ON A PERCEPTUAL RATE-DISTORTION MODEL FOR COLOR IMAGE CODING

Evgeny Gershikov and Moshe Porat

Department of Electrical Engineering, Technion - Israel Institute of Technology
Technion City, 32000, Haifa, Israel
phone: +972-4-8294725, fax: +972-4-8294799, email: eugeny@tx.technion.ac.il
web: http://visl.technion.ac.il/~eugeny
phone: +972-4-8294684, fax: +972-4-8295757, email: mp@ee.technion.ac.il
web: http://visl.technion.ac.il/mp

ABSTRACT

The Mean Square Error (MSE) is a common distortion measure used to assess image quality although it is not always justified when compared to the human observer. In this work we present a Rate-Distortion approach to color image compression based on subband transforms using perceptual optimization of the compression quality. This approach is based on minimization of the Weighted Mean Square Error (WMSE) of the encoded image, which better corresponds to quality assessment of the human eye. Based on the new approach, new optimized compression algorithms are introduced using the Discrete Cosine Transform and the Discrete Wavelet Transform. We compare the new algorithms to presently available algorithms such as JPEG. Our conclusion is that the new WMSE optimization approach offers superior performance when a human observer is considered.

1. BACKGROUND

Many color image coding algorithms rely on subband transforms (SBT) to enhance the compression performance. The range of such algorithms is wide: from those based on basic block transforms like the Discrete Cosine Transform (e.g., JPEG [1]) to more complex algorithms based on the Discrete Wavelet Transform (e.g., EZW - Embedded Zerotree Wavelet [2] and JPEG2000 [3],[4]). Recently, a Rate-Distortion (R-D) model has been introduced for SBT coders [5] based on the MSE as a distortion measure, which is not always well correlated with image quality assessment of the human eye. Other distortion measures have been proposed, such as calculating the MSE distortion after an intensity transformation and filtering [6] or using a non-linear transformation of the primaries, followed by filtering [7]. These measures, however, are too complicated for the proposed optimization process. Thus, in this work we develop a perceptual R-D model for subband transform coders based on the simpler WMSE as the distortion measure. We demonstrate the efficiency of the new model for subband coding by presenting a new type of compression algorithms based on perceptual optimization of the pre-processing stage and of the subband rates.

1.1 Rate-Distortion theory of subband transform coders

Consider a general subband transform coder for color images where the image samples are passed through the following steps:

1. Pre-processing by applying a Color Components Transform (CCT) to the RGB color components of the image. We denote the RGB components in vector form as $\mathbf{x} = [R\ G\ B]^T$ and the new color components as $\widetilde{\mathbf{x}} =$ $[C_1\ C_2\ C_3]^T$. The 3×3 size CCT matrix is denoted by $\mathbf{M}$. This stage can be written as:

$$\widetilde{\mathbf{x}} = \mathbf{M}\mathbf{x}. \qquad (1)$$

The goal of using a CCT transform is usually to reduce the high inter-color correlations of the RGB components [8], [9], although in some cases another choice of a CCT based on a correlation approach could be preferred [10].

2. A subband transform is applied to each color component and the subband coefficients are quantized. An independent uniform scalar quantizer for each subband is used.

3. Lossless post-quantization coding, e.g., entropy coding.

Assuming that a subband transform with B subbands is applied to each color component, it can be shown that the Rate-Distortion model of this algorithm is [11]:

$$d\left(\{R_{bi}\}, \mathbf{M}\right) = \frac{1}{3} \sum_{i=1}^{3} \sum_{b=0}^{B-1} \eta_b G_b \widetilde{\sigma}_{bi}^2 \varepsilon_i^2 e^{-aR_{bi}} \left((\mathbf{M}\mathbf{M}^T)^{-1}\right)_{ii}. \qquad (2)$$

Here $\widetilde{\sigma}_{bi}^2$ is the variance of subband b ($b \in \{0, 1, ..., B-1\}$) of the color component C_i ($i \in \{1, 2, 3\}$), G_b is its energy gain [12] and R_{bi} is the rate allocated to it. η_b is its sample rate, i.e., the relative part of the number of coefficients in it from the total number of samples in the color component. Finally, ε^2 is a constant dependent upon the distribution of the coded signal and a is $2ln2$.

