Estimating and Transmitting Regularization Parameters for Reducing Blocking Artifacts

Javier Mateos, Rafael Molina and Aggelos K. Katsaggelos

a) Departamento de Ciencias de la Computación e I.A., Universidad de Granada. 18071 Granada, Spain.

b) Department of Electrical and Computer Engineering, Northwestern University. 2145 Sheridan Road. Evanston, IL 60208-3118.

Abstract—High compression ratios for both still images and sequences of images are usually achieved by quantizing the block discrete cosine transform (BDCT) coefficients of the intensity or displaced frame diferent frames. This block based processing and quantization yields images that exhibit annoying blocking artifacts. In this paper, we propose a method based on the hierarchical Bayesian approach for the reconstruction of BDCT compressed images and the estimation of the related parameters both at the coder and decoder. We examine how to combine the parameters estimated at the coder and decoder and test the method on real images.

Key Words: Reconstruction, post-processing, image coding.

I. INTRODUCTION

Block-transform coding is by far the most popular approach for image compression. Evidence of this fact is that both the Joint Photographic Experts Group (JPEG) and the Motion Pictures Experts Group (MPEG) recommend the use of the block discrete cosine transform (BDCT) for compressing still and sequences of motion images, respectively [1]. However, such a compression method results in blocking artifacts for high compression ratios.

This artifact manifests itself as an artificial discontinuity 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. It constitutes a serious bottleneck for many important visual communication applications that require visually pleasing images at very high compression ratios. The recent progress in VLSI technology makes us believe that the incorporation of recovery algorithms at the decoders is a very promising approach to bridging the conflicting requirements of high-quality images and high compression ratios. In the past various algorithms have been proposed to improve quality of block-transform compressed images in the decoder without increasing the bit-rate. In the JPEG standard [1] a technique for predicting the AC coefficients is recommended in Annex-K 8.2. However, in areas with sharp intensity transitions such a prediction scheme fails. In [2], [3], methods to reconstruct the image using spatially-invariant prior knowledge about the original image along with the transmitted data are proposed. A MAP approach based on a compound Gauss Markov image model and mean field annealing is proposed in [4]. In [5], a POCS based recovery algorithm is proposed which uses a spatially-adaptive constraint that enforces between-block smoothness.

The method proposed in [5] requires the estimation of the regularization parameter. In this paper we formulate a similar algorithm within the hierarchical Bayesian paradigm and perform the estimation of the regularization parameters which represent the image and noise variances [6]. 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 within the hierarchical Bayesian approach to the reconstruction problem.

The rest of the paper is organized as follows. In section II the mathematical background for this paper is established. Section III describes the prior and noise models for the reconstruction problem. In section IV we describe the use of the hierarchical approach to the hyperparameter estimation and the recovery algorithm is presented for both the reconstructed and the original image. In section V we examine how the parameters obtained from the original and reconstructed images can be combined following the hierarchical Bayesian paradigm. Experimental results are presented in section VI and, finally, section VII concludes the paper.

II. MATHEMATICAL DEFINITION OF THE RECOVERY PROBLEM

Throughout this paper a digital N x N image is treated as a N² x 1 vector in the Rᴺ² space by lexicographic ordering either the rows or the columns. The BDCT is viewed as a linear transformation from Rᴺ² to Rᴺ². Then, for an image f we can write F = BF, where F is the BDCT of f and B is the BDCT matrix. To achieve a bit-rate reduction for transmission, each element of F is quantized. This quantization operator can be described mathematically by a mapping or an operator, Q from Rᴺ² to Rᴺ². The input-output relation of the coder can be modeled by G = QBF. Due to the unitary property of the DCT matrices, the BDCT matrix is also unitary and the inverse transform can be simply expressed by Bᵀ where t denotes the transpose of a matrix. In the receiver only the quantized BDCT coefficients G are available and the output of a conventional decoder is
\( g = B^r G \).

The reconstruction problem calls for finding an estimate of \( f \) given \( g, Q \) and, possibly, knowledge about \( f \). In this work we propose to reconstruct the artificial block boundaries, where the blocking artifact is more visible, keeping the rest of the image unchanged. Following the Bayesian paradigm a degradation model is needed which describes how the blocky image is obtained from the real one. In addition, a model describing our knowledge about the original image, the image before coding is also required. Let us study these models in detail.

