REDUCTION OF BLOCKING ARTIFACTS IN BLOCK TRANSFORMED COMPRESSED COLOR IMAGES

J. Mateos\textsuperscript{a}, C. Ilia, B. Jiménez, R. Molina\textsuperscript{a} and A. K. Katsaggelos\textsuperscript{b}

(a) Departamento de Ciencias de la Computación e Inteligencia Artificial, Universidad de Granada. 18071 Granada, Spain.
(b) Department of Electrical and Computer Engineering, Northwestern University. 2145 Sheridan Road. Evanston, IL 60208-3118.

ABSTRACT

In this paper we use the information in the chrominance bands to reconstruct color block transformed compressed images. For the luminance and the two chrominance channels, we define a reconstruction problem and show how to estimate the unknown hyperparameters and reconstruct each band automatically. The method is tested on real images.

1. INTRODUCTION

The block transformed (BT) techniques, and the block DCT in particular, are the most popular coding technique used by image compression applications. This approach is used to calculate the frequency components of a given signal sampled at a given sampling rate. In order to reduce the bit-rate and achieve compression before transmission, the transformed data are quantized. Block-based transform may generate a significant artifact called "blocking artifact", and this effect is accentuated at 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 correlation. To increase the use of the BT based coders several algorithms have been developed that reduce the blocking artifact \cite{5,6,8}. In \cite{8} a POCS based recovery algorithm that uses a spatially-adaptive constraint enforcing between-block smoothness is proposed. The method proposed in \cite{8}, like most of the methods referenced above, requires the estimation of unknown parameters. These parameters are usually estimated using ad hoc techniques. In \cite{1,2} we formulate algorithms similar to the one in \cite{8} within the hierarchical Bayesian paradigm. This algorithms reconstruct the image and estimate the regularization parameters, which represent the image and noise variances \cite{3}, at the same time.

Most of the proposed algorithms to reduce the blocking artifact deal only with grey scale images, and if applied to color images, they only process the Y (luminance) band in the YCbCr format; that is, they do not modify the Cb and Cr bands. In the JPEG specifications for color images \cite{5} each band is coded independently, that is the same coding is performed on the chrominance and the luminance, but using different quantization tables. In JPEG, a subsampling technique is also applied to chrominance bands, usually noted as 4:2:2 or 4:1:1. The format 4:2:2 specifies that for every four samples of Y information there are two samples of Cb and Cr while the 4:1:1 format, for every four samples of Y, there is one sample from Cb and Cr.

In this paper we extend the algorithm developed in \cite{2} to color images by processing each band of the JPEG compressed color image. In section 2 we introduce the elements needed to perform the reconstruction. In section 3 we describe the algorithm which simultaneously estimates the color image and the unknown hyperparameters and, in section 4, we show some experimental tests. Finally, section 5 concludes the paper.

2. IMAGE AND NOISE MODELS

Let $f$ be a three band color image, whose components will be denoted $f^I$, $I \in \{Y,Cb, Cr\}$ and let $g$ be the corresponding compressed image. Let $U \times V$ be the original size of each band and assume that $k \times k$ blocks are used to codify each band, where $U$ and $V$ are multiples of $k$.

For a 4:1:1 subsampling, the one used in this paper, we note that for the luminance band, $Y$, a $M_Y \times N_Y$ image is processed with $M_Y = U$ and $N_Y = V$ and, for the chrominance bands, the images are $M_I \times N_I$.
Figure 1: Distribution of boundary pixels and weights for each band \( I \), \( I \in \{ Y, Cb, Cr \} \).

where \( M_I = U/2 \) and \( N_I = V/2 \) with \( I \in \{ Cb, Cr \} \). However, in both cases the blocks used are \( k \times k \). Note that, although we are assuming a 4:1:1 subsampling, the proposed algorithm will be independent of the used subsampling allowing to use any other subsampling, or no subsampling at all, simply by using the corresponding value of \( M_I \) and \( N_I \), \( I \in \{ Y, Cb, Cr \} \).

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.