Optimal rates allocation for the subbands can be found by minimizing the expression of Equation (2) under the rate constraint:

$$\sum_{i=1}^{3} \alpha_i \sum_{b=0}^{B-1} \eta_b R_{bi} = R \qquad (3)$$

for some total image rate R. Here down-sampling factors α_i have been used. The optimal rates under the rate constraint of (3) as well as non-negativity constraints are:

$$R_{bi} = \frac{R}{\sum\limits_{j=1}^{3} \alpha_j \xi_j} + \frac{1}{a} ln \left(\frac{\frac{\varepsilon_i^2 G_b \widetilde{\sigma}_{bi}^2 \left((\mathbf{M}\mathbf{M}^T)^{-1}\right)_{ii}}{\alpha_i}}{\prod_{k=1}^{3} \left(\frac{\left((\mathbf{M}\mathbf{M}^T)^{-1}\right)_{kk} \varepsilon_k^2 GM_k^{Act}}{\alpha_k} \right)^{\frac{\alpha_k \xi_k}{\sum\limits_{j=1}^{3} \alpha_j \xi_j}}} \right) \qquad (4)$$

for $b \in Act_i$ where

$$Act_i \triangleq \{b \in [0, B-1] \mid R_{bi} > 0\}.$$

$$\xi_i \triangleq \sum_{b \in Act_i} \eta_b, \quad GM_i^{Act} \triangleq \prod_{b \in Act_i} (G_b \widetilde{\sigma}_{bi}^2)^{\frac{\eta_b}{\xi_i}}. \qquad (5)$$

Act_i denotes the set of non zero (or active) rates of C_i.

2. THE PERCEPTUAL R-D MODEL

We assume here that we are given the visual perception weights corresponding to the subbands of a certain subband transform in a color space. Such a space can be, for example, YCbCr. We now wish to derive an expression for the WMSE distortion of a coder based on the subband transform. The same coder described in Subsection 1.1 is assumed, so that a CCT is applied to the RGB color components of the image prior to coding and the actual image data compression is performed in another color space denoted C1C2C3. We denote by $\mathbf{Y_b} = [y_b^Y \; y_b^{Cb} \; y_b^{Cr}]^T$ the vector of the SBT coefficients at some index in subband b in the YCbCr color space. Similarly, the vectors of subband b coefficients in the RGB and C1C2C3 spaces are denoted $\mathbf{Y_b^{RGB}} = [y_b^R \; y_b^G \; y_b^B]^T$ and $\widetilde{\mathbf{Y}}_\mathbf{b} = [y_b^{C_1} \; y_b^{C_2} \; y_b^{C_3}]^T$ respectively. Due to the linearity of the SBT the following relationship holds:

$$\widetilde{\mathbf{Y}}_\mathbf{b} = \bar{\mathbf{M}}\mathbf{Y_b} \Rightarrow \mathbf{Y_b} = \bar{\mathbf{M}}^{-1}\widetilde{\mathbf{Y}}_\mathbf{b}, \qquad (6)$$

where $\bar{\mathbf{M}}$ stands for the CCT matrix from YCbCr to C1C2C3. If $\mathbf{M}$ is the CCT matrix from RGB to C1C2C3, and $\mathbf{M_{YCbCr}}$ is the RGB to YCbCr matrix, then:

$$\bar{\mathbf{M}} = \mathbf{M} \cdot \mathbf{M_{YCbCr}^{-1}}. \qquad (7)$$

Since the SBT coefficients are lossy encoded, errors are introduced between the reconstructed coefficients $\mathbf{Y_b^{rec}}$ in the YCbCr color space and the original ones. The error covariance matrices for subband b in the YCbCr and C1C2C3 domains respectively are:

$$\begin{aligned} \mathbf{Er_b} &= E\left[(\mathbf{Y_b} - \mathbf{Y_b^{rec}})(\mathbf{Y_b} - \mathbf{Y_b^{rec}})^T\right], \\ \widetilde{\mathbf{Er_b}} &= E\left[(\widetilde{\mathbf{Y}}_\mathbf{b} - \widetilde{\mathbf{Y}}_\mathbf{b}^{\mathbf{rec}})(\widetilde{\mathbf{Y}}_\mathbf{b} - \widetilde{\mathbf{Y}}_\mathbf{b}^{\mathbf{rec}})^T\right]. \end{aligned} \qquad (8)$$

