A Bayesian approach for the estimation and transmission of regularization parameters for reducing blocking artifacts

Javier Mateos, Aggelos K. Katsaggelos and Rafael Molina

Abstract

With block-based compression approaches for both still images and sequences of images annoying blocking artifacts are exhibited, primarily at high compression ratios. They are due to the independent processing (quantization) of the block transformed values of the intensity or the displaced frame difference. In this paper, we propose the application of the hierarchical Bayesian paradigm to the reconstruction of block discrete cosine transform (BDCT) compressed images and the estimation of the required parameters. We derive expressions for the iterative evaluation of these parameters applying the evidence analysis within the hierarchical Bayesian paradigm. The proposed method allows for the combination of parameters estimated at the coder and decoder. The performance of the proposed algorithms is demonstrated experimentally.

Keywords

Reconstruction, Bayesian models, evidence analysis, post-processing, image coding.

I. Introduction

Throughout this paper a digital $M \times N$ image is treated as a $(M \times N) \times 1$ vector in the $R^{M \times N}$ space by lexicographically ordering either its rows or columns. The block discrete cosine transform (BDCT) is a linear transformation from $R^{M \times N}$ to $R^{M \times N}$. That is, for an image $\mathbf{f}$ we can write

$$\mathbf{F} = \mathbf{B}\mathbf{f},$$

where $\mathbf{F}$ is the BDCT of $\mathbf{f}$ and $\mathbf{B}$ is the BDCT matrix. To achieve a bit-rate reduction, each element of $\mathbf{F}$ is quantized. This quantization operator represents mathematically a mapping, $\mathcal{Q}$ from $R^{M \times N}$ to

Preliminary versions of this work were presented at the 1996 Conference on Digital Compression Technologies and Video Communications [1] and the 1997 International Conference on Digital Signal Processing [2]. It was partially supported by the "Comisión Nacional de Ciencia y Tecnología" under contract TIC97-0989.

J. Mateos and R. Molina are with the Departamento de Ciencias de la Computación e I.A. Universidad de Granada. 18071 Granada, España.

A. K. Katsaggelos is with the Department of Electrical and Computer Engineering, Northwestern University, Evanston, IL 60208

$R^{M \times N}$. The input-output relation of the coder can be modeled by

$$\mathbf{G} = \mathcal{Q}\mathbf{B}\mathbf{f}.$$

Due to the unitary property of the DCT matrices, the BDCT matrix is also unitary and the inverse transform can be simply expressed by $\mathbf{B}^t$, where t denotes the transpose of a matrix. In the receiver only the quantized BDCT coefficients $\mathbf{G}$ are available and the output of a conventional decoder is $\mathbf{g} = \mathbf{B}^t\mathbf{G}$. Such a compression method results in blocking artifacts for high compression ratios.

These artifacts manifest themselves as artificial discontinuities between adjacent blocks. It is a direct result of the independent processing of the blocks which does not take into account the between-block pixel correlations. They constitute a serious bottleneck for many important visual communication applications that require visually pleasing images at very high compression ratios. The reconstruction problem calls for finding an estimate of $\mathbf{f}$ given $\mathbf{g}$, $\mathcal{Q}$ and possibly knowledge about $\mathbf{f}$. The advances in VLSI technology will result in the incorporation of recovery algorithms at the decoders, and will bridge the conflicting requirements of high-quality images and high compression ratios.

In the past various algorithms have been proposed to improve the quality of block-transform compressed images at the decoder without increasing the bit-rate. In fact, in the JPEG standard [3] a technique for predicting the AC coefficients is recommended. However, in areas with sharp intensity transitions such a prediction scheme performs poorly.

In the mid eighties Reeves and Lim [4], Ramamurthi and Gersho [5] and Baskurt, Prost and Goutte [6] applied filtering and restoration techniques to the blocking artifact removal problem. However, it has been in this decade when the reconstruction of block compressed images in the spatial domain has flourished and a large number of papers have been published using a large variety of methods on the spatial domain.

Sauer pointed out in [7] that: "Classical spatially-invariant filtering techniques are of little use in removing this signal-dependent reconstruction error". A bit later, Zakhor [8] presented a reconstruction algorithm, commented by Reeves and Eddins [9], based on the POCS theory with two convex sets. One of the convex sets is equivalent to a convolution with an ideal low-pass filter and the other one deals with the quantization intervals of the transform coefficients. Conclusions were similar to Sauer's ones; "the low-pass filtering by itself could remove the blockiness but at the expense of increased blurriness". Yang, *et al.* [10] also used the POCS theory, as well as a constrained least-squares iteration, to smooth block boundaries. After each iteration a constraint on the quantization intervals of the transform coefficients was also applied. However, they used a non-adaptive model and only one regularization parameter so block boundaries on smooth and detailed zones were equally smoothed.

Taking into account the necessity for spatially adaptive techniques, different methods have been proposed. Stevenson [11] proposed a stochastic regularization technique using a statistical model for the image based on the convex Huber prior function. This function is quadratic for values smaller than a threshold, like block boundaries, and linear for greater values like natural discontinuities. Since prior and penalty are convex, the MAP estimation can be performed by gradient descent. This initial work has been developed in [12], [13], [14] and was the basis for the works of Luo *et al.* [15], where a Huber-Markov Random Field based model is optimized by POCS or using ICM and incorporating two different values for the threshold; one for block boundaries and another for the pixels inside the block. Another Markov Random Field based method, using Mean Field Annealing for the MAP estimation, was proposed by Özcelik, *et al.* [16] for the removal of blocking artifacts in still images, as well as, in video.

The method proposed in [10], was later extended in [17] to include spatial adaptivity. Based on these methods, Paek, Park and Lee [18] apply a similar technique to block boundaries and the pixels inside the block, and Kwak and Haddad [19] uses the method in [17] but cancel out the DCT-IDCT pair needed for the constraint on the quantization intervals of the transform coefficients.

The method proposed in [17], like most of the methods mentioned above, requires the estimation of unknown parameters. These parameters are usually estimated using ad hoc techniques. In this paper we formulate a similar algorithm to the one in [17] within the hierarchical Bayesian paradigm [20], [21]. This algorithm reconstructs the image and estimates the regularization parameters at the same time. We then proceed to estimate these parameters at the encoder using the original image. After transmission they can be combined at the decoder with the ones obtained from the reconstructed image. We show how this combination can be made, rigorously, within the hierarchical Bayesian approach to the reconstruction problem. A simpler idea was used in [22] to transmit to the decoder information of which macroblocks should be processed and which should remain unchanged.

The rest of the paper is organized as follows. In section II we introduce the required notation. Section III describes the hierarchical Bayesian approach to the reconstruction problem. The adaptive prior and noise models for the reconstruction problem are described in subsection III-A. The distributions of the unknown hyperparameters, that is, the second stage of the hierarchical Bayesian paradigm, are defined in subsection III-B. In section IV we describe the use of the evidence analysis for the simultaneous estimation of the hyperparameters and the reconstructed image. We also show how the parameters obtained from the original and reconstructed images can be combined following the evidence analysis. Experimental results are presented in section V and, finally, section VI concludes the paper.

II. Notation

Let us assume that $k \times k$ blocks are used to transform the $M \times N$ image, where M and N are multiples of k. Since for the removing of blocking artifacts we will be only operating on the block boundary pixels let us introduce the needed notation to characterize these image pixels.

Let $\mathbf{f}_{cl}$ be a column vector formed by stacking all the elements of $\mathbf{f}$ which are on the left of a block boundary column, cl, but not on a four-block intersection (see figure 1), that is,

$$\begin{aligned}\mathbf{f}_{cl} \;=\; & \{\mathbf{f}(u) \mid u = (x, y) \text{ with } x = k * i + l,\ y = k * j,\ i = 0, 1, 2, \ldots, M/k - 1, \\ & j = 1, 2, \ldots, N/k - 1,\ l = 2, 3, \ldots, k - 1\}.\end{aligned}$$

Let also $\mathbf{f}_{cr}$ be the column vector formed by stacking all the elements of $\mathbf{f}$ that are on the right of a block boundary column, cr, but not on a four-block intersection (see figure 1). The same way $\mathbf{g}_{cl}$ and $\mathbf{g}_{cr}$ are formed from $\mathbf{g}$.

We also define the vectors

$$\mathbf{f}_c^t = \{(\mathbf{f}_{cl}(i)\ \mathbf{f}_{cr}(i),\ i = 1, \ldots, p)\}, \qquad \mathbf{g}_c^t = \{(\mathbf{g}_{cl}(i)\ \mathbf{g}_{cr}(i),\ i = 1, \ldots, p)\},$$

with

$$p = (k - 2)(M/k) \times (N/k - 1). \tag{1}$$

Note that $\mathbf{f}_{cl}$, $\mathbf{f}_{cr}$, $\mathbf{g}_{cl}$ and $\mathbf{g}_{cr}$ are $p \times 1$ vectors.

In a similar way we define $\mathbf{f}_{ra}$, $\mathbf{g}_{ra}$, $\mathbf{f}_{rb}$ and $\mathbf{g}_{rb}$, the column vectors representing the rows above and below a row block boundary of $\mathbf{f}$ and $\mathbf{g}$, respectively (see figure 1). Related to them are the vectors

$$\mathbf{f}_r^t = \{(\mathbf{f}_{ra}(i)\ \mathbf{f}_{rb}(i),\ i = 1, \ldots, q)\}, \qquad \mathbf{g}_r^t = \{(\mathbf{g}_{ra}(i)\ \mathbf{g}_{rb}(i),\ i = 1, \ldots, q)\},$$

with

$$q = (k - 2)(N/k) \times (M/k - 1). \tag{2}$$

We further stack all the elements of $\mathbf{f}$ above a horizontal boundary and to the left of a vertical boundary, indicated by al in figure 1, into vector $\mathbf{f}_{al}$, that is,

$$\mathbf{f}_{al} = \{\mathbf{f}(u) \mid u = (x, y),\ \text{with } x = k * i,\ y = k * j,\ i = 1, 2, \ldots, M/k - 1,\ j = 1, 2, \ldots, N/k - 1\}.$$

Similarly we form vectors $\mathbf{f}_{ar}$, $\mathbf{f}_{br}$ and $\mathbf{f}_{bl}$, and the observation vectors for these pixels in a four-block boundary $\mathbf{g}_{al}$, $\mathbf{g}_{ar}$, $\mathbf{g}_{br}$ and $\mathbf{g}_{bl}$ (see figure 1). Using the vectors above, we also define

$$\begin{aligned}\mathbf{f}_x^t &= \{(\mathbf{f}_{al}(i)\ \mathbf{f}_{ar}(i)\ \mathbf{f}_{br}(i)\ \mathbf{f}_{bl}(i),\ i = 1, \ldots, m)\}, \\ \mathbf{g}_x^t &= \{(\mathbf{g}_{al}(i)\ \mathbf{g}_{ar}(i)\ \mathbf{g}_{br}(i)\ \mathbf{g}_{bl}(i),\ i = 1, \ldots, m)\},\end{aligned}$$

with

$$m = (M/k - 1) \times (N/k - 1). \tag{3}$$

$$\mathbf{f}^t = (\mathbf{f}_c^t \ \mathbf{f}_r^t \ \mathbf{f}_x^t) \ \text{ and } \ \mathbf{g}^t = (\mathbf{g}_c^t \ \mathbf{g}_r^t \ \mathbf{g}_x^t),$$

when needed. Note that originally $\mathbf{f}$ represented the complete original image but since the method we are proposing only modifies the pixels at the boundaries we redefine $\mathbf{f}$ as shown above.

III. Hierarchical Bayesian Paradigm

The hierarchical Bayesian paradigm is currently being applied to many areas of research related to image analysis. Buntine [23] has applied this theory to the construction of classification trees and Spiegelhalter and Lauritzen [24] to the problem of refining probabilistic networks. Buntine [25] and Cooper and Herkovsits [26] have used the same framework for constructing such networks. MacKay [27] and Buntine and Weigund [28] use the full Bayesian framework in backpropagation networks. This framework is also applied to interpolation (Gull [29] and Mackay [30]) and restoration (Molina *et al.* [20]) problems.

In the hierarchical approach to image reconstruction we have at least two stages. In the first stage, knowledge about the structural form of the noise and the structural behavior of the reconstructed image is used in forming $p(\mathbf{g} \mid \mathbf{f}, \beta)$ and $p(\mathbf{f} \mid \alpha)$, respectively. These noise and image models depend on the unknown hyperparameters or hypervectors α and β which are inverses of variances. In the second stage the hierarchical Bayesian paradigm defines a hyperprior on the hyperparameters, where information about these hyperparameters is included.