For \( I \in \{ Y, Cb, Cr \} \), let \( f^I_{\alpha} \) be a column vector defined by stacking all the elements of \( f^I \) which are at the left of a block boundary column, \( c_l \), but not in a four-block intersection (see Fig. 1). Analogously we define \( f^I_{\alpha} \) by stacking all the elements of \( f^I \) that are at the right of a block boundary column, \( c_r \), but not in a four-block intersection (see Fig. 1). Similarly \( g^I_{\alpha} \) and \( g^I_{\alpha} \) are defined from \( g^I \). The size of all these vectors is \( p^I = (k - 2)(M_I/k) \times (N_I/k - 1) \).

We also define the vectors \( f^I_{\alpha} = (f^I_{\alpha}, f^I_{\beta}) \) and \( g^I_{\alpha} = (g^I_{\alpha}, g^I_{\beta}) \). In a similar way we define \( f^I_{\alpha} \), \( g^I_{\alpha} \), \( f^I_{\beta} \) and \( g^I_{\beta} \), the \( q \times 1 \) column vectors representing the rows above and below a row block boundary of \( f^I \) and \( g^I \), respectively (see Fig. 1), with \( q = (k - 2)(N_I/k) \times (M_I/k - 1) \). Related to them are the vectors \( f^I_{\alpha} = (f^I_{\alpha}, f^I_{\alpha}) \) and \( g^I_{\alpha} = (g^I_{\alpha}, g^I_{\alpha}) \).

We further stack all the elements of \( f^I \) above a horizontal boundary and to the left of a vertical boundary, indicated by \( aI \) in Fig. 1, into \( m_I \times 1 \) vector \( f^I_{\alpha} \) with \( m_I = (M_I/k - 1) \times (N_I/k - 1) \). Similarly we form vectors \( f^I_{\alpha}, f^I_{\beta} \) and \( f^I_{\alpha} \). In a similar fashion we define the observation vectors for these pixels in a four-block boundary \( g^I_{\alpha}, g^I_{\alpha}, g^I_{\alpha} \) and \( g^I_{\alpha} \). Using the vectors above, we also define \( f^I_{\alpha} = (f^I_{\alpha}, f^I_{\alpha}, f^I_{\alpha}) \) and \( g^I_{\alpha} = (g^I_{\alpha}, g^I_{\alpha}, g^I_{\alpha}) \).

We shall use \( t^I = (f^I_{\alpha}, f^I_{\alpha}, f^I_{\alpha}) \) and \( g^I_{\alpha} = (g^I_{\alpha}, g^I_{\alpha}, g^I_{\alpha}) \), when needed. Note that originally \( f \) and \( g \) contained the original and compressed image, respectively, but, since we are only going to modify the pixels in the boundaries, we can use the above expression to denote \( f \) and \( g \).

To capture the vertical local properties of the image we define, for \( I \in \{ Y, Cb, Cr \} \), a \( p^I \times p^I \) diagonal matrix, \( W^I_c \), where the diagonal elements, \( \omega^I_c(i) \)'s, weight each image differences in block column vectors as shown in Fig. 1 (note that the corners pixels have not been included in this matrix). Analogously, we can define \( W^I_c \) to capture the horizontal local properties of the image without including the four block pixels. We also define the matrices \( W^I_{11}, W^I_{22}, W^I_{33} \) and \( W^I_{44} \) to weight the differences \( t^I_{\alpha} - t^I_{\alpha}, t^I_{\alpha} - t^I_{\alpha}, t^I_{\alpha} - t^I_{\alpha} \) and \( f^I_{\alpha} - f^I_{\alpha} \), as shown in Fig. 1, see [8] for details.

The Bayesian paradigm applied to our problem uses the model defined as

\[
p(f, g | \alpha, \alpha_r, \beta) = \prod_{I \in \{ Y, Cb, Cr \}} p(t^I | \alpha, \alpha_r)p(g^I | t^I, \beta),
\]

Let us now examine the terms involved in this equation.