III. IMAGE AND NOISE MODELS

Let us assume without loss of generality that \( 8 \times 8 \) blocks are used and the image is \( 512 \times 512 \). We define by \( \mathbf{x}_c \) and \( \mathbf{y}_c \) column vectors formed by the elements of \( \mathbf{f} \) and \( \mathbf{g} \), respectively, that are at the left of a block boundary column but not in a four-block intersection, that is, \( \mathbf{u}_c = \{ \mathbf{f}(u) \} \), \( \mathbf{y}_c = \{ \mathbf{g}(u) \} \), \( u = (x, y), x = 8 \times i, y = 8 \times j , i = 1, 2, \ldots, 64, j = 2, 3, \ldots, 7 \). We define by \( \mathbf{v}_c \) and \( \mathbf{y}_c \) column vectors formed by elements of \( \mathbf{f} \) and \( \mathbf{g} \), respectively, that are at the right of a block boundary column but not in a four-block intersection, that is, \( \mathbf{v}_c = \{ \mathbf{f}(u) \} \), \( \mathbf{y}_c = \{ \mathbf{g}(u) \} \), \( u = (x, y), x = 8 \times i + 1, y = 8 \times j + 1 , i = 1, 2, \ldots, 64, j = 2, 3, \ldots, 7 \).

In a similar way we define by \( \mathbf{u}_r, \mathbf{v}_r, \mathbf{x}_r \) and \( \mathbf{y}_r \), the column vectors representing the rows at the block boundaries of \( \mathbf{f} \) and \( \mathbf{g} \), respectively. We further define \( \mathbf{f}_u(a), \mathbf{f}_v(b), \mathbf{f}_u(c), \mathbf{f}_v(d) \) and \( \mathbf{g}_u(a), \mathbf{g}_v(b), \mathbf{g}_u(c), \mathbf{g}_v(d) \), \( u = (i, j) \) as the pixels in a four-block boundary of \( \mathbf{f} \) and \( \mathbf{g} \), respectively, as represented in Fig 1.

In order to capture the vertical local properties of the image we define a \( (512 \times 63) \times (512 \times 63) \) diagonal matrix, \( \mathbf{W}_c \), of the form

\[
\mathbf{W}_c = \begin{bmatrix}
\omega_c(1) & 0 & 0 & \cdots & 0 \\
0 & \omega_c(2) & 0 & \cdots & 0 \\
0 & 0 & \omega_c(3) & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & \omega_c(512 \times 63)
\end{bmatrix},
\]

where the \( \omega_c(i) \)'s, \( i = 1, 2, \ldots, (512 \times 63) \) will weight each block boundary pixel. Analogously, we can define \( \mathbf{W}_r \) to capture the horizontal local properties of the image (see [5], [7] for details).

With the above definitions, the prior knowledge about the smoothness in the block boundaries of the image has the form

\[
p(f | \alpha) \propto \alpha^{-\frac{p}{2}} \exp \left\{ -\frac{1}{2} \mathbf{W}_c (\mathbf{u}_c - \mathbf{v}_c)^2 \right\},
\]

with \( p = 961 \times 63 \) and \( \alpha \) measure of the roughness between two block boundaries.

Fidelity to the data at these block boundaries is expressed by

\[
p(g | f, \beta) \propto \beta^{-r} \exp \left\{ \frac{1}{2} (g - \mathbf{f}) \mathbf{W}_r (g - \mathbf{f})^T \right\}
\]

where \( N(g | f, \beta) \) is defined by

\[
\beta \left\{ \| \mathbf{u}_c - \mathbf{x}_r \|^2 + \| \mathbf{y}_c - \mathbf{y}_r \|^2 + \| \mathbf{u}_r - \mathbf{x}_c \|^2 + \| \mathbf{v}_c - \mathbf{y}_c \|^2 + \| \mathbf{v}_r - \mathbf{y}_r \|^2 + \| \mathbf{f}(a) - \mathbf{g}(a) \|^2 + \| \mathbf{f}(b) - \mathbf{g}(b) \|^2 + \| \mathbf{f}(c) - \mathbf{g}(c) \|^2 + \| \mathbf{f}(d) - \mathbf{g}(d) \|^2 \right\},
\]

\( \beta \) is defined as \( \beta^{-1} = \sigma_{\text{noise}}^2 \), the noise variance in the image, and \( r = 898 \times 63 \).