$E()$ stands here for statistic mean. Using (6), we can express $\widetilde{\mathbf{Er_b}}$ by $\mathbf{Er_b}$ as:

$$\mathbf{Er_b} = \bar{\mathbf{M}}^{-1}\widetilde{\mathbf{Er_b}}\bar{\mathbf{M}}^{-T}. \qquad (9)$$

The MSE distortions d_{bi} of the YCbCr color components in subband b are the diagonal elements of $\mathbf{Er_b}$ and thus:

$$d_{bi} = \mathbf{n_i}^T\widetilde{\mathbf{Er_b}}\mathbf{n_i}, \qquad (10)$$

where $\mathbf{n_i}$ is the ith row of $\bar{\mathbf{M}}^{-1}$ in column form. In a similar fashion the diagonal elements of $\widetilde{\mathbf{Er_b}}$ can be recognized as the MSE distortions $\tilde{d}_{bi}$ of the C_1, C_2, C_3 color components, given by [12]:

$$\tilde{d}_{bi} = \varepsilon_i^2 \tilde{\sigma}_{bi}^2 e^{-aR_{bi}}. \qquad (11)$$

Note that we continue here with the consistent notation of a tilde for the variables related to the C1C2C3 color space. Assuming that the quantization errors of the three color components in each subband in the C1C2C3 domain are uncorrelated, $\widetilde{\mathbf{Er_b}}$ becomes a diagonal matrix and (10) becomes:

$$d_{bi} = \sum_{k=1}^{3} \mathbf{n_{ik}^2}\tilde{d}_{bk} = \sum_{k=1}^{3} (\bar{\mathbf{M}}^{-1})_{ik}^2 \varepsilon_k^2 \tilde{\sigma}_{bk}^2 e^{-aR_{bk}} \qquad (12)$$

after substitution of (11) for $\tilde{d}_{bk}$. Now if, for the sake of convenience, we denote the YCbCr color components at each pixel as a vector $\mathbf{x^{YCbCr}} = [Y \; Cb \; Cr]^T$, then the WMSE of the ith color component $(\mathbf{x^{YCbCr}})_i$ is:

$$WMSE\left((\mathbf{x^{YCbCr}})_i\right) = \sum_{b=0}^{B-1} \eta_b G_b w_{bi} d_{bi}. \qquad (13)$$

As can be seen, this expression incorporates the energy gains of the subbands G_b as well as their sample rates η_b. Also the visual weights w_{bi} are included in the expression to provide varying significance to different subbands of the same color component as well as between color components. Defining the total WMSE as the average WMSE of the YCbCr color components, we get:

$$WMSE = \frac{1}{3}\sum_{i=1}^{3} WMSE\left((\mathbf{x^{YCbCr}})_i\right) = \frac{1}{3}\sum_{i=1}^{3}\sum_{b=0}^{B-1} \eta_b G_b w_{bi} d_{bi} \qquad (14)$$

and after substituting (12) for d_{bi} the expression becomes:

$$\begin{aligned} WMSE &= \frac{1}{3}\sum_{i=1}^{3}\sum_{b=0}^{B-1} \eta_b G_b w_{bi} \sum_{k=1}^{3} (\bar{\mathbf{M}}^{-1})_{ik}^2 \varepsilon_k^2 \tilde{\sigma}_{bk}^2 e^{-aR_{bk}} \\ &= \frac{1}{3}\sum_{b=0}^{B-1} \eta_b G_b \sum_{k=1}^{3} \varepsilon_k^2 \tilde{\sigma}_{bk}^2 e^{-aR_{bk}} \sum_{i=1}^{3} w_{bi} (\bar{\mathbf{M}}^{-1})_{ik}^2. \end{aligned} \qquad (15)$$

To simplify (15) we denote:

$$\psi_{bk} \triangleq \sum_{i=1}^{3} w_{bi} (\bar{\mathbf{M}}^{-1})_{ik}^2, \qquad (16)$$

so that the WMSE expression becomes:

$$WMSE = \frac{1}{3}\sum_{b=0}^{B-1}\sum_{k=1}^{3} \eta_b G_b \tilde{\sigma}_{bk}^2 \varepsilon_k^2 e^{-aR_{bk}} \psi_{bk}. \qquad (17)$$