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

\[
p(t^I | \alpha, \alpha_r) = \alpha^I_c \frac{\beta^I_c}{\alpha^I_r} \frac{\beta^I_r}{\alpha^I_c} \frac{P^I(t^I | \alpha^I_c, \alpha^I_r)}{P^I(t^I | \alpha^I_c, \alpha^I_r)} \times \exp \left\{ -P^I(t^I | \alpha^I_c, \alpha^I_r) \right\},
\]

with

\[
P^I(t^I | \alpha^I_c, \alpha^I_r) = \frac{\alpha^I_c(1 + \frac{1}{\omega^I_{21}(i)} + \frac{1}{\omega^I_{32}(i)})}{\omega^I_{11}(i)} + \frac{\alpha^I_r(1 + \frac{1}{\omega^I_{21}(i)} + \frac{1}{\omega^I_{32}(i)})}{\omega^I_{11}(i)}
\]

and \( \alpha^I_c \) and \( \alpha^I_r \) measures of the roughness between two block boundaries columns and rows, respectively.
The other component we have to consider is the fidelity to the observed data, which is given by

\[ p(g^l | f^l, \beta^l) \propto \beta^{I(p^{tr} + q^{tr} + 2m_I)} \exp \left\{ - N^I (g^l | f^l, \beta^l) \right\}, \]

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

\[ N^I (g^l | f^l, \beta^l) = \frac{1}{2} \beta \left\{ \| g_{cr}^l - f_{cr}^l \|^2 + \| g_{or}^l - f_{or}^l \|^2 \right\} + \frac{1}{2} \beta \left\{ \| g_{cb}^l - f_{cb}^l \|^2 + \| g_{or}^l - f_{or}^l \|^2 \right\} + \frac{1}{2} \beta \left\{ \| g_{or}^l - f_{or}^l \|^2 + \| g_{cb}^l - f_{cb}^l \|^2 \right\}, \]

and \( \beta^l \) is defined as \( \beta^{-1} = (\sigma^I_{\text{noise}})^2 \), the noise variance in the image.

If \( \alpha_c^l, \alpha_r^l \) and \( \beta^l \) are known, then following the Bayesian paradigm it is customary to select, as the reconstruction of \( f^l \), the image \( \hat{f}^l \) defined by

\[ \hat{f}^l = \arg \min_{\bar{f}^l} \left\{ C^I (\bar{f}^l, g^l | \alpha_c^l, \alpha_r^l, \beta^l) \right\}, \quad (1) \]

where \( C^I (\bar{f}^l, g^l | \alpha_c^l, \alpha_r^l, \beta^l) \) is defined as

\[ C^I (\bar{f}^l, g^l | \alpha_c^l, \alpha_r^l, \beta^l) = \frac{P^I (f^l | \alpha_c^l, \alpha_r^l)}{N^I (g^l | f^l, \beta^l)}. \quad (2) \]

Note that we are processing each band independently but information on the chrominance can be used to reconstruct luminance (see [7]).

### 3. PROPOSED ALGORITHM

For \( I \in \{ Y, C_b, C_r \} \), given \( \alpha_c^l, \alpha_r^l \) and \( \beta^l \), \( \hat{f}^l \) is estimated by differentiating Eq. (2) to obtain

\[ \hat{f}_{ci}^l(i) = 0.5[1 + \gamma_{ci}^l(i)]g_{ci}^l(i) + 0.5[1 - \gamma_{ci}^l(i)]g_{cr}^l(i), \quad (3) \]
\[ \hat{f}_{cr}^l(i) = 0.5[1 - \gamma_{cr}^l(i)]g_{ci}^l(i) + 0.5[1 + \gamma_{cr}^l(i)]g_{cr}^l(i), \quad (4) \]
\[ \hat{f}_{rb}^l(i) = 0.5[1 + \gamma_{rb}^l(i)]g_{ra}^l(i) + 0.5[1 - \gamma_{rb}^l(i)]g_{rb}^l(i), \quad (5) \]
\[ \hat{f}_{br}^l(i) = 0.5[1 - \gamma_{br}^l(i)]g_{ra}^l(i) + 0.5[1 + \gamma_{br}^l(i)]g_{rb}^l(i), \quad (6) \]

with \( \gamma_{ci}^l(i) = \beta^l (\beta^l + 4 \alpha_c^l (\omega_{ci}^l(i))^2)^{-1} \) and \( \gamma_{cr}^l(i) = \beta^l (\beta^l + 4 \alpha_r^l (\omega_{cr}^l(i))^2)^{-1} \) and, for pixels in a four-block intersection, the solution of the following system of equations

\[ \alpha_c^l \mathbf{W}_{x1}^l (\hat{f}_{ci}^l - \hat{f}_{cr}^l) + \alpha_r^l \mathbf{W}_{x2}^l (\hat{f}_{al}^l - \hat{f}_{ar}^l) - \beta^l (g_{cr}^l - \hat{f}_{al}^l) = 0, \quad (7) \]