If \( \alpha \) and \( \beta \) are known, then following the Bayesian paradigm it is customary to select, as the reconstruction of \( f \), the image \( \mathbf{f}_{(a,b)} \) defined by

\[
\mathbf{f}_{(a,b)} = \arg \min_f \{ M(f, g | \alpha, \beta) \}
\]

\[
= \arg \min_f \{ P(f | \alpha) + N(g | f, \beta) \}.
\]

IV. PROPOSED ALGORITHM

Given \( \alpha \) and \( \beta \), \( \mathbf{f}_{(a,b)} \) is estimated by differentiating Eq. (1) with respect to \( f \), obtaining

\[
\mathbf{u}_c(i) = 0.5[1 + \beta(1 + 4\alpha \omega_c^2(i)-1)] \mathbf{x}_c(i) + 0.5[1 - \beta(1 + 4\alpha \omega_c^2(i)-1)] \mathbf{y}_c(i) \tag{2}
\]

\[
\mathbf{v}_c(i) = 0.5[1 + \beta(1 + 4\alpha \omega_c^2(i)-1)] \mathbf{x}_c(i) + 0.5[1 + \beta(1 + 4\alpha \omega_c^2(i)-1)] \mathbf{y}_c(i) \tag{3}
\]

\[
\mathbf{u}_r(i) = 0.5[1 + \beta(1 + 4\alpha \omega_c^2(i)-1)] \mathbf{x}_r(i) + 0.5[1 - \beta(1 + 4\alpha \omega_c^2(i)-1)] \mathbf{y}_r(i) \tag{4}
\]

\[
\mathbf{v}_r(i) = 0.5[1 - \beta(1 + 4\alpha \omega_c^2(i)-1)] \mathbf{x}_r(i) + 0.5[1 + \beta(1 + 4\alpha \omega_c^2(i)-1)] \mathbf{y}_r(i) \tag{5}
\]

and, for pixels in a four-block boundary, the solution of the following system of equations

\[
\alpha \omega_c^2(a)(\mathbf{f}(a) - \mathbf{f}(b)) + \omega_b^2(a)(\mathbf{f}(a) - \mathbf{f}(d)) + \beta(\mathbf{f}(a) - \mathbf{g}(a)) = 0
\]

\[
\alpha \omega_c^2(a)(\mathbf{f}(b) - \mathbf{f}(a)) + \omega_b^2(c)(\mathbf{f}(b) - \mathbf{f}(c)) + \beta(\mathbf{f}(b) - \mathbf{g}(b)) = 0
\]

\[
\alpha \omega_c^2(c)(\mathbf{f}(c) - \mathbf{f}(d)) + \omega_b^2(c)(\mathbf{f}(c) - \mathbf{f}(b)) + \beta(\mathbf{f}(c) - \mathbf{g}(c)) = 0
\]

\[
\alpha \omega_c^2(c)(\mathbf{f}(d) - \mathbf{f}(c)) + \omega_b^2(a)(\mathbf{f}(d) - \mathbf{f}(a)) + \beta(\mathbf{f}(d) - \mathbf{g}(d)) = 0.
\]
It is obvious that in order to estimate the original image, that is, in order to solve Eq. (1) we need to know or estimate the unknown hyperparameters. Yang et al. [3, 5] proposed empirical procedures to estimate these parameters within the POCS and constrained least squares approaches to the reconstruction problem. In this paper we follow the so called hierarchical Bayesian paradigm, according to which prior knowledge about the unknown hyperparameters can be incorporated into the algorithm if available or they can be estimated without any prior knowledge.

This hierarchical paradigm applied to our problem uses the global probability defined as

$$p(\alpha, \beta, f, g) = p(\alpha)p(\beta)p(f | \alpha)p(g | f, \beta).$$

Let us examine all the term involved in this equation. Having defined in the previous section the image and degradation models, that is, $p(f | \alpha)$ and $p(g | f, \beta)$, we have to specify the probabilities on the hyperparameters.