Although in some cases it would be possible to know, from previous experience, relations between the hyperparameters, we shall study here the model where the global probability is defined as

$$p(\alpha, \beta, \mathbf{f}, \mathbf{g}) = p(\alpha)p(\beta)p(\mathbf{f} \mid \alpha)p(\mathbf{g} \mid \mathbf{f}, \beta). \tag{4}$$

Once $p(\alpha, \beta, \mathbf{f}, \mathbf{g})$ is defined, the Bayesian analysis can be carried out in two different ways. In the evidence framework, $p(\alpha, \beta, \mathbf{f}, \mathbf{g})$ is integrated over $\mathbf{f}$ to give the evidence $p(\alpha, \beta \mid \mathbf{g})$ which is then maximized over the hyperparameters and the reconstruction is obtained for these hyperparameters. In the MAP framework $p(\alpha, \beta, \mathbf{f}, \mathbf{g})$ is first integrated over α and β and then maximized with respect to $\mathbf{f}$.

In this work we shall adopt the evidence analysis instead of the MAP analysis. We have found that this analysis provides better results for restoration-reconstruction problems (see [20]). According to this approach the simultaneous estimation of $\mathbf{f}$, α and β is done as follows:

- Parameter estimation step:

$$\hat{\alpha}, \hat{\beta} = \arg\max_{\alpha,\beta} \{p(\alpha, \beta \mid \mathbf{g})\} \tag{5}$$

- Reconstruction step:

$$\mathbf{f}^{(\hat{\alpha},\hat{\beta})} = \arg\max_{\mathbf{f}} \left\{p(\mathbf{f} \mid \mathbf{g}, \hat{\alpha}, \hat{\beta})\right\}. \tag{6}$$

We examine next the components of the first and second stages for the deblocking problem.

A. Components of the First Stage: Noise and Image Models

A.1 Adaptive Image Model

In defining the model for the reconstructed image we take into account the local properties of the images. This provides us with a locally adaptive image model. To capture the vertical local properties of the image we define a $p \times p$ diagonal matrix, $\mathbf{W}_c$, with p defined in Eq. (1), of the form [17]

$$\mathbf{W}_c = \begin{bmatrix} \omega_c(1) & 0 & 0 & \cdots & 0 \\ 0 & \omega_c(2) & 0 & \cdots & \vdots \\ 0 & 0 & \ddots & 0 & \vdots \\ \vdots & \vdots & 0 & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & \omega_c(p) \end{bmatrix}, \tag{7}$$

where the $\omega_c(i)$'s, $i = 1, 2, \ldots, p$, determine the relative importance of the intensity differences across a vertical block boundary as shown in figure 2 (note that the pixels in a four block boundary are not weighted by this matrix). Analogously, we define $\mathbf{W}_r$ to capture the horizontal local properties of the image, as shown in figure 2. The size of this matrix is $q \times q$, where q is defined in Eq. (2).

For pixels in a four block intersection we take into account the differences $\mathbf{f}_{al} - \mathbf{f}_{ar}$, $\mathbf{f}_{ar} - \mathbf{f}_{br}$, $\mathbf{f}_{br} - \mathbf{f}_{bl}$ and $\mathbf{f}_{bl} - \mathbf{f}_{al}$, corresponding to column, row, column and row differences, respectively (see figures 1 and 2). The corresponding weight matrices are denoted by $\mathbf{W}_{x1}, \mathbf{W}_{x2}, \mathbf{W}_{x3}$ and $\mathbf{W}_{x4}$ with entries w_c, w_r, w_c and w_r respectively (see figure 2). All these matrices have size $m \times m$, where m is defined in Eq. (3).

Let us now consider the adaptive image model we will be using. We assume that the degree of smoothness in the vertical and horizontal directions is different, resulting in the use of two smoothing parameters. More specifically, for the vertical block pixels, we propose the use of

$$p(\mathbf{f}_c \mid \alpha_c) \propto \alpha_c^{\frac{p}{2}} \exp\{-A_c(\mathbf{f}_c \mid \alpha_c)\}, \tag{8}$$

with

$$A_c(\mathbf{f}_c \mid \alpha_c) = \alpha_c \parallel \mathbf{W}_c(\mathbf{f}_{cl} - \mathbf{f}_{cr}) \parallel^2, \tag{9}$$

where p has been defined in Eq. (1), and α_c measures the roughness between two vertical block boundaries.

For the horizontal block pixels, we propose the use of the following image model

$$p(\mathbf{f}_r \mid \alpha_r) \propto \alpha_r^{\frac{q}{2}} \exp\left\{-A_r(\mathbf{f}_r \mid \alpha_r)\right\},$$

with

$$A_r(\mathbf{f}_r \mid \alpha_r) = \alpha_r \parallel \mathbf{W}_r(\mathbf{f}_{ra} - \mathbf{f}_{rb}) \parallel^2, \tag{10}$$

where q has been defined in Eq. (2), and α_r measures the roughness between two horizontal block boundaries.

Finally, for the four block intersection pixels we have

$$p(\mathbf{f}_x \mid \alpha_c, \alpha_r) \quad \propto \quad F(\alpha_c, \alpha_r) \exp\left\{-A_x(\mathbf{f}_x \mid \alpha_c, \alpha_r)\right\},$$

where

$$\begin{aligned} A_x(\mathbf{f}_x \mid \alpha_c, \alpha_r) \quad = \quad & \frac{1}{2}\Big\{\alpha_c \parallel \mathbf{W}_{x1}(\mathbf{f}_{al} - \mathbf{f}_{ar}) \parallel^2 + \alpha_r \parallel \mathbf{W}_{x2}(\mathbf{f}_{ar} - \mathbf{f}_{br}) \parallel^2 \\ & + \alpha_c \parallel \mathbf{W}_{x3}(\mathbf{f}_{br} - \mathbf{f}_{bl}) \parallel^2 + \alpha_r \parallel \mathbf{W}_{x4}(\mathbf{f}_{bl} - \mathbf{f}_{al}) \parallel^2\Big\}, \end{aligned} \tag{11}$$

and

$$F(\alpha_c, \alpha_r) = \prod_{i=1}^{m} \left[\alpha_c \alpha_r \left(\alpha_c \left(\frac{1}{\omega_{x2}^2(i)} + \frac{1}{\omega_{x4}^2(i)}\right) + \alpha_r \left(\frac{1}{\omega_{x1}^2(i)} + \frac{1}{\omega_{x3}^2(i)}\right)\right)\right]^{\frac{1}{2}}. \tag{12}$$

The image model we therefore propose is given by

$$\begin{aligned} p(\mathbf{f} \mid \alpha_c, \alpha_r) \quad = \quad & p(\mathbf{f}_c, \mathbf{f}_r, \mathbf{f}_x \mid \alpha_c, \alpha_r) = p(\mathbf{f}_c \mid \alpha_c) p(\mathbf{f}_r \mid \alpha_r) p(\mathbf{f}_x \mid \alpha_c, \alpha_r) \\ \propto \quad & \alpha_c^{\frac{p}{2}} \alpha_r^{\frac{q}{2}} F(\alpha_c, \alpha_r) \exp\left\{-A(\mathbf{f} \mid \alpha_c, \alpha_r)\right\}, \end{aligned} \tag{13}$$

with

$$A(\mathbf{f} \mid \alpha_c, \alpha_r) = A_c(\mathbf{f}_c \mid \alpha_c) + A_r(\mathbf{f}_r \mid \alpha_r) + A_x(\mathbf{f}_x \mid \alpha_c, \alpha_r). \tag{14}$$

A.2 Noise Model

This model takes into account the fidelity to the observed data. It is given by

$$\begin{aligned} p(\mathbf{g} \mid \mathbf{f}, \beta) \quad = \quad & p(\mathbf{g}_c, \mathbf{g}_r, \mathbf{g}_x \mid \mathbf{f}_c, \mathbf{f}_r, \mathbf{f}_x, \beta) = p(\mathbf{g}_c \mid \mathbf{f}_c, \beta) p(\mathbf{g}_r \mid \mathbf{f}_r, \beta) p(\mathbf{g}_x \mid \mathbf{f}_x, \beta) \\ \propto \quad & \beta^p \beta^q \beta^{2m} \exp\left\{-(B_c(\mathbf{g}_c \mid \mathbf{f}_c, \beta) + B_r(\mathbf{g}_r \mid \mathbf{f}_r, \beta) + B_x(\mathbf{g}_x \mid \mathbf{f}_x, \beta))\right\} \\ = \quad & \beta^p \beta^q \beta^{2m} \exp\left\{-B(\mathbf{g} \mid \mathbf{f}, \beta)\right\}, \end{aligned}$$

where

$$B(\mathbf{g} \mid \mathbf{f}, \beta) = B_c(\mathbf{g}_c \mid \mathbf{f}_c, \beta) + B_r(\mathbf{g}_r \mid \mathbf{f}_r, \beta) + B_x(\mathbf{g}_x \mid \mathbf{f}_x, \beta), \tag{15}$$

with

$$B_c(\mathbf{g}_c \mid \mathbf{f}_c, \beta) = \frac{1}{2}\beta \left\{ \| \mathbf{g}_{cl} - \mathbf{f}_{cl} \|^2 + \| \mathbf{g}_{cr} - \mathbf{f}_{cr} \|^2 \right\}, \tag{16}$$

$$B_r(\mathbf{g}_r \mid \mathbf{f}_r, \beta) = \frac{1}{2}\beta \left\{ \| \mathbf{g}_{ra} - \mathbf{f}_{ra} \|^2 + \| \mathbf{g}_{rb} - \mathbf{f}_{rb} \|^2 \right\}, \tag{17}$$

$$B_x(\mathbf{g}_x \mid \mathbf{f}_x, \beta) = \frac{1}{2}\beta \left\{ \| \mathbf{g}_{al} - \mathbf{f}_{al} \|^2 + \| \mathbf{g}_{ar} - \mathbf{f}_{ar} \|^2 + \| \mathbf{g}_{br} - \mathbf{f}_{br} \|^2 + \| \mathbf{g}_{bl} - \mathbf{f}_{bl} \|^2 \right\}. \tag{18}$$

It is mentioned here that there is no need to use different β's for the columns, rows and four block intersections, as was the case with the α's. This is the case because the noise does not depend on the image characteristics.

B. Components of the Second Stage

Because of the attractiveness of the Bayesian machinery to perform conditional analysis, Berger [31] describes the possibility of using the Bayesian approach when very little prior information is available. According to it, in situations without prior information what is needed is a non informative prior on the hyperparameters (the term "non informative" prior is meant to imply that no information about the hyperparameters is contained in the prior). For the problem at hand we use improper non informative priors $p(\omega) \propto const$, where ω denotes α_c, α_r or β.

However, we propose to also incorporate precise prior knowledge about the value of the noise and prior variances. To do so we use as hyperprior the gamma distribution defined by

$$p(\omega) \propto \omega^{l-1} \exp[-al\omega],$$

where $l > 1$, ω denotes a hyperparameter and a and l are explained below. This distribution has the following properties

$$E[w] = a^{-1} \qquad \text{and} \qquad Var[w] = [a^2 l]^{-1}.$$

So, the mean of w, which represents the inverse of the prior or noise variance, is equal to $1/a$, and its variance decreases when l increases. l can then be understood as a measure of the certainty on the knowledge about the prior or noise variances (see [32], [33]).

We show in section IV-B how these distributions can be used to combine information between the coder and decoder.

IV. Hierarchical Bayesian Analysis

Having defined $p(\alpha_c, \alpha_r, \beta, \mathbf{f}, \mathbf{g})$, the Bayesian analysis is performed. As mentioned earlier, in the evidence framework, $p(\alpha_c, \alpha_r, \beta, \mathbf{f}, \mathbf{g})$ is integrated over $\mathbf{f}$ to give the evidence $p(\alpha_c, \alpha_r, \beta \mid \mathbf{g})$ which is maximized over the hyperparameters and the reconstruction is obtained for those hyperparameters.

The analysis is made in the following sections using both flat and gamma hyperpriors, which, as it will be explained later, corresponds to having no information about the parameters at the decoder and having some information about them from the coder at the decoder.