Clearly if the visual weights w_{bi} are all equal to 1, the WMSE expression of (17) should become the expression for the MSE in the YCbCr domain. This expression is given exactly by (2) with the difference that $\mathbf{M}$ there is to be replaced by $\bar{\mathbf{M}}$ in our case. From the comparison of equations (17) and (2) we conclude that $\psi_{bk} = \left((\bar{\mathbf{M}}\bar{\mathbf{M}}^T)^{-1}\right)_{kk}$ in that case, which means according to (16) that

$$\sum_{i=1}^{3} (\bar{\mathbf{M}}^{-1})_{ik}^2 = \left((\bar{\mathbf{M}}\bar{\mathbf{M}}^T)^{-1}\right)_{kk}. \qquad (18)$$

2.1 Basic optimization using the WMSE model

After deriving the WMSE expression, the next step is to use it to find the optimal rates and optimal CCT in the WMSE sense. First we wish to minimize the WMSE of (17) subject to the rate constraint $\sum_{i=1}^{3}\sum_{b=0}^{B-1} \eta_b R_{bi} = R$, resulting in the following Lagrangian (λ is the Lagrange multiplier):

$$\begin{aligned} L\left(\{R_{bi}\}, \bar{\mathbf{M}}, \lambda\right) &= \frac{1}{3}\sum_{b=0}^{B-1}\sum_{k=1}^{3} \eta_b G_b \tilde{\sigma}_{bk}^2 \varepsilon_k^2 e^{-aR_{bk}} \psi_{bk} \\ &+ \lambda\left(\sum_{i=1}^{3}\sum_{b=0}^{B-1} \eta_b R_{bi} - R\right), \end{aligned} \qquad (19)$$

which is minimized by the optimal rates given by:

$$R_{bi} = \frac{R}{3} + \frac{1}{a} ln \left(\frac{G_b \widetilde{\sigma}_{bi}^2 \varepsilon_i^2 \psi_{bi}}{\prod_{k=1}^3 \left(GM_k \varepsilon_k^2 \Psi_k \right)^{\frac{1}{3}}} \right). \qquad (20)$$

Here

$$GM_k \triangleq \prod_{b=0}^{B-1} \left(G_b \widetilde{\sigma}_{bk}^2 \right)^{\eta_b} \quad \text{and} \quad \Psi_k \triangleq \prod_{b=0}^{B-1} \left(\psi_{bk} \right)^{\eta_b}. \qquad (21)$$

Note that no constraints for non-negativity of the rates were used here, which means that high image rates R are assumed. As for the optimal CCT matrix $\bar{\mathbf{M}}$: it can be found by minimizing the target function

$$f(\bar{\mathbf{M}}) = \prod_{k=1}^3 (GM_k \Psi_k) = \prod_{k=1}^3 \prod_{b=0}^{B-1} \left(G_b \widetilde{\sigma}_{bk}^2 \psi_{bk} \right)^{\eta_b}. \qquad (22)$$

We should remind here that ψ_{bk} is a function of $\bar{\mathbf{M}}$ given in (16). Also the variances $\widetilde{\sigma}_{bk}^2$ depend on $\bar{\mathbf{M}}$, or more specifically on $\mathbf{M}$. These variances are the diagonal elements of the subband b covariance matrix in the C1C2C3 domain:

$$\widetilde{\mathbf{\Lambda}}_{\mathbf{b}} \triangleq E \left[\left(\widetilde{\mathbf{Y}}_{\mathbf{b}} - \tilde{\mu}_{\mathbf{Y_b}} \right) \left(\widetilde{\mathbf{Y}}_{\mathbf{b}} - \tilde{\mu}_{\mathbf{Y_b}} \right)^T \right] \quad \tilde{\mu}_{\mathbf{Y_b}} \triangleq E \left[\widetilde{\mathbf{Y}}_{\mathbf{b}} \right], \qquad (23)$$

and can also be expressed using the $\mathbf{M}$ matrix and the subband b covariance matrix in the RGB domain:

$$\mathbf{\Lambda}_{\mathbf{b}} \triangleq E \left[\left(\mathbf{Y}_{\mathbf{b}}^{\mathbf{RGB}} - \mu_{\mathbf{Y_b}}^{\mathbf{RGB}} \right) \left(\mathbf{Y}_{\mathbf{b}}^{\mathbf{RGB}} - \mu_{\mathbf{Y_b}}^{\mathbf{RGB}} \right)^T \right]$$
$$\mu_{\mathbf{Y_b}}^{\mathbf{RGB}} \triangleq E \left[\mathbf{Y}_{\mathbf{b}}^{\mathbf{RGB}} \right] \qquad (24)$$