In situations without prior information about the hyperparameters $\alpha$ and $\beta$ what is needed is a non informative prior (the term non informative is meant to imply that no information about the hyperparameters is contained in the priors). For the problem at hand we can use improper non informative priors

$$p(\omega) \propto \text{const over } [0, \infty)$$

where $\omega$ denotes a hyperparameter. However, it is also possible to incorporate precise prior knowledge about the value of the noise and prior variances. With the improper hyperpriors, the method proposed amounts to selecting $\alpha$ and $\beta$, as the maximum likelihood estimates, mle, of $\alpha$ and $\beta$ from $p(g | \alpha, \beta)$.

Let us describe the estimation process in detail. Let us fix $\alpha$ and $\beta$ and expand $M(f, g | \alpha, \beta)$ around $f_{(\alpha, \beta)}$. We then have,

$$p(\alpha, \beta | g) \propto p(\alpha, \beta | g) \propto \exp\left\{\frac{1}{2}M(f_{(\alpha, \beta)}, g | \alpha, \beta)\right\} \times$$

$$\frac{1}{\alpha^{\frac{d}{2}}\beta^{\frac{d}{2}}} \int_{f} \exp\left\{ -\frac{1}{2}M(f - f_{(\alpha, \beta)}, g | \alpha, \beta) \right\} df$$

where $Q(\alpha, \beta)$ is a matrix which can be represented by a diagonal matrix with elements, $q_{ii}(\alpha, \beta) = 2\beta I_{2 \times 2} + \alpha A_{i}$ where $A_{i}$ is defined as

$$\begin{bmatrix} 4\omega^{2}(i) & -4\omega^{2}(i) \\ -4\omega^{2}(i) & 4\omega^{2}(i) \end{bmatrix}$$

and depending on pixel location $i \in u_{c}$ or $i \in u_{r}$, respectively. For pixel locations $i$ in a four-block intersection, $q_{ii}(\alpha, \beta) = 2\beta I_{4 \times 4} + 2\alpha A_{i}$ where $A_{i}$ is defined as

$$\begin{bmatrix} A_{2 \times 2}(a, c) & 0 & -2\omega^{2}(a) \\ 0 & -2\omega^{2}(c) & 0 \\ -2\omega^{2}(a) & 0 & A_{2 \times 2}(c, a) \end{bmatrix}$$

with

$$A_{2 \times 2}(x, y) = \begin{bmatrix} \omega^{2}(x) + \omega^{2}(y) & -2\omega^{2}(x) \\ -2\omega^{2}(y) & \omega^{2}(x) + \omega^{2}(y) \end{bmatrix}.$$ 

Differentiating - $\log p(\alpha, \beta | g)$ with respect to $\alpha$ and $\beta$ we have

$$2 \parallel W_{c}(u_{c} - v_{c}) \parallel^{2} + 2 \parallel W_{r}(u_{r} - v_{r}) \parallel^{2} +$$

$$\parallel W_{c}(f(c) - f(d)) \parallel^{2} + \parallel W_{r}(f(d) - f(a)) \parallel^{2} +$$

$$\max_{\alpha} \left( \frac{1}{2} p(\alpha) \right) = \frac{p}{\alpha}$$

$$\parallel u_{c} - x_{c} \parallel^{2} + \parallel v_{c} - y_{c} \parallel^{2} + \parallel u_{r} - x_{r} \parallel^{2} +$$

$$\parallel v_{r} - y_{r} \parallel^{2} + \parallel f(a) - g(a) \parallel^{2} + \parallel f(b) - g(b) \parallel^{2} +$$

$$\parallel f(c) - g(c) \parallel^{2} + \parallel f(d) - g(d) \parallel^{2} +$$

$$\sum_{i} \max_{\alpha} \left( \frac{1}{2} p(\alpha) \right) = \frac{1}{2\beta}$$

In summary, the following iterative algorithm can be used to recover $f$, an image with reduced blocking artifact.

1. Set $f^{0} = g$. Choose $\alpha^{0}$ and $\beta^{0}$.
2. For $k = 1, 2, \ldots$
   (a) Estimate $\alpha^{k}$ and $\beta^{k}$ by substituting $\alpha^{k-1}$ and $\beta^{k-1}$ in the left hand side of Eqs. (8) and (9), respectively.
   (b) For $i = 1, 2, \ldots$, compute $f^{k}$ from Eqs. (2), (3), (4), (5) and (6).
3. Goto 2 until $\parallel f^{k} - f^{k-1} \parallel$ is less than a prescribed bound.