A. Evidence Analysis for Flat Priors

As it has been already mentioned, in the evidence approach $\hat{\alpha}_c$, $\hat{\alpha}_r$ and $\hat{\beta}$ are first selected as

$$\hat{\alpha}_c, \hat{\alpha}_r, \hat{\beta} = \arg \max_{\alpha_c, \alpha_r, \beta} p(\alpha_c, \alpha_r, \beta \mid \mathbf{g}), \tag{19}$$

and then the reconstruction, $\mathbf{f}^{(\hat{\alpha}_c, \hat{\alpha}_r, \hat{\beta})}$, is obtained as

$$\mathbf{f}^{(\hat{\alpha}_c, \hat{\alpha}_r, \hat{\beta})} = \arg \min_{\mathbf{f}} M(\mathbf{f}, \mathbf{g} \mid \hat{\alpha}_c, \hat{\alpha}_r, \hat{\beta}), \tag{20}$$

where

$$M(\mathbf{f}, \mathbf{g} \mid \alpha_c, \alpha_r, \beta) = M_c(\mathbf{f}_c, \mathbf{g}_c \mid \alpha_c, \beta) + M_r(\mathbf{f}_r, \mathbf{g}_r \mid \alpha_r, \beta) + M_x(\mathbf{f}_x, \mathbf{g}_x \mid \alpha_c, \alpha_r, \beta), \tag{21}$$

with

$$M_c(\mathbf{f}_c, \mathbf{g}_c \mid \alpha_c, \beta) = A_c(\mathbf{f}_c \mid \alpha_c) + B_c(\mathbf{g}_c \mid \mathbf{f}_c, \beta), \tag{22}$$

$$M_r(\mathbf{f}_r, \mathbf{g}_r \mid \alpha_r, \beta) = A_r(\mathbf{f}_r \mid \alpha_r) + B_r(\mathbf{g}_r \mid \mathbf{f}_r, \beta), \tag{23}$$

$$M_x(\mathbf{f}_x, \mathbf{g}_x \mid \alpha_c, \alpha_r, \beta) = A_x(\mathbf{f}_x \mid \alpha_c, \alpha_r) + B_x(\mathbf{g}_x \mid \mathbf{f}_x, \beta), \tag{24}$$

and $A_u(.)$ and $B_u(.)$, $(u \in \{c, r, x\})$, defined in Eqs. (9)–(11) and (16)–(18), respectively. With the flat hyperpriors, Eq. (19) amounts to selecting $\hat{\alpha}_c$, $\hat{\alpha}_r$, $\hat{\beta}$ as the maximum likelihood estimates (*mle*) of α_c, α_r, β from $p(\mathbf{g} \mid \alpha_c, \alpha_r, \beta)$.

Let us now examine the process to estimate the hyperparameters. To do so, we may use the row, column and four block intersection observation pixels to obtain the posterior distribution of the hyperparameters, Eq. (19), or use only the row and column pixels. In this case we will have an approximation of the posterior distribution whose maximization will provide another estimation of the hyperparameters. Let us describe both approaches in details.

A.1 Using row, column and four block intersection pixels to estimate the hyperparameters

To estimate the hyperparameters and the image we will proceed in two steps:

1. Estimate $\alpha_c, \alpha_r, \beta$ by (see Appendix A-A)

$$\hat{\alpha}_c, \hat{\alpha}_r, \hat{\beta} = \arg \max_{\alpha_c, \alpha_r, \beta} p(\mathbf{g} \mid \alpha_c, \alpha_r, \beta) = \arg \max_{\alpha_c, \alpha_r, \beta} p(\mathbf{g}_c \mid \alpha_c, \beta) p(\mathbf{g}_r \mid \alpha_r, \beta) p(\mathbf{g}_x \mid \alpha_c, \alpha_r, \beta). \tag{25}$$

2. Use $\hat{\alpha}_c, \hat{\alpha}_r, \hat{\beta}$ in Eq. (20) to obtain the reconstructed image (see Appendix A-C).

The following algorithm, which takes into account the information provided by the four block intersections pixels together with the row and column pixels, is proposed for the simultaneous estimation of the hyperparameters and the image.

Algorithm 1:

1. Choose α_c^0, α_r^0 and β^0.
2. Compute $\mathbf{f}^{(\alpha_c^0, \alpha_r^0, \beta^0)}$ by solving equations (A-11)–(A-18).
3. For $k = 1, 2, \ldots$

(a) Estimate α_c^k, α_r^k and β^k by substituting α_c^{k-1}, α_r^{k-1} and β^{k-1} in the right hand side of Eqs. (A-2), (A-3) and (A-4).

(b) Compute $\mathbf{f}^{(\alpha_c^k, \alpha_r^k, \beta^k)}$ by solving equations (A-11)–(A-18).

4. Go to 3 until $\| \mathbf{f}^{(\alpha_c^k, \alpha_r^k, \beta^k)} - \mathbf{f}^{(\alpha_c^{k-1}, \alpha_r^{k-1}, \beta^{k-1})} \|$ is less than a prescribed bound.

The convergence of this algorithm is established in Appendix A-A.

A.2 Using only rows and columns to estimate the hyperparameters

The inclusion of $p(\mathbf{g}_x \mid \alpha_c, \alpha_r, \beta)$ in Eq. (25) slows the estimation process down, since the inversion of $(M/k - 1)(N/k - 1)$ 4×4 matrices is required. For this reason we modify the estimation process by using the following two steps

1. Estimate $\alpha_c, \alpha_r, \beta$ by (see Appendix A-B)

$$\hat{\alpha}_c, \hat{\alpha}_r, \hat{\beta} = \arg \max_{\alpha_c, \alpha_r, \beta} p(\mathbf{g}_c \mid \alpha_c, \beta) p(\mathbf{g}_r \mid \alpha_r, \beta). \tag{26}$$

2. Use $\hat{\alpha}_c, \hat{\alpha}_r, \hat{\beta}$ in Eq. (20) to obtain the reconstructed image (see Appendix A-C).

Notice that Eq. (26) is in fact an approximation of Eq. (25).

The following algorithm is proposed for the simultaneous estimation of the hyperparameters and the image using this approximation.

Algorithm 2:

1. Choose α_c^0, α_r^0 and β^0.
2. Compute $\mathbf{f}_c^{(\alpha_c^0, \alpha_r^0, \beta^0)}$ and $\mathbf{f}_r^{(\alpha_c^0, \alpha_r^0, \beta^0)}$ from Eqs. (A-11), (A-12) and Eqs. (A-13), (A-14), respectively.
3. For $k = 1, 2, \ldots$

(a) Estimate α_c^k, α_r^k and β^k by substituting α_c^{k-1}, α_r^{k-1} and β^{k-1} in the right hand side of Eqs. (A-6), (A-7) and (A-8).

(b) Compute $\mathbf{f}_c^{(\alpha_c^k, \alpha_r^k, \beta^k)}$ and $\mathbf{f}_r^{(\alpha_c^k, \alpha_r^k, \beta^k)}$ from Eqs. (A-11), (A-12) and Eqs. (A-13), (A-14), respectively

4. Go to 3 until $\| \mathbf{f}_c^{(\alpha_c^k,\alpha_r^k,\beta^k)} - \mathbf{f}_c^{(\alpha_c^{k-1},\alpha_r^{k-1},\beta^{k-1})} \| + \| \mathbf{f}_r^{(\alpha_c^k,\alpha_r^k,\beta^k)} - \mathbf{f}_r^{(\alpha_c^{k-1},\alpha_r^{k-1},\beta^{k-1})} \|$ is less than a prescribed bound.

5. Using α_c^k, α_r^k, β^k calculate $\mathbf{f}_x^{(\alpha_c^k,\alpha_r^k,\beta^k)}$ by solving equations (A-15)–(A-18).

The convergence of this algorithm is established by realizing that it corresponds to the EM algorithm, where the complete data are the observations $\mathbf{g}$ and the unknown reconstruction $\mathbf{f}$, that is $\mathbf{z}^t = (\mathbf{f}^t \quad \mathbf{g}^t)$ and

$$\mathbf{g} = (\mathbf{I} \quad 0)\,\mathbf{z}.$$

Details are provided in [20].

B. Combining information from the coder: Gamma Priors

It is clear that the described process for estimating the image and the hyperparameters can also be performed at the coder, where we use the original image $\mathbf{f}$ as observation $\mathbf{g}$ and again flat hyperpriors for the hyperparameters. In this case Eq. (20) becomes

$$\begin{aligned} \mathbf{f}^{(\alpha_c,\alpha_r,\beta)cod} &= \arg\min_{\mathbf{z}} \{M(\mathbf{z},\mathbf{f} \mid \alpha_c,\alpha_r,\beta)\} \\ &= \arg\min_{\mathbf{z}} \{A(\mathbf{z} \mid \alpha_c,\alpha_r) + B(\mathbf{f} \mid \mathbf{z},\beta)\}, \end{aligned} \tag{27}$$

and the hyperparameters are also estimated using the original image as observation, that is,

$$\hat{\alpha}_c^{cod}, \hat{\alpha}_r^{cod}, \hat{\beta}^{cod} = \arg\max_{\alpha_c,\alpha_r,\beta} \int_{\mathbf{z}} p(\mathbf{z},\mathbf{f} \mid \alpha_c,\alpha_r,\beta) d\mathbf{z}. \tag{28}$$

It is clear that to obtain $\hat{\alpha}_c^{cod}$, $\hat{\alpha}_r^{cod}$ and $\hat{\beta}^{cod}$ we only need to run Algorithm 1 or Algorithm 2 using the original image as observation.

A (quantized) version of $\hat{\alpha}_c^{cod}$, $\hat{\alpha}_r^{cod}$ and $\hat{\beta}^{cod}$ is received by the decoder, and denoted respectively by m_c^{cod}, m_r^{cod} and n^{cod}. They are used as prior information in guiding the estimation of the hyperparameters at the decoder. More specifically, they are used in defining the following hyperpriors for each hyperparameter

$$p(\alpha_c) \propto \alpha_c^{l(m_c^{cod})-1} \exp[-l(m_c^{cod})\alpha_c / m_c^{cod}], \tag{29}$$

$$p(\alpha_r) \propto \alpha_r^{l(m_r^{cod})-1} \exp[-l(m_r^{cod})\alpha_r / m_r^{cod}], \tag{30}$$

$$p(\beta) \propto \beta^{l(n^{cod})-1} \exp[-l(n^{cod})\beta / n^{cod}]. \tag{31}$$

Following again the hierarchical Bayesian approach to the reconstruction problem and using the gamma distributions in Eqs. (29)–(31), we perform the estimation of the hyperparameters and the reconstruction using the following two steps

1. Estimate $\alpha_c, \alpha_r, \beta$ by (see Appendix B-A)

$$\hat{\alpha}_c, \hat{\alpha}_r, \hat{\beta} = \arg \max_{\alpha_c, \alpha_r, \beta} p(\alpha_c)p(\alpha_r)p(\beta)p(\mathbf{g}_c \mid \alpha_c, \beta)p(\mathbf{g}_r \mid \alpha_r, \beta). \tag{32}$$

where $p(\alpha_c)$, $p(\alpha_r)$ and $p(\beta)$ have been defined in Eqs. (29)–(31).

2. Use $\hat{\alpha}_c, \hat{\alpha}_r, \hat{\beta}$ in Eq. (20) to obtain the reconstructed image (see Appendix A-C).

We notice that Eq. (32) is again an approximation of the true posterior distribution of the hyperparameters and that it would also be possible to use

$$\hat{\alpha}_c, \hat{\alpha}_r, \hat{\beta} = \arg \max_{\alpha_c, \alpha_r, \beta} p(\alpha_c)p(\alpha_r)p(\beta)p(\mathbf{g}_c \mid \alpha_c, \beta)p(\mathbf{g}_r \mid \alpha_r, \beta)p(\mathbf{g}_x \mid \alpha_c, \alpha_r, \beta). \tag{33}$$

The derivation of the parameter estimation step when Eq. (33) is used instead of Eq. (32) is similar to the process described in Appendix B-A and it will therefore not be shown here. We notice that the reconstruction step is the same for the flat and gamma hyperprior cases.

Using steps 1 and 2 above the following algorithm is proposed for the simultaneous estimation of the hyperparameters and the image assuming gamma hyperpriors.

Algorithm 3:

1. Choose α_c^0, α_r^0 and β^0.
2. Compute $\mathbf{f}_c^{(\alpha_c^0, \alpha_r^0, \beta^0)}$ and $\mathbf{f}_r^{(\alpha_c^0, \alpha_r^0, \beta^0)}$ from Eqs. (A-11), (A-12) and Eqs. (A-13), (A-14), respectively.
3. For $k = 1, 2, \ldots$

(a) Estimate α_c^k, α_r^k and β^k by substituting α_c^{k-1}, α_r^{k-1} and β^{k-1} in the right hand side of Eqs. (B-2), (B-3) and (B-4).

