{g} - \mathbf{D}\mathbf{f}_{(\underline{\alpha},\beta)} \|^2 + \beta trace[\mathbf{Q}(\underline{\alpha}, \beta)^{-1} \mathbf{D}^t\mathbf{D}] = p \quad (18) \end{aligned}$$

and since

$$\begin{aligned} &\sum_{u,v\in\{l,h\}} \alpha_{uv} trace[\mathbf{Q}(\underline{\alpha}, \beta)^{-1} \mathbf{C}^t \mathbf{W}_{uv}^t \mathbf{W}_{uv} \mathbf{C}] \\ &+ \beta trace[\mathbf{Q}(\underline{\alpha}, \beta)^{-1} \mathbf{D}^t\mathbf{D}] = trace[\mathbf{I}_p] = p, \end{aligned}$$

and

$$\begin{aligned} &\sum_{u,v\in\{l,h\}} \alpha_{uv} trace[\mathbf{P}(\underline{\alpha})^{-1} \mathbf{C}^t \mathbf{W}_{uv}^t \mathbf{W}_{uv} \mathbf{C}] = \\ &trace[\mathbf{I}_p] = p, \end{aligned}$$

we have at the maximum likelihood estimate (mle)

$$\sum_{u,v\in\{l,h\}} \alpha_{uv} \parallel \mathbf{W}_{uv}\mathbf{C}\mathbf{f}_{(\alpha,\beta)} \parallel^2 + \beta \parallel \mathbf{g} - \mathbf{D}\mathbf{f}_{(\underline{\alpha},\beta)} \parallel^2 = p,$$

which means that at the mle a fraction of the observations are used to estimate the missfit to the components of the prior model ($\alpha_{uv} \parallel \mathbf{W}_{uv}\mathbf{C}\mathbf{f}_{(\underline{\alpha},\beta)} \parallel^2$, with $u, v \in \{l, h\}$) and another proportion is used for the missfit to the noise model ($\beta \parallel \mathbf{g} - \mathbf{D}\mathbf{f}_{(\underline{\alpha},\beta)} \parallel^2$).

Let us examine the use of the EM-algorithm [6] with $\mathcal{X}^t = (\mathbf{f}^t, \mathbf{g}^t)$ and $\mathcal{Y} = \mathbf{g} = [\mathbf{0}\ \mathbf{I}]^\mathbf{t}\mathcal{X}$ to iteratively increase $\mathcal{L}_\mathbf{g}(\underline{\alpha}, \beta)$. The application of the EM-algorithm to our problems produces Eqs. 15 and 16 where the old values of the hyperparameters are used on the left hand side of equations to obtain the new ones on the right hand side of the equations. Unfortunately, those equations are highly nonlinear and so the new values of the hyperparameters are difficult to find.

Let us, however, consider the iterative EM equations corresponding to using one hyperparameter for the image model (α) and one for the noise (β) (see [15]), we have

$$\begin{aligned}
\left[\frac{1}{\alpha}\right]_{new} &= \left[\frac{1}{p}\left\{\parallel \mathbf{C}\mathbf{f}_{(\alpha,\beta)} \parallel^2 + \right.\right. \\
&\quad \left.\left. trace[(\alpha\mathbf{C}^t\mathbf{C} + \beta\mathbf{D}^t\mathbf{D})^{-1}\mathbf{C}^t\mathbf{C}]\right\}\right]_{old} \\
&= \left[\frac{1}{\alpha} + \frac{1}{p}\left\{\parallel \mathbf{C}\mathbf{f}_{(\alpha,\beta)} \parallel^2 + \right.\right. \\
&\quad \left.\left. trace[(\alpha\mathbf{C}^t\mathbf{C} + \beta\mathbf{D}^t\mathbf{D})^{-1}\mathbf{C}^t\mathbf{C}] - \frac{p}{\alpha}\right\}\right]_{old} \quad (19) \\
\left[\frac{1}{\beta}\right]_{new} &= \left[\frac{1}{p}\left\{\parallel \mathbf{g} - \mathbf{D}\mathbf{f}_{(\alpha,\beta)} \parallel^2 + \right.\right. \\
&\quad \left.\left. trace[(\alpha\mathbf{C}^t\mathbf{C} + \beta\mathbf{D}^t\mathbf{D})^{-1}\mathbf{D}^t\mathbf{D}]\right\}\right]_{old} \\
&= \left[\frac{1}{\beta} + \frac{1}{p}\left\{\parallel \mathbf{g} - \mathbf{D}\mathbf{f}_{(\alpha,\beta)} \parallel^2 + \right.\right. \\
&\quad \left.\left. trace[(\alpha\mathbf{C}^t\mathbf{C} + \beta\mathbf{D}^t\mathbf{D})^{-1}\mathbf{D}^t\mathbf{D}] - \frac{p}{\beta}\right\}\right]_{old} \quad (20)
\end{aligned}$$

where $\mathbf{f}_{(\alpha,\beta)}$ has been defined in Eq. 3 and $[\,]_{new}$ and $[\,]_{old}$ mean the evaluation of the expressions for the new and old values of α and β respectively.