according to $\widetilde{\sigma}_{bk}^2 = \mathbf{m_k}^T \mathbf{\Lambda_b m_k}$, where $\mathbf{m_k}$ denotes the kth row of the $\mathbf{M}$ matrix in vector form. Thus the target function $f(\bar{\mathbf{M}})$ can be rewritten as:

$$f(\bar{\mathbf{M}}) = \prod_{k=1}^3 \prod_{b=0}^{B-1} \left((\mathbf{m_k}^T \mathbf{\Lambda_b m_k}) G_b \psi_{bk} \right)^{\eta_b}. \qquad (25)$$

2.2 Optimal rates with down-sampling

When considering possible down-sampling of some of the color components, the rate constraint becomes (3) and the Lagrangian that incorporates this constraint as well as constraints for the non-negativity of the subband rates is:

$$L(\{R_{bi}\}, \mathbf{M}, \lambda, \{\mu_{bi}\}) = \frac{1}{3} \sum_{b=0}^{B-1} \sum_{k=1}^3 \eta_b G_b \widetilde{\sigma}_{bk}^2 \varepsilon_k^2 e^{-aR_{bk}} \psi_{bk}$$
$$+ \lambda \left(\sum_{i=1}^3 \alpha_i \sum_{b=0}^{B-1} \eta_b R_{bi} - R \right) - \sum_{i=1}^3 \sum_{b=0}^{B-1} \mu_{bi} R_{bi}, \qquad (26)$$

where μ_{bi} are the Lagrange multipliers for the new constraints. The active rates that minimize (26) are:

$$R_{bi} = \frac{R}{\sum_{j=1}^3 \alpha_j \xi_j} + \frac{1}{a} ln \left(\frac{\frac{\varepsilon_i^2 G_b \widetilde{\sigma}_{bi}^2 \psi_{bk}}{\alpha_i}}{\prod_{k=1}^3 \left(\frac{GM_k^{Act} \varepsilon_k^2 \Psi_k^{Act}}{\alpha_k} \right)^{\frac{\alpha_k \xi_k}{\sum_{j=1}^3 \alpha_j \xi_j}}} \right), \qquad (27)$$

where $\Psi_k^{Act} \triangleq \prod_{b \in Act_k} \left(\psi_{bk} \right)^{\frac{\eta_b}{\xi_k}}$ and GM_k^{Act} is as given in (5).

3. PERCEPTUALLY OPTIMIZED COMPRESSION

We present a general approach to color image compression using a subband transform with perceptual optimization of the CCT and the subband rates allocation. The approach consists of the stages described in the beginning of Section 1.1. The main difference here is that in the pre-processing stage the perceptually optimal CCT transform is applied to the color components and in the quantization stage the perceptually optimal rates allocation is used. We demonstrate the approach both for the DCT (Discrete Cosine Transform) in Subsection 3.1 and the DWT (Discrete Wavelet Transform) in Subsection 3.2.

3.1 The proposed DCT-based compression algorithm

Since the DCT is a subband transform, the Rate-Distortion theory of Section 2 is readily suitable for such a case. To find the DCT visual weights we use the CSF (Contrast Sensitivity Function) curves of the human visual system for the YCbCr color space, that can be found, for example, in [12]. To convert the cpd (cycle per degree) units of these graphs into spatial frequency units for the DCT, the equations proposed in [13] are adopted. We consider for example 256×256 images displayed as $64mm \times 64mm$ on a display with dot pitch of 0.25mm. The viewing distance is assumed to be four times the image height [14], i.e., in this example 25cm. The stages of the proposed algorithm are as follows.

1. Find the optimal CCT $\mathbf{M}$ by minimizing (25).
2. Apply the CCT $\mathbf{M}$ to the RGB color components of the image to obtain the new color components C_1, C_2, C_3.
3. Apply the DCT block transform to each color component $C_i, i \in \{1, 2, 3\}$.
4. Calculate the optimal rates according to (27) substituting there the used CCT matrix and the variances of the DCT subbands. To find the active subbands, the algorithm presented in [11] could be used.
5. Quantize the DCT coefficients using a uniform scalar quantizer in each subband. The (optimal) quantization steps are found using an iterative algorithm [5].
6. Use post-quantization coding similar to the one used in JPEG. Adaptive Huffman coding is employed and the codes are sent with the image data. This stage is lossless and does not affect the image distortion.