(b) Compute $\mathbf{f}_c^{(\alpha_c^k, \alpha_r^k, \beta^k)}$ and $\mathbf{f}_r^{(\alpha_c^k, \alpha_r^k, \beta^k)}$ from Eqs. (A-11), (A-12) and Eqs. (A-13), (A-14), respectively.

4. Go to 3 until $\| \mathbf{f}_c^{(\alpha_c^k, \alpha_r^k, \beta^k)} - \mathbf{f}_c^{(\alpha_c^{k-1}, \alpha_r^{k-1}, \beta^{k-1})} \| + \| \mathbf{f}_r^{(\alpha_c^k, \alpha_r^k, \beta^k)} - \mathbf{f}_r^{(\alpha_c^{k-1}, \alpha_r^{k-1}, \beta^{k-1})} \|$ is less than a prescribed bound.
5. Using α_c^k, α_r^k, β^k calculate $\mathbf{f}_x^{(\alpha_c^k, \alpha_r^k, \beta^k)}$ by solving equations (A-15)–(A-18).

The proof of the convergence of this algorithm is again based on the fact that it is an EM algorithm. (see [34]).

Assuming that $p \simeq p - 2$ and $q \simeq q - 2$, we can write the Equations (B-2), (B-3) and (B-4) as

$$\frac{1}{\alpha_c^k} = \mu_c \frac{1}{m_c^{cod}} + (1 - \mu_c)\frac{1}{\alpha_c^{k,dec}}, \tag{34}$$

$$\frac{1}{\alpha_r^k} = \mu_r \frac{1}{m_r^{cod}} + (1 - \mu_r)\frac{1}{\alpha_r^{k,dec}}, \tag{35}$$

$$\frac{1}{\beta^k} = \nu \frac{1}{n^{cod}} + (1 - \nu)\frac{1}{\beta^{k,dec}}, \tag{36}$$

where

$$\mu_c = \frac{2l(m_c^{cod})}{2l(m_c^{cod})+p}, \tag{37}$$

$$\mu_r = \frac{2l(m_r^{cod})}{2l(m_r^{cod})+q}, \tag{38}$$

$$\nu = \frac{l(n^{cod})}{l(n^{cod})+p+q}, \tag{39}$$

and $\alpha_c^{k,dec}$, $\alpha_r^{k,dec}$ and $\beta^{k,dec}$ are calculated by substituting α_c^{k-1}, α_r^{k-1} and β^{k-1} in the right hand side of Eqs. (A-6), (A-7) and (A-8).

We note now that μ_c, μ_r and ν can be understood as normalized confidence parameters. They take values in the interval $[0, 1)$ and so when they are zero no confidence is put on the corresponding transmitted mean (full estimation at the decoder, that is, m_c^{cod}, m_r^{cod} and n^{cod} are not transmitted or used), 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 at the decoder). It is then clear that $l(m_c^{cod})$, $l(m_r^{cod})$ and $l(n^{cod})$ depend on the confidence we have in their corresponding means, m_c^{cod}, m_r^{cod} and n^{cod}.

V. Experimental Results

In obtaining the required weights, $\omega_c(i)$ and $\omega_r(i)$, the pixel at location (i) is treated as a random variable with mean $\mu(i)$ and variance $\sigma^2(i)$. The mean serves as a measure of the local brightness and the variance as a measure of the local detail. Several forms of the weighting functions have been previously suggested [17] . The one we use in this work is given by

$$\omega_c(i) = \log\left(1 + \frac{\sqrt{\mu(i)}}{1+\sigma(i)}\right). \tag{40}$$

It represents a compressed range function which decreases as a function of $\sigma(i)$ and captures the intensity changes since, in our experiments, the blocking artifacts seem to be more visible in bright rather than in dark areas of the image. We follow [17] in obtaining $\mu(i)$ and $\sigma(i)$. The DC coefficient in each block is the average (to a constant) of the pixel intensity within this block. Let us consider the vertical boundary l_c and let DC_L and DC_R denote the DC coefficient of its left and right block, respectively. Then, the estimate of the mean $\mu_c(l_c)$ is given by

$$\hat{\mu}_c(l_c) = \frac{DC_L + DC_R}{2 \times k^2}. \tag{41}$$

The mean $\mu_r(l_r)$ is estimated in a similar fashion using in this case the DC coefficients of the blocks above and below this horizontal boundary, l_r.

The variance $\sigma_c^2(l_c)$ at the vertical block boundary l_c is estimated by

$$\hat{\sigma}_c^2(l_c) = \frac{VAC_L + VAC_R}{2 \times k^2}, \tag{42}$$

where VAC_L and VAC_R are the sums of the squared AC coefficients in the first column of the blocks to the left and to the right of the boundary l_c, respectively ($\sigma_r^2(l_r)$ is estimated in a similar fashion). Equations (41) and (42) are used in Eq. (40) to define the entries of the weight matrices $\mathbf{W}_c$. A similar process can be carried out to obtain $\mathbf{W}_r$ and $\mathbf{W}_x$. The scaled weight maps for the *Lena* image corresponding to $\mathbf{W}_c$ and $\mathbf{W}_r$ are shown in Figs. 3(a) and 3(b), respectively.

Two quantization matrices are used in the experiments presented here. They correspond to compression ratios of approximately 0.50 bpp and 0.30 bpp using the JPEG compression standard ('moderate' and 'high' compression ratio, respectively).

Experiments were carried out on a set of *de facto* standard 256 graylevel images of size 512×512 pixels in order to evaluate the performance of the proposed methods and compare them with other existing methods, such as, the POCS method proposed in [17] and the method proposed by Choy *et. al.* [35], which performs the estimation of all the DCT coefficients (in implementing this method we tried various windows sizes and chose the one which provided the best PSNR reconstruction). Results are presented with three of the images (*airplane*, *Lena* and *peppers*) representing various types and spatial activities.

In the first experiment the parameters are estimated at the decoder using Algorithms 1 and 2, with and without projecting the image coefficients on the quantization intervals after each iteration. In addition the POCS algorithms in [17] and the WLS algorithm in [35] are tested. The PSNR results are summarized in Table I. Similar results are obtained with the use of Algorithms 1 and 2, and the results they provide are better than the ones obtained by the other two algorithms. Furthermore, the process of projecting the image coefficients on the quantization intervals does not improve the performance of Algorithms 1 and 2. Figures 4 and 5 show the center part of the *Lena* image compressed at 0.29 bpp and the center part of the *peppers* image compressed at 0.32 bpp, respectively. They also show the reconstructed images using four different methods. Although the PSNR is used widely as an objective metric for measuring image quality, it does not fully reflect the visual quality of an image. Since the improvement of the visual quality of the image is the objective of the work described in this paper, the quality of the presented reconstructed images should be considered in addition to the reported PSNR values.

Next experiments were performed for estimating $\hat{\alpha}_c^{cod}$, $\hat{\alpha}_r^{cod}$ and $\hat{\beta}^{cod}$ by Algorithms 1 and 2 at the coder with the use of the original image as observation, and then using these parameters in Eq. (20) to

obtain the reconstruction. The results are shown in Table II. It can be seen that the PSNR improves slightly in this case.

The parameters obtained at the coder and the decoder were then combined. The same normalized confidence parameters μ_c and μ_r, defined in Eqs. (37) and (38), were used for α_c and α_r. The values used in the experiments were $\mu_c = \mu_r = \mu \in \{0.0, 0.1, ..., 1\}$. The normalized confidence parameter ν, defined in Eq. (39), belongs to the same range. The $3-D$ plot in Fig. 6 shows the PSNR as a function of μ and ν for the *Lena* highly compressed image. The center part of the compressed image and the best reconstruction, corresponding to the parameter values $m_c^{cod} = \alpha_c^{cod} = 30.82^{-1}$, $m_r^{cod} = \alpha_r^{cod} = 5.36^{-1}$ and $n^{cod} = \beta^{cod} = 36.36^{-1}$ with $\mu = 0.9$ and $\nu = 0.0$ is displayed in Fig. 7b. The corresponding PSNR is 31.40dB. Similar results are obtained using other high compressed images showing that best reconstructions in terms of PSNR are obtained using μ between 0.7 and 1.0 and $\nu = 0.0$.

From the above experiment we observe that using only the parameters obtained at the coder ($\mu = \nu = 1$, $m_c^{cod} = \alpha_c^{cod}$, $m_r^{cod} = \alpha_r^{cod}$ and $n^{cod} = \beta^{cod}$) we obtain better PSNR than when we estimate the parameters at the decoder and use $\mu = \nu = 0$. However, neither performing the estimation strictly at the coder nor strictly at the decoder provide the best PSNR. This may be due to the different characteristics of the images at the coder and at the decoder. When the estimation is performed at the coder, the noise variance of the image is usually smaller than the one estimated at the decoder, at least for many highly compressed images. It is, therefore usually better to estimate it at the decoder in order to obtain the best reconstruction. On the other hand, the prior variance at the coder is more accurate than the one estimated at the decoder since it was obtained from the original image and should be used at the decoder. These reasons could explain why the highest PSNR is obtained when μ is near 1.0 and $\nu = 0.0$.

However, most of the moderately compressed images tested show another behavior. In these cases, the better reconstruction in PSNR is obtained when both μ and ν are near 1.0. This may be due to the fact that, for moderate compression rates, the noise parameter value obtained at the coder is similar to the real one for the compressed image thus obtaining better reconstructions. Results on combining parameters obtained at the coder and the decoder are summarized in Table III.

The same experiment was performed using two quantized versions of $\alpha_c^{cod^{-1}}$, $\alpha_r^{cod^{-1}}$ and $\beta^{cod^{-1}}$ as $m_c^{cod^{-1}}$, $m_r^{cod^{-1}}$ and $n^{cod^{-1}}$, respectively. The obtained results, summarized in Tables IV and V, were similar to the ones obtained using the unquantized values.

VI. Conclusions

A new spatially-adaptive image recovery algorithm based on the Bayesian hierarchical approach has been proposed to decode BDCT based compressed image. Using this approach we have shown how to estimate the unknown hyperparameters using well grounded estimation procedures and how to incorporate from vague to precise knowledge about the unknown parameters into the recovery process. In addition we have shown how to rigorously combine the hyperparameters estimated at the coder and transmitted to the decoder, with those estimated at the decoder. The experimental results demonstrate and compare the performance of the algorithms with other algorithms. They show a good improvement in term of the PSNR metric and the visual quality of the reconstructed images, and an improved performance over the two algorithms it was compared against. The added advantage of the proposed algorithms is that no arbitrary choice of the hyperparameters is required, since they are rigorously estimated. Although the proposed method only includes smoothness between blocks, the extension to within blocks smoothness is currently being investigated.

Appendix

I. Derivation of the Evidence Analysis for Flat Hyperpriors

A. Parameter estimation step in Algorithm 1.

We now describe in details the estimation of α_c, α_r and β for the flat hyperprior case when Algorithm 1 is used. Note that we are taking into account the row, column and four block intersection pixels to estimate the hyperparameters.