We notice that these equations correspond to the application of a gradient descendent method on $1/\alpha$ and $1/\beta$ and that the step on the gradient used is $1/(p\alpha^2)$ and $1/(p\beta^2)$ for $1/\alpha$ and $1/\beta$ respectively. Notice that for a given function y depending on x, $y(x)$, we have that

$$\frac{dy}{d(1/x)} = \frac{dy}{dx}\frac{dx}{d(1/x)} = -x^2\frac{dy}{dx}$$

Let us adapt this method to the multichannel problem. Multiplying and dividing the right hand side of Eq. 15 by α_{uv} we have

$$\begin{aligned}
&\frac{1}{p(\alpha_{uv})}\left[\parallel \mathbf{W}_{uv}\mathbf{C}\mathbf{f}_{(\underline{\alpha},\beta)} \parallel^2 + \right. \\
&\quad \left. trace[\mathbf{Q}(\underline{\alpha},\beta)^{-1}\mathbf{C}^t W_{uv}^t\mathbf{W}_{uv}\mathbf{C}]\right] = 1/\alpha_{uv}, \quad (21)
\end{aligned}$$

where

$$p(\alpha_{uv}) = \alpha_{uv} trace[\mathbf{P}(\underline{\alpha})^{-1}\mathbf{C}^t\mathbf{W}_{uv}^t\mathbf{W}_{uv}\mathbf{C}],$$

(we have removed the dependency on $\underline{\alpha}$ of $p(\alpha_{uv})$ to simplify the notation).

Notice that $p(\alpha_{uv}) = p$ if we have only one image parameter and that

$$\sum_{u,v\in\{l,h\}} p(\alpha_{uv}) = p$$

Then we can use the following equations to estimate the hyperparameters, where the old values are used in the right hand side of the equations to obtain the new ones on the left hand side

$$\begin{aligned}
\left[\frac{1}{\alpha_{uv}}\right]_{new} &= \frac{1}{p(\alpha_{uv})}\left[\parallel \mathbf{W}_{uv}\mathbf{C}\mathbf{f}_{(\underline{\alpha},\beta)} \parallel^2 + \right. \\
&\quad \left. trace[\mathbf{Q}(\underline{\alpha},\beta)^{-1}\mathbf{C}^t\mathbf{W}_{uv}^t\mathbf{W}_{uv}\mathbf{C}]\right]_{old} \\
&= \left[\frac{1}{\alpha_{uv}} + \frac{1}{p(\alpha_{uv})}\left\{\parallel \mathbf{W}_{uv}\mathbf{C}\mathbf{f}_{(\alpha,\beta)} \parallel^2 + \right.\right. \\
&\quad trace[\mathbf{Q}(\underline{\alpha},\beta)^{-1}\mathbf{C}^t\mathbf{W}_{uv}^t\mathbf{W}_{uv}\mathbf{C}] \\
&\quad \left.\left. -trace[\mathbf{P}(\underline{\alpha})^{-1}\mathbf{C}^t\mathbf{W}_{uv}^t\mathbf{W}_{uv}\mathbf{C}]\right\}\right]_{old} \\
&\qquad \text{for } u, v \in \{l, h\} \quad (22) \\
\left[\frac{1}{\beta}\right]_{new} &= \left[\frac{1}{\beta} + \frac{1}{p}\left\{\parallel \mathbf{g} - \mathbf{D}\mathbf{f}_{(\underline{\alpha},\beta)} \parallel^2 + \right.\right. \\
&\quad \left.\left. trace[\mathbf{Q}(\underline{\alpha},\beta)^{-1}\mathbf{D}^t\mathbf{D}] - \frac{p}{\beta}\right\}\right]_{old} \quad (23)
\end{aligned}$$

This method is again a gradient descent one. We have used it in our experiments and have not observed any convergence problem, however, it would always be possible to use smaller steps as to guarante convergence.