It is of interest to compare the performance of this algorithm to other DCT-based compression algorithms, such as the MSE optimized algorithm proposed in [11] and to JPEG. A comparison for several images is summarized in Table 1. We consider here the above algorithm with WMSE R-D optimization of the rates allocation and CCT as well as another version of the algorithm that uses optimal rates in the YCbCr color space. The PSPNR (Peak Signal to Perceptual Noise Ratio) measure used here is:

$$PSPNR \triangleq 10 log_{10} \frac{255^2}{WMSE} , \qquad (28)$$

where $WMSE$ for each color component is calculated in the DWT domain in the YCbCr color space according to the visual weights suggested in [12]. Then the average PSPNR of the 3 color components is taken.

It can be concluded from the table that the WMSE optimized algorithm with the optimal CCT achieves the highest

		PSPNR				CR
	Image	**WMSE Alg. YCbCr**	**WMSE Alg. Opt.**	**MSE Alg.**	**JPEG**	
	Lena	39.425	40.622	38.863	37.566	31.63
	Peppers	39.633	39.631	38.140	36.567	29.65
	Baboon	42.016	42.535	39.204	36.101	13.63
	Cat	41.336	43.082	41.305	39.926	18.53
	Sails	41.010	42.908	39.663	37.550	13.07
	Monarch	39.796	40.152	38.692	37.521	23.24
	Goldhill	42.933	43.371	41.882	40.607	11.05
	Mean	**40.878**	**41.757**	**39.678**	**37.977**	

Table 1: Perceptually-based results (PSPNR) for (from left to right): The DCT-based WMSE optimized algorithm in the YCbCr domain; The same algorithm with optimal CCT; The MSE optimized algorithm; JPEG. The compression ratio for each image (CR) is shown in the right column.

PSPNR, which is a gain of 2.08dB compared to the MSE optimized algorithm, and 3.78dB above JPEG on average. The use of the optimal CCT in the WMSE sense increases the performance by 0.88dB on average when perceptually optimal rates are employed. Another comparison of interest is of the standard or *objective* distortions of the algorithms, i.e., the PSNR (Peak Signal to Noise Ratio). The average performance for the same images of Table 1 is presented in Table 2. As expected, the MSE optimized algorithm is superior here, but what is perhaps less intuitive is the fact that the use of the optimal CCT for the algorithm based on the WMSE optimization slightly decreases the PSNR. Despite this, both WMSE algorithms outperform JPEG with a gain of 0.87dB in the PSNR without using the optimal CCT and even higher (1.08dB) with the optimal CCT. We conclude this section by presenting a visual comparison of the algorithms in Fig. 1 for the Baboon image. It can be seen that the WMSE algorithm provides results that are perceptually superior to the MSE algorithm. Yet both algorithms outperform JPEG.

Mean PSNR			
WMSE Alg. YCbCr	**WMSE Alg. Opt.**	**MSE Alg.**	**JPEG**
30.002	**29.800**	**30.854**	**28.927**

Table 2: Standard (PSNR) mean results for (from left to right): The DCT-based WMSE algorithm in the YCbCr domain; The WMSE algorithm using the optimal CCT; The MSE optimized algorithm; JPEG.
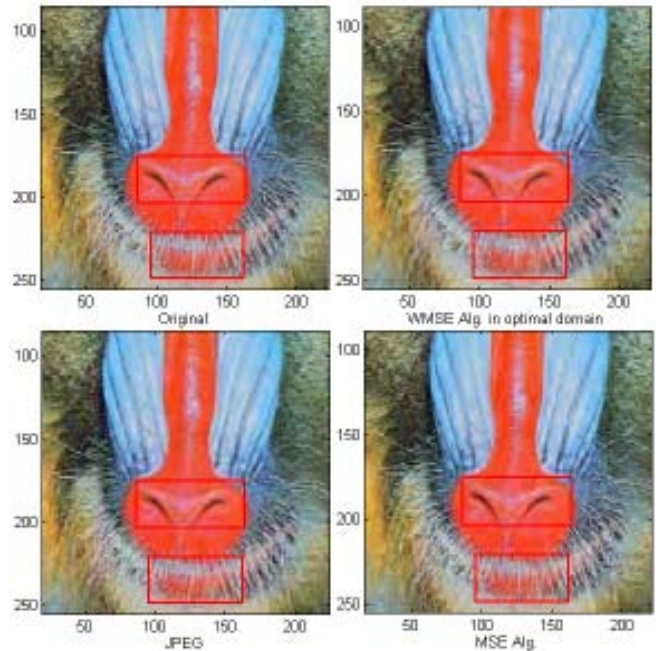