First we take into account that

$$\begin{aligned} p(\alpha_c, \alpha_r, \beta \mid \mathbf{g}) &\propto \text{[for the flat hyperprior]} \int_{\mathbf{f}} p(\mathbf{f}, \mathbf{g} \mid \alpha_c, \alpha_r, \beta) d\mathbf{f} \\ &= \int_{\mathbf{f}_c} p(\mathbf{f}_c, \mathbf{g}_c \mid \alpha_c, \beta) d\mathbf{f}_c \int_{\mathbf{f}_r} p(\mathbf{f}_r, \mathbf{g}_r \mid \alpha_r, \beta) d\mathbf{f}_r \int_{\mathbf{f}_x} p(\mathbf{f}_x, \mathbf{g}_x \mid \alpha_c, \alpha_r, \beta) d\mathbf{f}_x \\ &= p(\mathbf{g}_c \mid \alpha_c, \beta) p(\mathbf{g}_r \mid \alpha_r, \beta) p(\mathbf{g}_x \mid \alpha_c, \alpha_r, \beta) = p(\mathbf{g} \mid \alpha_c, \alpha_r, \beta). \end{aligned} \tag{A-1}$$

Let us then proceed to calculate $p(\mathbf{g} \mid \alpha_c, \alpha_r, \beta)$ taking into account that

$$p(\mathbf{g} \mid \alpha_c, \alpha_r, \beta) = \alpha_c^{\frac{p}{2}} \beta^p \alpha_r^{\frac{q}{2}} \beta^q F(\alpha_c, \alpha_r) \beta^{2m} \int_{\mathbf{f}} \exp[-M(\mathbf{f}, \mathbf{g} \mid \alpha_c, \alpha_r, \beta)] \, d\mathbf{f},$$

where $F(\alpha_c, \alpha_r)$ has been defined in Eq. (12). Let us fix α_c,α_r and β and expand $M(\mathbf{f}, \mathbf{g} \mid \alpha_c, \alpha_r, \beta)$ around $\mathbf{f}^{(\alpha_c, \alpha_r, \beta)}$. We then have,

$$\begin{aligned} &M(\mathbf{f}, \mathbf{g} \mid \alpha_c, \alpha_r, \beta) = \\ &= \alpha_c \parallel \mathbf{W}_c(\mathbf{f}_{cl} - \mathbf{f}_{cr}) \parallel^2 + \alpha_r \parallel \mathbf{W}_r(\mathbf{f}_{ra} - \mathbf{f}_{rb}) \parallel^2 \end{aligned}$$

$$
\begin{aligned}
&+ \quad \frac{1}{2}\Big\{\alpha_c \parallel \mathbf{W}_{x1}(\mathbf{f}_{al} - \mathbf{f}_{ar}) \parallel^2 + \alpha_r \parallel \mathbf{W}_{x2}(\mathbf{f}_{ar} - \mathbf{f}_{br}) \parallel^2 \\
&\quad + \alpha_c \parallel \mathbf{W}_{x3}(\mathbf{f}_{br} - \mathbf{f}_{bl}) \parallel^2 + \alpha_r \parallel \mathbf{W}_{x4}(\mathbf{f}_{bl} - \mathbf{f}_{al}) \parallel^2\Big\} \\
&+ \quad \frac{1}{2}\beta\Big\{\parallel \mathbf{g}_{cl} - \mathbf{f}_{cl} \parallel^2 + \parallel \mathbf{g}_{cr} - \mathbf{f}_{cr} \parallel^2 + \parallel \mathbf{g}_{ra} - \mathbf{f}_{ra} \parallel^2 + \parallel \mathbf{g}_{rb} - \mathbf{f}_{rb} \parallel^2 \\
&+ \quad \parallel \mathbf{g}_{al} - \mathbf{f}_{al} \parallel^2 + \parallel \mathbf{g}_{ar} - \mathbf{f}_{ar} \parallel^2 + \parallel \mathbf{g}_{br} - \mathbf{f}_{br} \parallel^2 + \parallel \mathbf{g}_{bl} - \mathbf{f}_{bl} \parallel^2\Big\} \\
&= \quad \alpha_c \parallel \mathbf{W}_c(\mathbf{f}_{cl}^{(\alpha_c,\alpha_r,\beta)} - \mathbf{f}_{cr}^{(\alpha_c,\alpha_r,\beta)}) \parallel^2 + \alpha_r \parallel \mathbf{W}_r(\mathbf{f}_{ra}^{(\alpha_c,\alpha_r,\beta)} - \mathbf{f}_{rb}^{(\alpha_c,\alpha_r,\beta)}) \parallel^2 \\
&+ \quad \frac{1}{2}\Big\{\alpha_c \parallel \mathbf{W}_{x1}(\mathbf{f}_{al}^{(\alpha_c,\alpha_r,\beta)} - \mathbf{f}_{ar}^{(\alpha_c,\alpha_r,\beta)}) \parallel^2 + \alpha_r \parallel \mathbf{W}_{x2}(\mathbf{f}_{ar}^{(\alpha_c,\alpha_r,\beta)} - \mathbf{f}_{br}^{(\alpha_c,\alpha_r,\beta)}) \parallel^2 \\
&\quad + \alpha_c \parallel \mathbf{W}_{x3}(\mathbf{f}_{br}^{(\alpha_c,\alpha_r,\beta)} - \mathbf{f}_{bl}^{(\alpha_c,\alpha_r,\beta)}) \parallel^2 + \alpha_r \parallel \mathbf{W}_{x4}(\mathbf{f}_{bl}^{(\alpha_c,\alpha_r,\beta)} - \mathbf{f}_{al}^{(\alpha_c,\alpha_r,\beta)}) \parallel^2\Big\} \\
&+ \quad \frac{1}{2}\beta\Big\{\parallel \mathbf{g}_{cl} - \mathbf{f}_{cl}^{(\alpha_c,\alpha_r,\beta)} \parallel^2 + \parallel \mathbf{g}_{cr} - \mathbf{f}_{cr}^{(\alpha_c,\alpha_r,\beta)} \parallel^2 + \parallel \mathbf{g}_{ra} - \mathbf{f}_{ra}^{(\alpha_c,\alpha_r,\beta)} \parallel^2 + \parallel \mathbf{g}_{rb} - \mathbf{f}_{rb}^{(\alpha_c,\alpha_r,\beta)} \parallel^2 \\
&+ \quad \parallel \mathbf{g}_{al} - \mathbf{f}_{al}^{(\alpha_c,\alpha_r,\beta)} \parallel^2 + \parallel \mathbf{g}_{ar} - \mathbf{f}_{ar}^{(\alpha_c,\alpha_r,\beta)} \parallel^2 + \parallel \mathbf{g}_{br} - \mathbf{f}_{br}^{(\alpha_c,\alpha_r,\beta)} \parallel^2 + \parallel \mathbf{g}_{bl} - \mathbf{f}_{bl}^{(\alpha_c,\alpha_r,\beta)} \parallel^2\Big\} \\
&+ \quad \frac{1}{2}(\mathbf{f}_c - \mathbf{f}_c^{(\alpha_c,\alpha_r,\beta)})^t \mathbf{Q}_c(\alpha_c,\beta)(\mathbf{f}_c - \mathbf{f}_c^{(\alpha_c,\alpha_r,\beta)}) + \frac{1}{2}(\mathbf{f}_r - \mathbf{f}_r^{(\alpha_c,\alpha_r,\beta)})^t \mathbf{Q}_r(\alpha_r,\beta)(\mathbf{f}_r - \mathbf{f}_r^{(\alpha_c,\alpha_r,\beta)}) \\
&+ \quad \frac{1}{2}(\mathbf{f}_x - \mathbf{f}_x^{(\alpha_c,\alpha_r,\beta)})^t \mathbf{Q}_x(\alpha_c,\alpha_r,\beta)(\mathbf{f}_x - \mathbf{f}_x^{(\alpha_c,\alpha_r,\beta)}) \\
&= \quad M(\mathbf{f}^{(\alpha_c,\alpha_r,\beta)}, \mathbf{g} \mid \alpha_c, \alpha_r, \beta) \\
&+ \quad \frac{1}{2}(\mathbf{f}_c - \mathbf{f}_c^{(\alpha_c,\alpha_r,\beta)})^t \mathbf{Q}_c(\alpha_c,\beta)(\mathbf{f}_c - \mathbf{f}_c^{(\alpha_c,\alpha_r,\beta)}) + \frac{1}{2}(\mathbf{f}_r - \mathbf{f}_r^{(\alpha_c,\alpha_r,\beta)})^t \mathbf{Q}_r(\alpha_r,\beta)(\mathbf{f}_r - \mathbf{f}_r^{(\alpha_c,\alpha_r,\beta)}) \\
&+ \quad \frac{1}{2}(\mathbf{f}_x - \mathbf{f}_x^{(\alpha_c,\alpha_r,\beta)})^t \mathbf{Q}_x(\alpha_c,\alpha_r,\beta)(\mathbf{f}_x - \mathbf{f}_x^{(\alpha_c,\alpha_r,\beta)}),
\end{aligned}
$$

where $\mathbf{Q}_c(\alpha_c,\beta)$ is a block diagonal matrix whose diagonal 2×2 matrices are $\mathbf{q}_c(\alpha_c,\beta)(i) = \beta\mathbf{I}_{2\times 2} + \alpha_c\mathbf{A}_c(i)$ with

$$
\mathbf{A}_c(i) = \begin{bmatrix} 2\omega_c^2(i) & -2\omega_c^2(i) \\ -2\omega_c^2(i) & 2\omega_c^2(i) \end{bmatrix},
$$

$i = 1, 2, \ldots, p$, $\mathbf{Q}_r(\alpha_r,\beta)$ is a block diagonal matrix whose diagonal 2×2 matrices are $\mathbf{q}_r(i)(\alpha_r,\beta) = \beta\mathbf{I}_{2\times 2} + \alpha_r\mathbf{A}_r(i)$ with

$$
\mathbf{A}_r(i) = \begin{bmatrix} 2\omega_r^2(i) & -2\omega_r^2(i) \\ -2\omega_r^2(i) & 2\omega_r^2(i) \end{bmatrix},
$$

$i = 1, 2, \ldots, q$, and $\mathbf{Q}_x(\alpha_c,\alpha_r,\beta)$ is a block diagonal matrix whose diagonal 4×4 matrices are $\mathbf{q}_x(i)(\alpha_c,\alpha_r,\beta) = \beta\mathbf{I}_{4\times 4} + \alpha_c\mathbf{A}_{xc}(i) + \alpha_r\mathbf{A}_{xr}(i)$ with

$$
\mathbf{A}_{xc}(i) = \begin{bmatrix} \omega_{x1}^2(i) & -\omega_{x1}^2(i) & 0 & 0 \\ -\omega_{x1}^2(i) & \omega_{x1}^2(i) & 0 & 0 \\ 0 & -\omega_{x2}^2(i) & \omega_{x2}^2(i) & 0 \\ 0 & 0 & -\omega_{x3}^2(i) & \omega_{x3}^2(i) \end{bmatrix},
$$

and

$$\mathbf{A}_{xr}(i) = \begin{bmatrix} \omega_{x4}^2(i) & 0 & 0 & -\omega_{x4}^2(i) \\ 0 & \omega_{x2}^2(i) & -\omega_{x2}^2(i) & 0 \\ 0 & 0 & \omega_{x3}^2(i) & -\omega_{x3}^2(i) \\ -\omega_{x4}^2(i) & 0 & 0 & \omega_{x4}^2(i) \end{bmatrix},$$

$i = 1, 2, \ldots, m$.

We then have

$$\begin{aligned} & p(\mathbf{g} \mid \alpha_c, \alpha_r, \beta) \propto \\ & \quad \alpha_c^{\frac{p}{2}} \beta^p \alpha_r^{\frac{q}{2}} \beta^q F(\alpha_c, \alpha_r) \beta^{2m} \exp\{-M(\mathbf{f}^{(\alpha_c,\alpha_r,\beta)}, \mathbf{g} \mid \alpha_c, \alpha_r, \beta)\} \\ \times & \int_{\mathbf{f}_c} \exp\left\{-\frac{1}{2}(\mathbf{f}_c - \mathbf{f}_c^{(\alpha_c,\alpha_r,\beta)})^t \mathbf{Q}_c(\alpha_c, \beta)(\mathbf{f}_c - \mathbf{f}_c^{(\alpha_c,\alpha_r,\beta)})\right\} d\mathbf{f}_c \\ \times & \int_{\mathbf{f}_r} \exp\left\{-\frac{1}{2}(\mathbf{f}_r - \mathbf{f}_r^{(\alpha_c,\alpha_r,\beta)})^t \mathbf{Q}_r(\alpha_r, \beta)(\mathbf{f}_r - \mathbf{f}_r^{(\alpha_c,\alpha_r,\beta)})\right\} d\mathbf{f}_r \\ \times & \int_{\mathbf{f}_x} \exp\left\{-\frac{1}{2}(\mathbf{f}_x - \mathbf{f}_x^{(\alpha_c,\alpha_r,\beta)})^t \mathbf{Q}_x(\alpha_c, \alpha_r, \beta)(\mathbf{f}_x - \mathbf{f}_x^{(\alpha_c,\alpha_r,\beta)})\right\} d\mathbf{f}_r \\ = & \alpha_c^{\frac{p}{2}} \beta^p \alpha_r^{\frac{q}{2}} \beta^q F(\alpha_c, \alpha_r) \beta^{2m} \exp\{-M(\mathbf{f}^{(\alpha_c,\alpha_r,\beta)}, \mathbf{g} \mid \alpha_c, \alpha_r, \beta)\} \\ \times & [det\, \mathbf{Q}_c(\alpha_c, \beta)]^{-\frac{1}{2}} [det\, \mathbf{Q}_r(\alpha_r, \beta)]^{-\frac{1}{2}} [det\, \mathbf{Q}_x(\alpha_c, \alpha_r, \beta)]^{-\frac{1}{2}}. \end{aligned}$$