\[ \alpha_c^l \mathbf{W}_{x1}^l (\hat{f}_{ci}^l - \hat{f}_{cr}^l) + \alpha_r^l \mathbf{W}_{x2}^l (\hat{f}_{al}^l - \hat{f}_{cr}^l) - \beta^l (g_{cr}^l - \hat{f}_{al}^l) = 0, \quad (8) \]

\[ \alpha_c^l \mathbf{W}_{x1}^l (\hat{f}_{ci}^l - \hat{f}_{cr}^l) + \alpha_r^l \mathbf{W}_{x2}^l (\hat{f}_{al}^l - \hat{f}_{cr}^l) - \beta^l (g_{cr}^l - \hat{f}_{al}^l) = 0, \quad (9) \]

\[ \alpha_c^l \mathbf{W}_{x1}^l (\hat{f}_{ci}^l - \hat{f}_{cr}^l) + \alpha_r^l \mathbf{W}_{x2}^l (\hat{f}_{al}^l - \hat{f}_{cr}^l) - \beta^l (g_{cr}^l - \hat{f}_{al}^l) = 0. \quad (10) \]

It is obvious that in order to estimate the original image, that is, in order to solve Eq. (1), it is necessary to know or estimate the unknown hyperparameters. The method proposed amounts to selecting, for \( I \in \{ Y, C_b, C_r \} \), \( \alpha_c^l, \alpha_r^l \) and \( \beta^l \), as the maximum likelihood estimates, mle, of \( \alpha_c^l, \alpha_r^l \) and \( \beta^l \) from \( p(g^l | \alpha_c^l, \alpha_r^l, \beta^l) \).

Note that including the data about the four-block boundaries in the estimation process of \( \alpha_c^l, \alpha_r^l \) and \( \beta^l \) has two problems. First, those data are less reliable than the ones in the rows and columns since they are mixed. Furthermore, it needs the inversion of \((M_I/k - 1)/N_I/k - 1) \times 4 \times 4\) matrices for each band \( I \in \{ Y, C_b, C_r \} \). Due to these reasons we proceeded to estimate \( \hat{\alpha}_c^l, \hat{\alpha}_r^l \) and \( \hat{\beta}^l \) from the data from rows and columns but not in the intersections (see [2] for details). Then, the E-M algorithm [4] to estimate \( \hat{\alpha}_c^l, \hat{\alpha}_r^l \) and \( \hat{\beta}^l \) gives the following iterative scheme to update the hyperparameters when we use the old values on the right hand side of the following equations

\[ \frac{p_I}{\hat{\alpha}_c^l} = \sum_{i=1}^{p_I} \text{tr} [q_{c}^{-1} (\hat{\alpha}_c^l, \hat{\beta}^l(i)) \mathbf{A}_c^l(i)] + 2 \| \mathbf{W}_1^l (\hat{f}_{ci}^l - \hat{f}_{cr}^l) \|^2 \quad (11) \]

\[ \frac{q_I}{\hat{\alpha}_r^l} = \sum_{i=1}^{q_I} \text{tr} [q_{r}^{-1} (\hat{\alpha}_r^l, \hat{\beta}^l(i)) \mathbf{A}_r^l(i)] + 2 \| \mathbf{W}_2^l (\hat{f}_{al}^l - \hat{f}_{rb}^l) \|^2 \quad (12) \]

\[ \frac{2(p_I + q_I)}{\hat{\beta}^l} = \sum_{i=1}^{p_I} \text{tr} [q_{c}^{-1} (\hat{\alpha}_c^l, \hat{\beta}^l(i))] + \sum_{i=1}^{q_I} \text{tr} [q_{r}^{-1} (\hat{\alpha}_r^l, \hat{\beta}^l(i))] \]

\[ \quad + \| g_{ci}^l - \hat{f}_{ci}^l \|^2 + \| g_{cr}^l - \hat{f}_{cr}^l \|^2 + \| g_{ra}^l - \hat{f}_{ra}^l \|^2 + \| g_{rb}^l - \hat{f}_{rb}^l \|^2, \quad (13) \]

where

\[ q_{c}^l (\hat{\alpha}_c^l, \hat{\beta}^l(i)) = \beta^l \mathbf{I}_{2 \times 2} + \alpha_c^l \mathbf{A}_c^l(i), \]