Figure 1: Compression results for the Baboon (zoomed in) at 0.881 bpp. Original image (top left); Image compressed by the WMSE optimized algorithm (top right, PSPNR=36.93dB); Image compressed by JPEG (bottom left, PSPNR=33.59dB); Image compressed by the MSE optimized algorithm (bottom right, PSPNR=35.56dB). As expected, the WMSE algorithm outperforms the other methods, especially in the marked areas.

3.2 The proposed DWT-based compression algorithm

We consider here the DWT decomposition by Daubechies 9/7 filter bank. No tiling [3] is used. The choice of the visual weights is according to [12]. The stages of the proposed algorithm are:

1. Find the optimal CCT $\mathbf{M}$ by minimizing (25).
2. Apply the CCT $\mathbf{M}$ to the RGB color components of the image to obtain the new color components C_1, C_2, C_3.
3. Apply the DWT tree decomposition up to the required depth of the tree (3, 4, 5 or higher according to image size) to each color component C_i, $i \in \{1,2,3\}$.
4. Calculate the optimal rates according to (27) substituting there the CCT matrix and the variances, the sample rates and energy gains of the DWT subbands. The determination of the active subbands is the same as for the DCT-based algorithm of Section 3.1.
5. Quantize the DWT coefficients by a uniform quantizer with a central dead-zone in each subband. Use optimal quantization steps.
6. Use the post-quantization coding of the EZW algorithm [2] on the quantized subband coefficients. This stage is lossless and includes bit plane coding with the use of zero trees. The bit plane coding is split into two passes (dominant and subordinate) and a separate arithmetic coder is employed for each pass.

It is of interest to compare the proposed algorithm to JPEG2000. We consider the JPEG2000 implementation by the JasPer software package [15] and another version of the implementation with fixed visual weighting at subband level
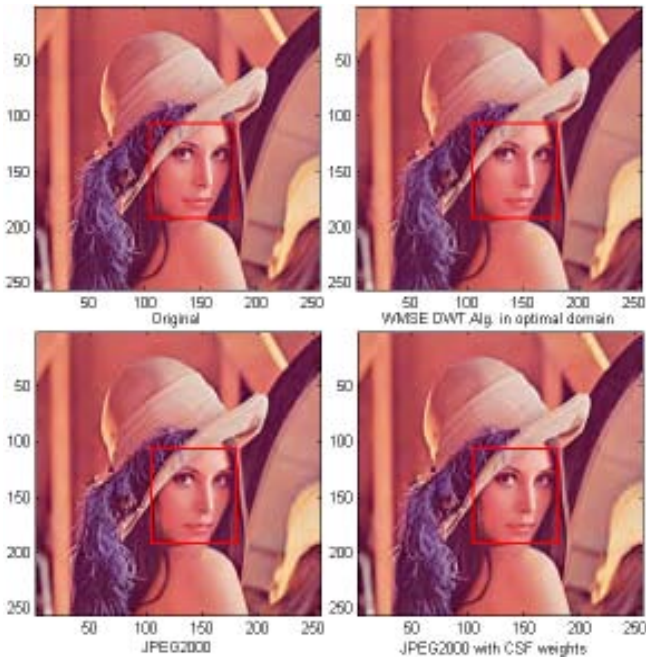
Figure 2: Compression results for Lena at 0.519 bpp. Original image (top left); Image compressed by the DWT-based WMSE optimized algorithm (top right, PSPNR=19.66dB); Image compressed by JPEG2000 (bottom left, PSPNR=19.10dB); Image compressed by JPEG2000 with CSF weights (bottom right, PSPNR=19.23dB).

using the CSF weights of [12]. The visual results for the Lena image are shown in Fig. 2. The PSNR results here are 30.66dB for the proposed WMSE optimized algorithm, 29.96dB for JPEG2000 and 29.73dB for JPEG2000 with CSF weights. We conclude that the usage of CSF weights decreases the PSNR of JPEG2000, but slightly improves its visual performance. Also the proposed algorithm produces an image that perceptually outperforms JPEG2000.