4. EXPERIMENTAL RESULTS

In order to show the behaviour of the proposed algorithm, we have used the original 256×256 "Cameraman" image, blurred by a motion blur over 9 pixels. It was degraded by additive Gaussian noise to achieve 10, 20 and 30dB SNR (noise variances of $\beta^{-1} = 216.1$, $\beta^{-1} = 64$, and $\beta^{-1} = 6.25$, respectively). In all cases, the initial image was the degraded one. For a comparison, we have also applied the maximum likelihood restoration method to these degraded images.

Method	Iterations	$ISNR$	β^{-1}
MLE	40	7.6038	184.98
1 param.	30	7.2863	214.03
2 params.	50	7.2844	214.09
4 params.	60	7.2241	213.95

10 dB

Method	Iterations	$ISNR$	β^{-1}
MLE	35	7.9889	49.84
1 param.	35	8.8162	63.64
2 params.	50	8.8086	63.68
4 params.	70	8.2787	64.73

20 dB

Method	Iterations	$ISNR$	β^{-1}
MLE	30	6.2049	3.91
1 param.	35	8.8181	6.37
2 params.	41	8.8006	6.42
4 params.	90	9.1402	6.57

30 dB

Table 1: iterations required, ISNR, and noise variance estimations for the "Cameraman" image and different SNRs.

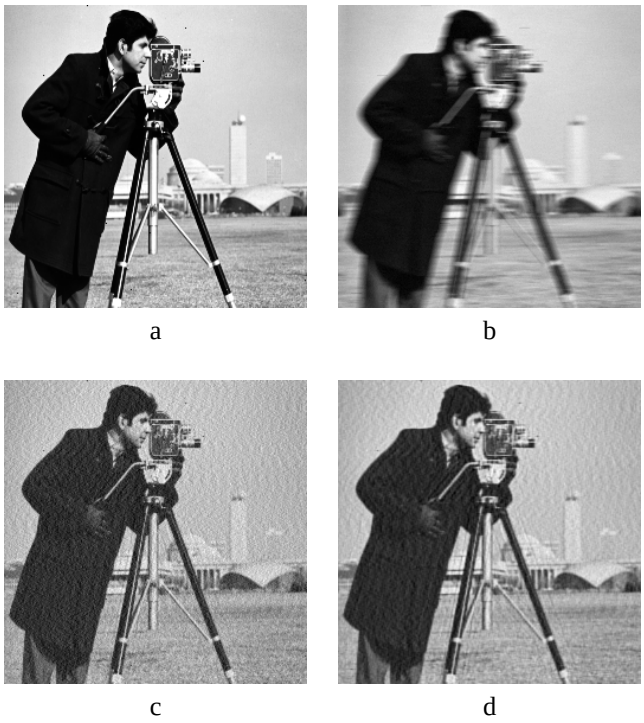
For the purpose of objetively testing the performance of the image restoration algorithms, the Improvement in Signal to Noise Ratio (ISNR) will be used. This metric is given by

$$ISNR = 10 \log_{10} \left\{ \frac{\sum_{m,n} [\mathbf{f}(m,n) - \mathbf{g}(m,n)]^2}{\sum_{m,n} \left[\mathbf{f}(m,n) - \mathbf{f}_{(\underline{\hat{\alpha}},\hat{\beta})}(m,n)\right]^2} \right\}$$

In any case, it is very important to consider the behaviour of the algorithms from the viewpoint of edge preservation, which can be a basic indicator of improvement in quality for subjetive comparison of restoration algorithms

The values of ISNR, the required number of iterations needed to achieve convergence in parameter estimation, and the corresponding values of the estimated noise variance are shown in table 1. We have included the results obtained by maximum likelihood and the proposed algorithm using only one parameter for all the bands of the image; using two parameters, one for the ll band, and a different one for the lh, hl, and hh. The last row of every sub-table contains the result obtained using a different parameter for each subband.

From this table we can see that the proposed method obtains better estimations of the noise variance, very close to the real value, giving less noisy images than maximum likelihood method. We can see that the ISNR is also better for the proposed method.



a b c d

Figure 2: (a) Original image. (b) Noisy-blurred image for 9-point motion blur at 30dB. (c) Maximum likelihood restoration. (d) Restoration obtained with the proposed method.

Fig. 2 shows the original "Cameraman" image, the degraded image at 30 dB, the restoration obtained by maximum likelihood and the restoration obtained with the proposed method using four prior-model parameters, a different one for each band. We can see that the solution proposed gives smoother solutions but the noise is much better removed.