Differentiating $-\log[p(\mathbf{g} \mid \alpha_c, \alpha_r, \beta)]$ with respect to α_c, α_r and β we obtain

$$\begin{aligned} \frac{1}{\hat{\alpha}_c} = & \left\{2 \parallel \mathbf{W}_c(\mathbf{f}_{cl}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})} - \mathbf{f}_{cr}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})}) \parallel^2 + \parallel \mathbf{W}_{x1}(\mathbf{f}_{al}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})} - \mathbf{f}_{ar}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})}) \parallel^2 \right. \\ & + \parallel \mathbf{W}_{x3}(\mathbf{f}_{br}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})} - \mathbf{f}_{bl}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})}) \parallel^2 + \sum_{i=1}^{p} tr[\mathbf{q}_c^{-1}(\hat{\alpha}_c, \hat{\beta})(i)\mathbf{A}_c(i)] \\ & \left. + \sum_{i=1}^{m} tr[\mathbf{q}_x^{-1}(\hat{\alpha}_c, \hat{\alpha}_r, \hat{\beta})(i)\mathbf{A}_{xc}(i)]\right\} / \left[p + 2m + 2\sum_{i=1}^{m} \frac{\hat{\alpha}_c}{\hat{\alpha}_c + K(i)\hat{\alpha}_r}\right], \end{aligned} \tag{A-2}$$

$$\begin{aligned} \frac{1}{\hat{\alpha}_r} = & \left\{2 \parallel \mathbf{W}_r(\mathbf{f}_{ra}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})} - \mathbf{f}_{rb}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})}) \parallel^2 + \parallel \mathbf{W}_{x2}(\mathbf{f}_{ar}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})} - \mathbf{f}_{br}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})}) \parallel^2 \right. \\ & + \parallel \mathbf{W}_{x4}(\mathbf{f}_{bl}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})} - \mathbf{f}_{al}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})}) \parallel^2 + \sum_{i=1}^{q} tr[\mathbf{q}_r^{-1}(\hat{\alpha}_r, \hat{\beta})(i)\mathbf{A}_r(i)] \\ & \left. + \sum_{i=1}^{m} tr[\mathbf{q}_x^{-1}(\hat{\alpha}_c, \hat{\alpha}_r, \hat{\beta})(i)\mathbf{A}_{xr}(i)]\right\} / \left[q + 2m + 2\sum_{i=1}^{m} \frac{K(i)\hat{\alpha}_r}{\hat{\alpha}_c + K(i)\hat{\alpha}_r}\right], \end{aligned} \tag{A-3}$$

$$\begin{aligned} \frac{1}{\hat{\beta}} = & \left\{\parallel \mathbf{g}_{cl} - \mathbf{f}_{cl}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})} \parallel^2 + \parallel \mathbf{g}_{cr} - \mathbf{f}_{cr}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})} \parallel^2 + \parallel \mathbf{g}_{ra} - \mathbf{f}_{ra}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})} \parallel^2 \right. \\ & + \parallel \mathbf{g}_{rb} - \mathbf{f}_{rb}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})} \parallel^2 + \parallel \mathbf{g}_{al} - \mathbf{f}_{al}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})} \parallel^2 + \parallel \mathbf{g}_{ar} - \mathbf{f}_{ar}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})} \parallel^2 \\ & + \parallel \mathbf{g}_{br} - \mathbf{f}_{br}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})} \parallel^2 + \parallel \mathbf{g}_{bl} - \mathbf{f}_{bl}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})} \parallel^2 + \sum_{i=1}^{p} tr[\mathbf{q}_c^{-1}(\hat{\alpha}_c, \hat{\beta})(i)] \\ & \left. + \sum_{i=1}^{q} tr[\mathbf{q}_r^{-1}(\hat{\alpha}_r, \hat{\beta})(i)] + \sum_{i=1}^{m} tr[\mathbf{q}_x^{-1}(\hat{\alpha}_c, \hat{\alpha}_r, \hat{\beta})(i)]\right\} / [2(p + q + 2m)]. \end{aligned} \tag{A-4}$$

where

$$K(i) = \frac{\omega_{x2}^2(i)\omega_{x4}^2(i)[\omega_{x1}^2(i) + \omega_{x3}^2(i)]}{\omega_{x1}^2(i)\omega_{x3}^2(i)[\omega_{x2}^2(i) + \omega_{x4}^2(i)]}.$$

Obviously, Eqs. (A-2)–(A-4) characterize $\hat{\alpha}_c$, $\hat{\alpha}_r$ and $\hat{\beta}$ for the flat hyperpriors but they do not provide a way to estimate them. However, we notice that the substitution of the old values of α_c, α_r and β on the right hand side of Eqs. (A-2)–(A-4) to obtain the new values on the left hand side, is a gradient descent method to decrease $-\log p(\alpha_c, \alpha_r, \beta|\mathbf{g})$.

B. Parameter estimation step in Algorithm 2.

As described in section IV-A, the inclusion of $p(\mathbf{g}_x \mid \alpha_c, \alpha_r, \beta)$ in the estimation process of α_c, α_r and β requires the inversion of $(M/k-1)(N/k-1)$ 4×4 matrices. So, as it has been mentioned in Eq. (26), we propose to perform the estimation of α_c, α_r and β as

$$\hat{\alpha}_c, \hat{\alpha}_r, \hat{\beta} = \arg \max_{\alpha_c, \alpha_r, \beta} p(\mathbf{g}_c \mid \alpha_c, \beta) p(\mathbf{g}_r \mid \alpha_r, \beta). \tag{A-5}$$

Let us proceed to calculate $p(\mathbf{g}_c \mid \alpha_c, \beta)p(\mathbf{g}_r \mid \alpha_r, \beta)$ in Eq. (A-5), taking into account that

$$p(\mathbf{g}_c \mid \alpha_c, \beta)p(\mathbf{g}_r \mid \alpha_r, \beta) = \alpha_c^{\frac{p}{2}} \beta^p \alpha_r^{\frac{q}{2}} \beta^q \int_{\mathbf{f}_c, \mathbf{f}_r} \exp[-M_c(\mathbf{f}_c, \mathbf{g}_c \mid \alpha_c, \beta) + M_r(\mathbf{f}_r, \mathbf{g}_r \mid \alpha_r, \beta)]\, d\mathbf{f}_c\, d\mathbf{f}_r.$$

Let us fix α_c,α_r and β and expand $M_c(\mathbf{f}_c, \mathbf{g}_c \mid \alpha_c, \beta) + M_r(\mathbf{f}_r, \mathbf{g}_r \mid \alpha_r, \beta)$ around $\mathbf{f}_c^{(\alpha_c, \alpha_r, \beta)}$ and $\mathbf{f}_r^{(\alpha_c, \alpha_r, \beta)}$. We then have

$$\begin{aligned} p(\mathbf{g}_c \mid \alpha_c, \beta)p(\mathbf{g}_r \mid \alpha_r, \beta) \propto & \\ & \alpha_c^{\frac{p}{2}} \beta^p \alpha_r^{\frac{q}{2}} \beta^q \exp\{-M_c(\mathbf{f}_c^{(\alpha_c, \alpha_r, \beta)}, \mathbf{g}_c \mid \alpha_c, \beta)\} \exp\{-M_r(\mathbf{f}_r^{(\alpha_c, \alpha_r, \beta)}, \mathbf{g}_r \mid \alpha_r, \beta)\} \\ \times \quad & [det\, \mathbf{Q}_c(\alpha_c, \beta)]^{-\frac{1}{2}} [det\, \mathbf{Q}_r(\alpha_r, \beta)]^{-\frac{1}{2}}. \end{aligned}$$

Differentiating $-\log[p(\mathbf{g}_c \mid \alpha_c, \beta)p(\mathbf{g}_r \mid \alpha_r, \beta)]$ with respect to α_c, α_r and β we obtain $\hat{\alpha}_c$, $\hat{\alpha}_r$ and $\hat{\beta}$, that is,

$$\frac{p}{\hat{\alpha}_c} = 2 \parallel \mathbf{W}_c(\mathbf{f}_{cl}^{(\hat{\alpha}_c, \hat{\alpha}_r, \hat{\beta})} - \mathbf{f}_{cr}^{(\hat{\alpha}_c, \hat{\alpha}_r, \hat{\beta})}) \parallel^2 + \sum_{i=1}^{p} tr[\mathbf{q}_c^{-1}(\hat{\alpha}_c, \hat{\beta})(i)\mathbf{A}_c(i)], \tag{A-6}$$

$$\frac{q}{\hat{\alpha}_r} = 2 \parallel \mathbf{W}_r(\mathbf{f}_{ra}^{(\hat{\alpha}_c, \hat{\alpha}_r, \hat{\beta})} - \mathbf{f}_{rb}^{(\hat{\alpha}_c, \hat{\alpha}_r, \hat{\beta})}) \parallel^2 + \sum_{i=1}^{q} tr[\mathbf{q}_r^{-1}(\hat{\alpha}_r, \hat{\beta})(i)\mathbf{A}_r(i)], \tag{A-7}$$

$$\begin{aligned} \frac{2(p+q)}{\hat{\beta}} = & \parallel \mathbf{g}_{cl} - \mathbf{f}_{cl}^{(\hat{\alpha}_c, \hat{\alpha}_r, \hat{\beta})} \parallel^2 + \parallel \mathbf{g}_{cr} - \mathbf{f}_{cr}^{(\hat{\alpha}_c, \hat{\alpha}_r, \hat{\beta})} \parallel^2 + \parallel \mathbf{g}_{ra} - \mathbf{f}_{ra}^{(\hat{\alpha}_c, \hat{\alpha}_r, \hat{\beta})} \parallel^2 \\ & + \parallel \mathbf{g}_{rb} - \mathbf{f}_{rb}^{(\hat{\alpha}_c, \hat{\alpha}_r, \hat{\beta})} \parallel^2 + \sum_{i=1}^{p} tr[\mathbf{q}_c^{-1}(\hat{\alpha}_c, \hat{\beta})(i)] + \sum_{i=1}^{q} tr[\mathbf{q}_r^{-1}(\hat{\alpha}_r, \hat{\beta})(i)]. \end{aligned} \tag{A-8}$$

Since these equations do not provide a way to estimate the values of the hyperparameter, Algorithm 2 is used to estimate them.

C. Reconstruction step

Let us examine the reconstruction step. Given α_c, α_r and β, from Eq. (19), $\mathbf{f}_c^{(\alpha_c,\alpha_r,\beta)}$ is calculated by differentiating Eq. (22) with respect to $\mathbf{f}_c$, obtaining

$$2\alpha_c \mathbf{W}_c^t \mathbf{W}_c (\mathbf{f}_{cl}^{(\alpha_c,\alpha_r,\beta)} - \mathbf{f}_{cr}^{(\alpha_c,\alpha_r,\beta)}) - \beta(\mathbf{g}_{cl} - \mathbf{f}_{cl}^{(\alpha_c,\alpha_r,\beta)}) = 0, \quad \text{(A-9)}$$

$$2\alpha_c \mathbf{W}_c^t \mathbf{W}_c (\mathbf{f}_{cr}^{(\alpha_c,\alpha_r,\beta)} - \mathbf{f}_{cl}^{(\alpha_c,\alpha_r,\beta)}) - \beta(\mathbf{g}_{cr} - \mathbf{f}_{cr}^{(\alpha_c,\alpha_r,\beta)}) = 0, \quad \text{(A-10)}$$

which can be solved easily since we only have to invert 2×2 matrices. The solution is

$$\mathbf{f}_{cl}^{(\alpha_c,\alpha_r,\beta)}(i) = 0.5[1 + \beta(\beta + 4\alpha_c\omega_c^2(i))^{-1}]\mathbf{g}_{cl}(i) + 0.5[1 - \beta(\beta + 4\alpha_c\omega_c^2(i))^{-1}]\mathbf{g}_{cr}(i), \quad \text{(A-11)}$$

$$\mathbf{f}_{cr}^{(\alpha_c,\alpha_r,\beta)}(i) = 0.5[1 - \beta(\beta + 4\alpha_c\omega_c^2(i))^{-1}]\mathbf{g}_{cl}(i) + 0.5[1 + \beta(\beta + 4\alpha_c\omega_c^2(i))^{-1}]\mathbf{g}_{cr}(i), \quad \text{(A-12)}$$

for $i = 1, 2, \ldots, p$. Furthermore, differentiating Eq. (23) with respect to $\mathbf{f}_r$, we obtain