4. SUMMARY

An optimized perceptually-based model for the Rate-Distortion function of color subband coders has been introduced and derived. The new model approximates the WMSE distortion of an image in a given color space, such as YCbCr. This distortion is then minimized to achieve perceptual optimization of the compression. Based on the Rate-Distortion model, new algorithms have been introduced consisting of a pre-processing stage of applying a CCT, followed by a subband transform, quantization and lossless post-processing. The algorithms optimize the CCT in the pre-processing stage of the compression and the quantization tables used in the coding stage with respect to WMSE. The DCT-based algorithm, outperforms both JPEG and the corresponding MSE optimized algorithm. It has been demonstrated that even when a relatively basic algorithm is used in the post-processing stage (introduced for EZW), superior results are achieved by the DWT-based algorithm compared to other algorithms based on the DWT, such as JPEG2000. This holds even when the same WMSE distortion is used in both the JPEG2000 and the proposed algorithm.

REFERENCES

[1] G. K. Wallace, "The JPEG Still Picture Compression Standard", *IEEE Trans. on Consumer Electronics* , vol. 38, pp. xviii–xxxiv, 1992.

[2] J. M. Shapiro, "Embedded Image Coding Using Zerotrees of Wavelet Coefficients", *IEEE Trans. on Signal Processing*, vol. 41, no. 12, pp 3345–3462, 1993.

[3] M. Rabbani and R. Joshi, "An overview of the JPEG 2000 still image compression standard", *Signal Processing: Image Communication*, vol. 17, no. 1, pp 3–48, 2002.

[4] JPEG 2000 Part I: Final Draft International Standard (ISO/IECFDIS15444-1), ISO/IEC JTC1/SC29/WG1 N1855, Aug. 2000.

[5] E. Gershikov and M. Porat, "A Rate-Distortion Approach to Optimal Color Image Compression", in *Proc. EUSIPCO 2006*, Florence, Italy, September 4-8, 2006.

[6] J. Mannos and D. Sakrison, "The effects of a visual fidelity criterion of the encoding of images", *IEEE Trans. on Inform. Theory*, vol. 20, no. 4, pp. 525–536, Jul. 1974.

[7] O. D. Faugeras, "Digital Color Image Processing Within the Framework of a Human Visual Model", *IEEE Trans. on Acoustics Speech and Signal Processing*, vol. ASSP-27, no. 4, pp. 380–393, 1979.

[8] H. Kotera and K. Kanamori, "A Novel Coding Algorithm for Representing Full Color Images by a Single Color Image", *J. Imaging Technology*, vol. 16, pp. 142–152, Aug. 1990.

[9] J. O. Limb and C.B. Rubinstein, "Statistical Dependence Between Components of A Differentially Quantized Color Signal", *IEEE Trans. on Comm.*, vol. Com-20, pp. 890–899, Oct. 1971.

[10] E. Gershikov, E. Lavi-Burlak and M. Porat, "Correlation-Based Approach to Color Image Compression", in press, *Signal Processing: Image Communication*.

[11] E. Gershikov and M. Porat, "On Color Transforms and Bit Allocation for Optimal Subband Image Compression", *Signal Processing: Image Communication*, vol. 22, no. 1, pp 1–18, Jan. 2007.

[12] D. S. Taubman and M. W. Marcellin, *JPEG2000: image compression, fundamentals, standards and practice*, Kluwer Academic Publishers, 2002.

[13] C. Y. Wang, S. M. Lee and L. W. Chang, "Designing JPEG quantization tables based on human visual system", *Signal Processing: Image Communication*, vol. 16, no. 5, pp. 501–506, 2001.

[14] K. N. Ngan, K. S. Leong and H. Singh, "Adaptive cosine transform coding of images in perceptual domain", *IEEE Trans. on Acoustics, Speech, and Signal Processing*, vol. 37, no. 11, pp. 1743–1750, Nov. 1989.

[15] http://www.ece.uvic.ca/ mdadams/jasper