5. CONCLUSIONS

6. REFERENCES

[1] H.C. Andrews and B. R. Hunt, *Digital Image Restoration*, Prentice Hall, New York, 1977.

[2] M.R. Banham, N. P. Galatsanos, H.L. Gonzalez and A.K. Katsaggelos (1994) "Multichannel Restoration of Single Channel Images Using a Wavelet-Based Subband Decomposition", *IEEE Trans on Image Processing* vol 3, pp 1-13.

[3] M.R. Banham and A.K. Katsaggelos (1994) "Spatially Adaptive wavelet-based multiscale image restoration" *IEEE Trans on Image Processing* vol 5, pp 619-634.

[4] M.M. Chang, M.I. Sezan, A.M. Tekalp and M.J. Berg (1996) "Bayesian segmentation of multislice brain

magnetic resonance imaging using three-dimensional Gibbsian priors", *Optical Engineering*, vol. 35, No. 11, pp 3206–3221.

[5] I. Daubechies, (1988) "Orthonormal bases of compactly supported wavelets, *Commun. Pure Appl. Math.*, vol. 41, pp-909-996.

[6] A.P. Dempster, N.M. Laird and D.B. Rubin, "Maximum Likelihood from Incomplete Data", *J.Royal Statist. Soc. B*, vol. 39, pp. 1-38, 1972.

[7] Galatsanos, N.P. and Chin, R. T., (1989) "Digital Restoration of Multichannel Images", *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, No. 3, pp. 415–421.

[8] N.P. Galatsanos and A.K. Katsaggelos, "Methods for Choosing the Regularization Parameter and Estimating the Noise Variance in Image Restoration and their Relation", *IEEE Trans. on Image Processing*, vol. 1, pp. 322-336, 1992.

[9] Galatsanos, N.P., Katsaggelos, A.K., Chin and R. T., Hillery, A.D., (1991) "Least Squares Restoration of Multichannel Images", *IEEE Trans. on Signal Processing*, vol. 39, No. 10, pp. 2222–2236.

[10] Guo, Y.P., Lee, H.P. and Teo, C.L., (1996) "Multichannel Image Restorations Using Iterative Algorithm in Space Domain", *Image and Video Computing*, vol. 14, pp. 389–400.

[11] M.G. Kang and A.K. Katsaggelos (1995) "Frequency-domain adaptive iterative image restoration and evaluation of the regularization parameter", *Optical Engineering*, vol 33, pp 3222-3232

[12] M.G. Kang and A.K. Katsaggelos (1997) "Simultaneous Multichannel Image Restoration and estimation of the regularization parameters", *IEEE Trans on Image Processing* vol 6, pp 774-778.

[13] A.K. Katsaggelos, editor, *Digital Image Restoration*, Springer Series in Information Sciences, vol. 23, Springer-Verlag, 1991.

[14] A.K. Katsaggelos, K.T. Lay and N.P. Galatsanos (1993) "A general framework for frequency domain multichannel signal processing" *IEEE Trans on Image Processing* vol 2, pp 417-420.

[15] R. Molina, A.K. Katsaggelos and J. Mateos (1998) "Bayesian and Regularization Methods for Hyperparameter Estimation in Image Restoration", *IEEE Trans on Image Processing*, accepted for publication.

[16] R. Molina and J. Mateos (1997), "Multichannel Image Restoration in Astronomy", to appear in *Vistas in Astronomy*, Special issue edited by R. Molina, F. Murtagh and A. Heck , *From Information Fusion to Data Mining*.

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

[18] R.R. Schultz and R.L. Stevenson, (1995) "Stochastic Modeling and Estimation of Multispectral Image Data", *IEEE Trans. on Image Processing*, vol. **IP**4, No. 8, pp. 1109–1119.

[19] I.M. Stephanakis, (1997) "Regularized image restoration in multiresolution spaces", *Optical Engineering*, vol 36, 1738-1744.

[20] P.P. Vaidyanathan (1993) *Multirate Systems and Filter Banks*, Englewoods Cliffs, NJ: Prentice Hall.

[21] M.E. Zervakis, (1992), "Optimal Restoration of Multichannel Images Based on Constrained Mean-Square Estimation, *Journal of Visual Communication and Image Representation*, vol. 3, pp 392-411.

[22] **W**. Zhu and N.P. Galatsanos, (1995) "Regularized Multichannel Restoration Using Cross-Validation" *Graphical Models and Image Processing*, vol. 57, pp 38-54.