$$\mathbf{f}_{ra}^{(\alpha_c,\alpha_r,\beta)}(i) = 0.5[1 + \beta(\beta + 4\alpha_r\omega_r^2(i))^{-1}]\mathbf{g}_{ra}(i) + 0.5[1 - \beta(\beta + 4\alpha_r\omega_r^2(i))^{-1}]\mathbf{g}_{rb}(i), \quad \text{(A-13)}$$

$$\mathbf{f}_{rb}^{(\alpha_c,\alpha_r,\beta)}(i) = 0.5[1 - \beta(\beta + 4\alpha_r\omega_r^2(i))^{-1}]\mathbf{g}_{ra}(i) + 0.5[1 + \beta(\beta + 4\alpha_r\omega_r^2(i))^{-1}]\mathbf{g}_{rb}(i), \quad \text{(A-14)}$$

for $i = 1, 2, \ldots, q$. Finally, $\mathbf{f}_x^{(\alpha_c,\alpha_r,\beta)}$ is obtained by differentiating Eq. (24) with respect to $\mathbf{f}_x$, obtaining

$$\alpha_c \mathbf{W}_{x1}^t \mathbf{W}_{x1} (\mathbf{f}_{al}^{(\alpha_c,\alpha_r,\beta)} - \mathbf{f}_{ar}^{(\alpha_c,\alpha_r,\beta)}) + \alpha_r \mathbf{W}_{x4}^t \mathbf{W}_{x4} (\mathbf{f}_{al}^{(\alpha_c,\alpha_r,\beta)} - \mathbf{f}_{bl}^{(\alpha_c,\alpha_r,\beta)}) - \beta(\mathbf{g}_{al} - \mathbf{f}_{al}^{(\alpha_c,\alpha_r,\beta)}) = 0, \quad \text{(A-15)}$$

$$\alpha_c \mathbf{W}_{x1}^t \mathbf{W}_{x1} (\mathbf{f}_{ar}^{(\alpha_c,\alpha_r,\beta)} - \mathbf{f}_{al}^{(\alpha_c,\alpha_r,\beta)}) + \alpha_r \mathbf{W}_{x2}^t \mathbf{W}_{x2} (\mathbf{f}_{ar}^{(\alpha_c,\alpha_r,\beta)} - \mathbf{f}_{br}^{(\alpha_c,\alpha_r,\beta)}) - \beta(\mathbf{g}_{ar} - \mathbf{f}_{ar}^{(\alpha_c,\alpha_r,\beta)}) = 0, \quad \text{(A-16)}$$

$$\alpha_c \mathbf{W}_{x3}^t \mathbf{W}_{x3} (\mathbf{f}_{br}^{(\alpha_c,\alpha_r,\beta)} - \mathbf{f}_{bl}^{(\alpha_c,\alpha_r,\beta)}) + \alpha_r \mathbf{W}_{x2}^t \mathbf{W}_{x2} (\mathbf{f}_{br}^{(\alpha_c,\alpha_r,\beta)} - \mathbf{f}_{ar}^{(\alpha_c,\alpha_r,\beta)}) - \beta(\mathbf{g}_{br} - \mathbf{f}_{br}^{(\alpha_c,\alpha_r,\beta)}) = 0, \quad \text{(A-17)}$$

$$\alpha_c \mathbf{W}_{x3}^t \mathbf{W}_{x3} (\mathbf{f}_{bl}^{(\alpha_c,\alpha_r,\beta)} - \mathbf{f}_{br}^{(\alpha_c,\alpha_r,\beta)}) + \alpha_r \mathbf{W}_{x4}^t \mathbf{W}_{x4} (\mathbf{f}_{bl}^{(\alpha_c,\alpha_r,\beta)} - \mathbf{f}_{al}^{(\alpha_c,\alpha_r,\beta)}) - \beta(\mathbf{g}_{bl} - \mathbf{f}_{bl}^{(\alpha_c,\alpha_r,\beta)}) = 0, \quad \text{(A-18)}$$

which can be solved easily since we only have to invert 4×4 matrices to find $(\mathbf{f}_{al}(i), \mathbf{f}_{ar}(i), \mathbf{f}_{br}(i), \mathbf{f}_{bl}(i))$, $i = 1, 2, \ldots, m$.

II. Derivation of the Evidence Analysis for Gamma Hyperpriors

A. Parameter estimation step

In this section we use the gamma distributions, defined in Eqs. (29), (30) and (31), as hyperpriors for the hyperparameters, instead of the flat distributions.

First we take into account that

$$\begin{aligned} p(\alpha_c, \alpha_r, \beta \mid \mathbf{g}) &\propto p(\alpha_c)p(\alpha_r)p(\beta) \int_{\mathbf{f}} p(\mathbf{f}, \mathbf{g} \mid \alpha_c, \alpha_r, \beta)\, d\mathbf{f} \\ &= p(\alpha_c)p(\alpha_r)p(\beta) \int_{\mathbf{f}_c} p(\mathbf{f}_c, \mathbf{g}_c \mid \alpha_c, \beta)\, d\mathbf{f}_c \int_{\mathbf{f}_r} p(\mathbf{f}_r, \mathbf{g}_r \mid \alpha_r, \beta)\, d\mathbf{f}_r \int_{\mathbf{f}_x} p(\mathbf{f}_x, \mathbf{g}_x \mid \alpha_c, \alpha_r, \beta)\, d\mathbf{f}_x \\ &= p(\alpha_c)p(\alpha_r)p(\beta)p(\mathbf{g}_c \mid \alpha_c, \beta)p(\mathbf{g}_r \mid \alpha_r, \beta)p(\mathbf{g}_x \mid \alpha_c, \alpha_r, \beta). \end{aligned}$$

Following the same approach as with the flat hyperprior approximation, we estimate α_c, α_r and β by

$$\hat{\alpha}_c, \hat{\alpha}_r, \hat{\beta} = \arg \max_{\alpha_c, \alpha_r, \beta} p(\alpha_c)p(\alpha_r)p(\beta)p(\mathbf{g}_c \mid \alpha_c, \beta)p(\mathbf{g}_r \mid \alpha_r, \beta). \tag{B-1}$$

Then,

$$\begin{aligned} p(\alpha_c)p(\alpha_r)p(\beta)p(\mathbf{g}_c \mid \alpha_c, \beta)p(\mathbf{g}_r \mid \alpha_r, \beta) &= \alpha_c^{\frac{p}{2}+l(m_c^{cod})-1} \alpha_r^{\frac{q}{2}+l(m_r^{cod})-1} \beta^{p+q+l(n^{cod})-1} \\ \times \int_{\mathbf{f}_c, \mathbf{f}_r} \exp \Bigg[& -M_c(\mathbf{f}_c, \mathbf{g}_c \mid \alpha_c, \beta) - M_r(\mathbf{f}_r, \mathbf{g}_r \mid \alpha_r, \beta) - \frac{l(m_c^{cod})\alpha_c}{m_c^{cod}} \\ & - \frac{l(m_r^{cod})\alpha_r}{m_r^{cod}} - \frac{l(n^{cod})\beta}{n^{cod}} \Bigg] d\mathbf{f}_c\, d\mathbf{f}_r. \end{aligned}$$

That is

$$\begin{aligned} & p(\alpha_c)p(\alpha_r)p(\beta)p(\mathbf{g}_c \mid \alpha_c, \beta)p(\mathbf{g}_r \mid \alpha_r, \beta) \propto \\ & \quad \alpha_c^{\frac{p}{2}+l(m_c^{cod})-1} \alpha_r^{\frac{q}{2}+l(m_r^{cod})-1} \beta^{p+q+l(n^{cod})-1} \exp\{-M_c(\mathbf{f}_c^{(\alpha_c,\alpha_r,\beta)}, \mathbf{g}_c \mid \alpha_c, \beta)\} \\ \times & \quad \exp\{-M_r(\mathbf{f}_r^{(\alpha_c,\alpha_r,\beta)}, \mathbf{g}_r \mid \alpha_r, \beta)\} \exp\left\{-\frac{l(m_c^{cod})\alpha_c}{m_c^{cod}}\right\} \exp\left\{-\frac{l(m_r^{cod})\alpha_r}{m_r^{cod}}\right\} \\ \times & \quad \exp\left\{-\frac{l(n^{cod})\beta}{n^{cod}}\right\} \int_{\mathbf{f}_c} \exp\left\{-\frac{1}{2}(\mathbf{f}_c - \mathbf{f}_c^{(\alpha_c,\alpha_r,\beta)})^t \mathbf{Q}_c(\alpha_c, \beta)(\mathbf{f}_c - \mathbf{f}_c^{(\alpha_c,\alpha_r,\beta)})\right\} d\mathbf{f}_c \\ \times & \quad \int_{\mathbf{f}_r} \exp\left\{-\frac{1}{2}(\mathbf{f}_r - \mathbf{f}_r^{(\alpha_c,\alpha_r,\beta)})^t \mathbf{Q}_r(\alpha_r, \beta)(\mathbf{f}_r - \mathbf{f}_r^{(\alpha_c,\alpha_r,\beta)})\right\} d\mathbf{f}_r \\ = & \quad \alpha_c^{\frac{p}{2}+l(m_c^{cod})-1} \alpha_r^{\frac{q}{2}+l(m_r^{cod})-1} \beta^{p+q+l(n^{cod})-1} \exp\{-M_c(\mathbf{f}_c^{(\alpha_c,\alpha_r,\beta)}, \mathbf{g}_c \mid \alpha_c, \beta)\} \\ & \quad \exp\{-M_r(\mathbf{f}_r^{(\alpha_c,\alpha_r,\beta)}, \mathbf{g}_r \mid \alpha_r, \beta)\} \exp\left\{-\frac{l(m_c^{cod})\alpha_c}{m_c^{cod}}\right\} \exp\left\{-\frac{l(m_r^{cod})\alpha_r}{m_r^{cod}}\right\} \\ & \quad \exp\left\{-\frac{l(n^{cod})\beta}{n^{cod}}\right\} [det\, \mathbf{Q}_c(\alpha_c, \beta)]^{-\frac{1}{2}} [det\, \mathbf{Q}_r(\alpha_r, \beta)]^{-\frac{1}{2}}. \end{aligned}$$

Differentiating $-\log[p(\alpha_c)p(\alpha_r)p(\beta)p(\mathbf{g}_c \mid \alpha_c, \beta)p(\mathbf{g}_r \mid \alpha_r, \beta)]$ with respect to α_c, α_r and β we obtain $\hat{\alpha}_c$, $\hat{\alpha}_r$ and $\hat{\beta}$, that is,

$$\begin{aligned} \left(\frac{p}{2} + l(m_c^{cod}) - 1\right) \frac{1}{\hat{\alpha}_c} &= \| \mathbf{W}_c(\mathbf{f}_{cl}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})} - \mathbf{f}_{cr}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})}) \|^2 + \frac{l(m_c^{cod})}{m_c^{cod}} \\ &+ \frac{1}{2} \sum_{i=1}^{p} tr[\mathbf{q}_c^{-1}(\hat{\alpha}_c, \hat{\beta})(i)\mathbf{A}_c(i)], \end{aligned} \tag{B-2}$$

$$\left(\frac{q}{2}+l(m_r^{cod})-1\right)\frac{1}{\hat{\alpha}_r} = \parallel \mathbf{W}_r(\mathbf{f}_{ra}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})}-\mathbf{f}_{rb}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})})\parallel^2+\frac{l(m_r^{cod})}{m_r^{cod}}$$
$$+ \frac{1}{2}\sum_{i=1}^{q} tr[\mathbf{q}_r^{-1}(\hat{\alpha}_r,\hat{\beta})(i)\mathbf{A}_r(i)], \quad \text{(B-3)}$$

$$\left(p+q+l(n^{cod})-1\right)\frac{1}{\hat{\beta}} = \frac{1}{2}\parallel \mathbf{g}_{cl}-\mathbf{f}_{cl}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})}\parallel^2+\frac{1}{2}\parallel \mathbf{g}_{cr}-\mathbf{f}_{cr}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})}\parallel^2$$
$$+ \frac{1}{2}\parallel \mathbf{g}_{ra}-\mathbf{f}_{ra}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})}\parallel^2+\frac{1}{2}\parallel \mathbf{g}_{rb}-\mathbf{f}_{rb}^{(\hat{\alpha}_c,\hat{\alpha}_r,\hat{\beta})}\parallel^2+\frac{l(n^{cod})}{n^{cod}}$$
$$+ \frac{1}{2}\sum_{i=1}^{p} tr[\mathbf{q}_c^{-1}(\hat{\alpha}_c,\hat{\beta})(i)]++\frac{1}{2}\sum_{i=1}^{q} tr[\mathbf{q}_r^{-1}(\hat{\alpha}_r,\hat{\beta})(i)]. \quad \text{(B-4)}$$

References

[1] J. Mateos, A. K. Katsaggelos, and R. Molina, "Parameter estimation in regularized reconstruction of BDCT compressed images for reducing blocking artifacts," in *Proceedings of the Conference on Digital Compression Technologies & Video Communications, SPIE Proc.*, 1996, vol. 2952, pp. 70–81.