It is clear that this process for estimating the image and the hyperparameters can also be performed at the coder.

In this case Eq. (1) becomes

$$f_{(\alpha, \beta)} = \arg \min_{f} \left\{ M(z, f | \alpha, \beta) \right\} = \arg \min_{f} \left\{ p(z | \alpha) + N(f | z, \beta) \right\},$$

and the iterative procedure described in the previous section can also be applied here. Let us then call $\alpha^{c}$ and $\beta^{c}$ the hyperparameters obtained by the iterative procedure described in the previous section with the use of the original image.

V. COMBINING INFORMATION FROM THE CODER

Let us assume we have been able to transmit the parameters $\alpha^{c}$ and $\beta^{c}$ to the decoder and want to combine them with the ones obtained in section IV which will be denoted $\alpha^{d}$ and $\beta^{d}$. Then, following the hierarchical Bayesian approach to the reconstruction problem and using Gamma distributions for the hyperparameters instead of the flat distribution given in Eq. (7) we would have the following iterative procedure.

1. Set $f^{0} = g$. Choose $\alpha^{0}$ and $\beta^{0}$.
2. For $k = 1, 2, \ldots$
   (a) Estimate $\alpha^{k}$ and $\beta^{k}$ by substituting $\alpha^{k-1}$ and $\beta^{k-1}$ in the left hand side of Eqs. (8) and (9), respectively.
   (b) $\alpha^{k} = \lambda \alpha^{c} + (1 - \lambda) \alpha^{d}$, $\beta^{k} = \lambda \beta^{c} + (1 - \lambda) \beta^{d}$.
TABLE I

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>PSNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>JPEG Blocky</td>
<td>29.58dB</td>
</tr>
<tr>
<td>AC Prediction</td>
<td>29.52dB</td>
</tr>
<tr>
<td>POCS Nonadaptive</td>
<td>30.32dB</td>
</tr>
<tr>
<td>POCS Adaptive</td>
<td>30.43dB</td>
</tr>
<tr>
<td>Proposed algorithm with</td>
<td>30.47dB</td>
</tr>
<tr>
<td>Estimation at the decoder</td>
<td></td>
</tr>
<tr>
<td>Estimation at the coder</td>
<td>30.51dB</td>
</tr>
</tbody>
</table>

PSNR of the reconstructed images with different algorithms.

(c) For \( i = 1, 2, \ldots \), compute \( f^i \) from Eqs. (2), (3), (4), (5) and (6).

3. Goto 2 until \( \| f^i - f^{i-1} \| \) is less than a prescribed bound.

It is clear that the way the parameters are weighted depend on how they improve the peak signal to noise ratio and the way they are transmitted. For instance, if the parameters are transmitted without loss and they result in a higher peak signal to noise ratio, then they should be used instead of the ones obtained at the decoder. However, if they are quantized, better results could be obtained by using a convex combination with the ones obtained at the decoder, as examined next.

VI. TEST EXAMPLES

In this section, experiments are presented in order to test the proposed recovery algorithm. The 512 \( \times \) 512 "Lena" image is used. The image was compressed using a JPEG based coder-decoder at a bit-rate of 0.24 bpp. For presentation purposes the center 256 \( \times \) 256 section of this image is shown in Fig. 2 having a PSNR of 29.58 dB.

The reconstructed image using the proposed algorithm at the coder is shown in Fig. 2. The corresponding PSNR is equal to 30.51dB. If the parameters are estimated at the decoder the corresponding PSNR is 30.47dB. If we use a linear convex combination of the above parameters, that is, with \( \lambda \in [0, 1] \), the PSNR moves from 30.51dB to 30.47dB as \( \lambda \) changes from 1 to 0. For comparison purposes, we show results with previous proposed algorithms in table I. The proposed approach outperformed the non POCS approach [3] and JPEG AC prediction [1] both in PSNR and visual quality of the image. The results seem also to be better than the spatially-adaptive POCS approach [5].

VII. 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. The performed tests show very good improvement in term of the PSNR metric and the visual quality of the image.