and

\[ q_{r}^l (\hat{\alpha}_r^l, \hat{\beta}^l(i)) = \beta^l \mathbf{I}_{2 \times 2} + \alpha_r^l \mathbf{A}_r^l(i), \]
with
\[ A_l^I(i) = \begin{bmatrix}
2\omega_l^2(i) & -2\omega_l^2(i) \\
-2\omega_l^2(i) & 2\omega_l^2(i)
\end{bmatrix}, \]
\[ A_r^I(i) = \begin{bmatrix}
2\omega_r^2(i) & -2\omega_r^2(i) \\
-2\omega_r^2(i) & 2\omega_r^2(i)
\end{bmatrix}. \]

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

1. Divide the image \( g \) into three different images, \( g^Y, g^{Cb} \) and \( g^{Cr} \), to represent each band of the image.

2. For \( I \in \{ Y, Cb, Cr \} \)
   (a) Choose an initial value for \( \alpha_c^I, \alpha_r^I \) and \( \beta^I \),
   (b) For \( k = 1, 2, \ldots \)
      i. Compute \( \hat{f}_c^{I,k} \) and \( \hat{f}_r^{I,k} \) from Eqs. (3)-(6).
      ii. Estimate \( \alpha_c^{I,k}, \alpha_r^{I,k} \) and \( \beta^{I,k} \) by substituting \( \alpha_c^{I,k-1}, \alpha_r^{I,k-1} \) and \( \beta^{I,k-1} \) in the right hand side of Eqs. (11), (12) and (13).
   (c) Go to step 2b until \( \| \hat{f}_c^{I,k} - \hat{f}_c^{I,k-1} \| + \| \hat{f}_r^{I,k} - \hat{f}_r^{I,k-1} \| \) is less than a prescribed bound.
   (d) Using the estimated \( \alpha_c^{I,k}, \alpha_r^{I,k} \), \( \beta^{I,k} \) calculate \( f^I \) by solving Eqs. (7)-(10).

3. The reconstructed image, \( f \), is the result of merge the three bands, \( f^Y, f^{Cb} \) and \( f^{Cr} \).

To calculate the distance between a reconstructed band and its original band, we have used the peak signal-to-noise-ratio (PSNR) defined as
\[ PSNR^I = 10\log_{10} \left[ \frac{M \cdot N \times 255^2}{\| f^I - f^I \|^2} \right]. \]

This measure is applied to each band independently since there are no standard visual quality improvement measures for color images. We also note that it is well known that PSNR does not always match the human perception for image quality.

In Table 1 we show the PSNR results for the compressed and the proposed algorithm. Figure 2 depicts the compressed and reconstructed central part of the Cb and Cr channels and figure 3 the central part of the luminance band. The proposed algorithm improves both the PSNR of each band and the visual quality of the color image.

5. CONCLUSION

In this paper we have proposed a new method to reconstruct color images compressed using BDCT. The method is fully automatic and does not require the setting of any parameters. All of them are estimated within the proposed algorithm and the convergence is guaranteed. The three bands are reconstructed independently based on the hierarchical Bayesian approach to image reconstruction.

Although the proposed method treats each band independently, the possibility of combining between band information is currently being investigated. Work is also in progress on the estimation of the parameters at the coder and the transmission of a quantized version of them to the decoder.

6. REFERENCES


<table>
<thead>
<tr>
<th>Band</th>
<th>Blocky</th>
<th>Reconstructed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>30.45</td>
<td>30.82</td>
</tr>
<tr>
<td>Cb</td>
<td>33.74</td>
<td>33.86</td>
</tr>
<tr>
<td>Cr</td>
<td>33.89</td>
<td>33.97</td>
</tr>
</tbody>
</table>

Table 1: PSNR of the compressed and reconstructed bands.

4. TEST EXAMPLES

In order to test this algorithm, we have used a 512×512 "Lena" image. This color image has 24 bits per pixel and has been compressed using a JPEG-based coder-decoder with quality factor equal to 10, to obtain a 0.30 bpp representation.


Figure 2: Lena compressed ((a) Cb and (b) Cr channel) and (c) and (d) its reconstruction.

Figure 3: (a) Lena compressed (Y channel) and (b) its reconstruction.