[2] J. Mateos, R. Molina, and A. K. Katsaggelos, "Estimating and transmitting regularization parameters for reducing blocking artifacts," in *Proceedings of the 13th International Conference on Digital Signal Processing*, 1997, pp. 209–212.

[3] W. B. Pennebaker and J. L. Mitchell, *JPEG still image compression standard*, Van Nostrand Reinhold, 1992.

[4] H. C. Reeves and J. S. Lim, "Reduction of blocking effects in image coding," *Optical Engineering*, vol. 23, pp. 34–37, 1984.

[5] G. Ramamurthi and A. Gersho, "Nonlinear space-variant postprocessing of block coded images," *IEEE Trans. on Acoustic, Speech and Signal Processing*, vol. 34, pp. 1258–1269, 1986.

[6] A. Baskurt, R. Prost, and R. Goutte, "Iterative constrained restoration of DCT-compressed images," *Signal Processing*, vol. 17, pp. 201–211, 1989.

[7] K. Sauer, "Enhancement of low bit-rate coded images using edge detection and estimation," *CVGIP: Graphical Models and Image Processing*, vol. 53, pp. 52–62, 1991.

[8] A. Zakhor, "Iterative procedures for reduction of blocking effects in transform image coding," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 2, pp. 91–95, 1992.

[9] S. J. Reeves and S. L. Eddins, "Comments on "Iterative procedures for reduction of blocking effects in transform image coding"," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 3, pp. 439–440, 1993.

[10] Y. Yang, N. P. Galatsanos, and A. K. Katsaggelos, "Regularized reconstruction to reduce blocking artifacts of block discrete cosine transform compressed images," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 3, pp. 421–432, 1993.

[11] R. L. Stevenson, "Reduction of coding artifacts in transform image coding," in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing ICASSP 93*, 1993, vol. 5, pp. 401–404.

[12] T. P. O'Rourke and R. L. Stevenson, "Improved image decompression for reduced transform coding artifacts," in *Proceedings of the Image and Video Processing II Conf., SPIE Proc.*, 1994, vol. 2182, pp. 90–101.

[13] T. P. O'Rourke and R. L. Stevenson, "Improved image decompression for reduced transform coding artifacts," *IEEE Trans. on Circuits and Systems for Video Technology*, vol. 5, pp. 490–499, 1995.

[14] R. L. Stevenson, "Reduction of coding artifacts in low-bit-rate video coding," in *Proceedings of the 38th Midwest Symposium on Circuits and Systems.*, 1995, pp. 854–857.

[15] J. Luo, C. W. Chen, K. J. Parker, and T. S. Huang, "Artifact reduction in low bit rate DCT-based image compression," *IEEE Trans. on Image Processing*, vol. 5, pp. 1363–1368, 1996.

[16] T. Özcelik, J. C. Brailean, and A. K. Katsaggelos, "Image and video compression algorithms based on recovery techniques using mean field annealing," *Proceedings of the IEEE*, vol. 83, pp. 304–316, 1995.

[17] Y. Yang, N. P. Galatsanos, and A. K. Katsaggelos, "Projection-based spatially-adaptive reconstruction of block-transform compressed images," *IEEE Trans. on Image Processing*, vol. 4, pp. 896–908, 1995.

[18] H. Paek, J.-W. Park, and S.-U. Lee, "Non-iterative post-processing technique for transform coded image sequence," in *Proceedings of the International Conference on Image Processing ICIP 95*, 1995, vol. 3, pp. 208–211.

[19] K. Y. Kwak and R.A Haddad, "Projection-based eigenvector decomposition for reduction of blocking artifacts of DCT coded image," in *Proceedings of the International Conference on Image Processing ICIP 95*, 1995, vol. 2, pp. 527–530.

[20] R. Molina, A. K. Katsaggelos, and J. Mateos, "Bayesian and regularization methods for hyperparameter estimation in image restoration," *IEEE Trans. on Image Processing*, vol. 8, pp. 231–246, 1999.

[21] R. Molina, "On the hierarchical Bayesian approach to image restoration. application to astronomical images," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 16, pp. 1222–1228, 1994.

[22] D.A. Silverstein and S.A. Klein, "Restoration of compressed images," in *Proceedings of the Image and Video Compression Conf.*, 1994, vol. SPIE Proc. 2186, pp. 56–64.

[23] W. L. Buntine, *A Theory of Learning Classification Rules*, Ph.D. thesis, University of Technology, Sydney, Australia, 1991.

[24] D. J. Spiegelhalter and S. L. Lauritzen, "Sequential updating of conditional probabilities on directed graphical structures," *Networks*, vol. 20, pp. 579–605, 1990.

[25] W. L. Buntine, "Theory refinement on Bayesian networks," in *Proc. of the Seventh Conference on Uncertainty in Artificial Intelligence*, 1991, pp. 52–60.

[26] G. F. Cooper and E. Herkovsits, "A Bayesian method for the induction of probabilistic networks from data," *Machine Learning*, vol. 9, pp. 309–347, 1992.

[27] D. J. C. MacKay, "A practical Bayesian framework for backprop networks," *Neural Computation*, vol. 4, pp. 448–472, 1992.

[28] W. L. Buntine and A. Weigund, "Bayesian back-propagation," *Complex Systems*, vol. 5, pp. 603–644, 1991.

[29] S. F. Gull, "Developments in maximum entropy data analysis," in *Maximum Entropy and Bayesian Methods*, J. Skilling, Ed., pp. 53–71. Cambridge, Kluwer, 1989.

[30] D. J. C. MacKay, "Bayesian interpolation," *Neural Computation*, vol. 4, pp. 415–447, 1992.

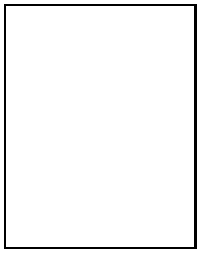
[31] J. O. Berger, *Statistical Decision Theory and Bayesian Analysis*, chapter 3 and 4, New York, Springer Verlag, 1985.

[32] R. Molina and A. K. Katsaggelos, "On the hierarchical Bayesian approach to image restoration and the iterative evaluation of the regularization parameter," in *Proc. of the Visual Communication and Image Processing'94*, 1994, pp. 244–251.

[33] R. M. Neal, *Bayesian Learning for Neural Networks*, Ph.D. thesis, University of Toronto, Department of Computer Science, 1995.

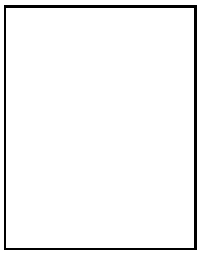
[34] K. T. Lay and A. K. Katsaggelos, "Image identification and restoration based on the expectation-maximization algorithm," *Optical Engineering*, vol. 29, pp. 436–445, 1990.

[35] S. S. O. Choy, Y.-H. Chan, and W.-C. Siu, "Reduction of block-transform image coding artifacts by using local statistics of transform coefficients," *IEEE Signal Processing Letters*, vol. 4, pp. 5–7, 1997.



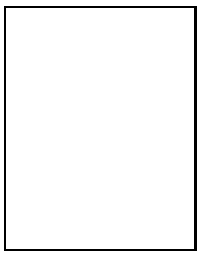
Javier Mateos was born in Granada, Spain, in 1968. He received his degree in Computer Science from the University of Granada in 1991. He completed his Ph. D. in Computer Science at the University of Granada in July 1998.

Since September 1992, he has been Assistant Professor at the Department of Computer Science and Artificial Intelligence of the University of Granada. His research interest include image restoration and image and video recovery and compression. He is a member of the AERFAI (Asociación Española de Reconocimento de Formas y Análisis de Imágenes).



Aggelos K. Katsaggelos received the Diploma degree in electrical and mechanical engineering from the Aristotelian University of Thessaloniki, Thessaloniki, Greece, in 1979 and the M.S. and Ph.D. degrees both in electrical engineering from the Georgia Institute of Technology, Atlanta, Georgia, in 1981 and 1985, respectively.

In 1985 he joined the Department of Electrical Engineering and Computer Science at Northwestern University, Evanston, IL, where he is currently professor, holding the Ameritech Chair of Information Technology. He is also the Director of the Motorola Center for Communications. During the 1986-1987 academic year he was an assistant professor at Polytechnic University, Department of Electrical Engineering and Computer Science, Brooklyn, NY. His current research interests include image and video recovery, video compression, motion estimation, boundary encoding, computational vision, and multimedia signal processing. Dr. Katsaggelos is a Fellow of the IEEE, an Ameritech Fellow, a member of the Associate Staff, Department of Medicine, at Evanston Hospital, and a member of SPIE. He is a member of the Board of Governors and the Publication Board of the IEEE Signal Processing Society, the IEEE TAB Magazine Committee, the Steering Committee of the *IEEE Transactions on Medical Imaging*, the IEEE Technical Committees on Visual Signal Processing and Communications, and Multimedia Signal Processing, and editor-in-chief of the *IEEE Signal Processing Magazine*. He has served as an Associate editor for the *IEEE Transactions on Signal Processing* (1990-1992), an area editor for the journal *Graphical Models and Image Processing* (1992-1995), a member of the Steering Committee of the *IEEE Transactions on Image Processing* (1992-1997), and a member of the IEEE Technical Committee on Image and Multi-Dimensional Signal Processing (1992-1998). He is the editor of *Digital Image Restoration* (Springer-Verlag, Heidelberg, 1991), co-author of *Rate-Distortion Based Video Compression* (Kluwer Academic Publishers, 1997), and co-editor of *Recovery Techniques for Image and Video Compression and Transmission,* (Kluwer Academic Publishers, 1998). He has served as the General Chairman of the 1994 Visual Communications and Image Processing Conference (Chicago, IL), and as technical program co-chair of the 1998 IEEE International Conference on Image Processing (Chicago, IL).



Rafael Molina was born in 1957 and received his degree in Mathematics (Statistics) in 1979. He completed his Ph. D. Thesis in 1983 in Optimal Design in Linear Models and became Associate Professor in Computer Science and Artificial Intelligence at the University of Granada in 1989. His areas of research interest are image restoration (applications to astronomy and medicine), parameter estimation, image compression, and blind deconvolution. He is a member of the IEEE, SPIE, Royal Statistical Society and AERFAI (Asociación Española de Reconocimento de Formas y Análisis de Imágenes) and currently the Dean of the Computer Engineering Faculty at the University of Granada.

Figure and Table Caption List

Fig. 1. Distribution of boundary pixels according to their position.

Fig. 2. Distribution of weights.

Fig. 3. Weights of the blocky *Lena* image: (a) vertical, (b) horizontal.

Fig. 4. (a) *Lena* compressed at 0.29 bpp. Reconstructions with (b) Algorithm 1, (c) Algorithm 2 + projection on the quantization intervals, (d) POCS and (e) WLS.

Fig. 5. (a) *peppers* image compressed at 0.32 bpp. Reconstructions with (b) Algorithm 1, (c) Algorithm 2 + projection on the quantization intervals, (d) POCS and (e) WLS.

Fig. 6. PSNR for different values of μ and ν on the *Lena* image compressed at 0.29 bpp.

Fig. 7. (a) *Lena* compressed at 0.29 bpp. (b) Reconstruction with $\alpha_c^{cod^{-1}} = 30.82$, $\alpha_r^{cod^{-1}} = 5.36$, $\beta^{cod^{-1}} = 36.36$, $\mu = 0.9$, $\nu = 0.0$. PSNR = 31.40 dB.

TABLE I

PSNR obtained by different reconstruction methods together with Algorithms 1 and 2.

TABLE II

PSNR obtained by estimating the parameters at the coder.

TABLE III

Best PSNR obtained by combining the parameters at the coder and the decoder.

TABLE IV

Best PSNR obtained by combining a 6 bits quantized versions of the parameters at the coder and the obtained at the decoder.

TABLE V

Best PSNR obtained by combining a 3 bits quantized versions of the parameters at the coder and the obtained at the decoder.