On a Probabilistic Approach to Rate Control for Optimal Color Image Compression and Video Transmission

Evgeny Gershikov
Department of Electrical Engineering, Braude Academic College of Engineering,
Karmiel 21982, Israel

Abstract:
Based on a recently introduced Rate-Distortion model for color image compression, optimal color coding and bit allocation are derived. We show that this Rate-Distortion model in conjunction with the probability distribution of subband coefficients can be used to develop an efficient algorithm for coding color images and video sequences. We demonstrate this approach for subband coding using the Discrete Cosine Transform (DCT) and a Laplacian distribution as the probability model. We show how the model can be used for rate-control, applicable to still images and to controlling the bit-rate or bandwidth of video transmission. Visual and quantitative results are presented and discussed to support the efficiency of our algorithms, which outperform other available compression systems.

Keywords: *color image coding, subband transforms, rate-distortion model, discrete cosine transform, Laplacian distribution, rate-control, video coding*

I. INTRODUCTION

Color image and video compression has become a major task in today's communication environment. Usually color images are represented by the three RGB color components, which are highly correlated [4], [7], [10], [15], [22]. Naturally, it is a naive approach to compress each color component separately. To improve the information distribution in the image data, usually a color components transform (CCT) is used. The RGB to YUV transform is employed for example in JPEG [20] and JPEG2000 [13], while the Karhunen Loeve transform (KLT) is used in [5], [8] and [21]. Nevertheless, these transforms are presently used arbitrarily since no optimization process has been proposed so far for color image compression. Part of the reason has been the lack of a model for color images and their Rate-Distortion (R-D) curve. Recently, such a model has been introduced for the analysis of color compression and its optimization [3]. The model has been proposed in the context of the widely used subband transform coders. Based on the model, a color compression algorithm has been presented, outperforming commonly used algorithms. In this work we present an improved compression algorithm based on the new Rate-Distortion theory and on a probability model for the distribution of the subband transform coefficients. We also present an application of the R-D model for rate-control of the compression. This application can be used to achieve a certain compression ratio or target rate. This approach to image compression is applicable to both still and video coding.

The structure of the work is as follows. In Subsections 1.1 and 1.2 we briefly review subband transforms and the Rate-Distortion theory of subband transform coders. In Section 2 we present the new compression algorithm for color images and compare its performance to that of presently available algorithms including [3]. In Section 3 the new rate-control algorithm is presented and its performance is measured for still images, and in Section 4 the algorithm is considered for video sequences. Finally, conclusions and a summary are given in Section 5.

1.1. Subband transforms - definitions

Subband transform coding is an efficient approach to image compression. Fig. 1 presents a filter bank interpretation of the general tree structured subband transform. The input signal $x[n]$ is decomposed by passing through a set of m analysis filters and down-sampling by a factor m. Then its low frequencies subband $y_0^{(1)}[n]$ is decomposed by the same filters and so on in an iterative fashion until depth D of the tree is reached. The signal can be reconstructed iteratively as shown in Fig. 2 by up-sampling the outputs $y_k^{(d)}[n]$ $(0 \le k \le m-1)$ of the analysis filters at level d by a factor of m, filtering them through a synthesis filter-

bank and summing up the results to obtain the low pass subband at level $d-1$ ($y_0^{(d-1)}[n]$). The original signal $x[n]$ can be considered as the subband $y_0^{(0)}$ in this context, i.e., it is obtained using a reconstruction algorithm on the subbands at level 1. Compression can be achieved by quantization of the subband components and possibly omission of the less significant subbands.
The DCT [14], the Discrete Wavelet Transform (DWT) [11] and filter-banks used for audio coding (e.g. [16]) readily fit into this typical structure.

1.2. The Rate-Distortion model

A brief presentation of the theory developed in [3] is summarized here. Given a color image in the RGB domain, we denote each pixel by a 3x1 vector $\boldsymbol{x} = [R\ G\ B]^T$. The RGB correlations are usually high for natural images [1], [4], [7], [10] and hence, a preprocessing stage of a color components transform (CCT) is usually performed prior to coding. Assuming that a CCT is applied to an image, denoted by a 3×3 matrix $\boldsymbol{M}$, we obtain for each pixel a new vector of 3 components $C1, C2, C3$, denoted $\tilde{\boldsymbol{x}} = [C1\ C2\ C3]^T$. Thus $\tilde{\boldsymbol{x}}$ is related to $\boldsymbol{x}$ by:

$$\tilde{\boldsymbol{x}} = \boldsymbol{M}\boldsymbol{x}. \qquad (1)$$

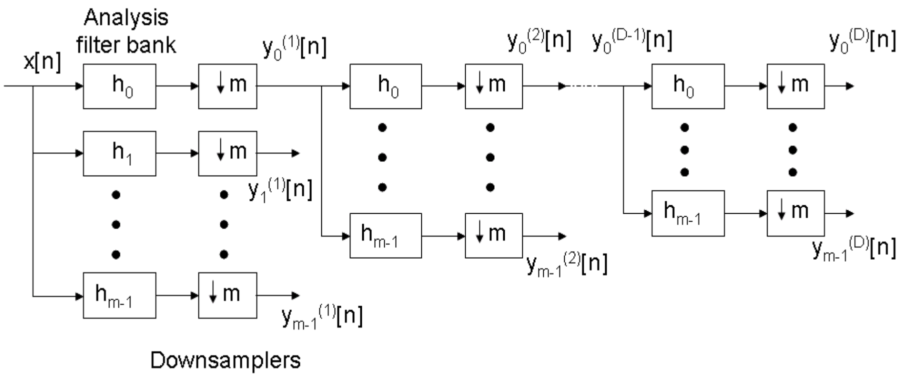


Figure 1. Tree structured subband transform: Analysis.

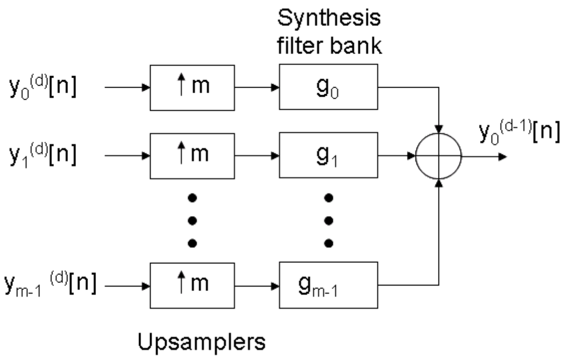


Figure 2. Tree structured subband transform: Synthesis.

Each component in the C1, C2, C3 color space is subband transformed, quantized and its samples are independently encoded (e.g., entropy coded). This description corresponds to typical image compression algorithms such as JPEG [20] and JPEG 2000 [13], when applied to a color image up to and including the quantization stage. Denote by $x_{\boldsymbol{rec}}$ the reconstructed image in the RGB domain, the error covariance matrix in the RGB domain $\boldsymbol{Er}$ is given by

$$\boldsymbol{Er} = E[(\boldsymbol{x} - x_{\boldsymbol{rec}})(\boldsymbol{x} - x_{\boldsymbol{rec}})^T]. \qquad (2)$$

Assume that we have B subbands in the transform for each color component, indexed by $b \in [0, B-1]$ and $i \in \{1,2,3\}$ is the index of the color component. Then the average MSE (Mean Square Error) between the original and reconstructed images in the RGB domain is:

$$MSE = \frac{1}{3} trace(\boldsymbol{Er}) = \frac{1}{3} \sum_{i=1}^{3} \sum_{b=0}^{B-1} \eta_b\, G_b\, \sigma_{bi}^2 \varepsilon_i^2 e^{-aR_{bi}} ((\boldsymbol{M}\boldsymbol{M}^T)^{-1})_{ii},$$

(3)

where R_{bi} stands for the rate allocated for the subband b of color component i and σ_{bi}^2 is this subband's variance. η_b is the sample rate of subband b, meaning the ratio of the number of coefficients in this subband and the total number of coefficients in each component. G_b is the energy gain of subband b equal to the squared norm of the subband's synthesis vectors [19]. ε_i^2 is a constant dependent on the distribution of the subband transform coefficients in each color component and, finally, $a \triangleq 2ln2$. Considering the optimization problem of minimizing (3) under the constraint $\sum_{i=1}^{3}\sum_{b=0}^{B-1}\eta_b R_{bi} = R$ for some total image rate R and using Lagrange multipliers method, the Lagrangian to be minimized is:

$$L(\{R_{bi}\}, \boldsymbol{M}, \lambda) = \frac{1}{3}\sum_{i=1}^{3}\sum_{b=0}^{B-1}\eta_b\, G_b \sigma_{bi}^2 \varepsilon_i^2 e^{-aR_{bi}}((\boldsymbol{M}\boldsymbol{M}^T)^{-1})_{ii} + \lambda\left(\sum_{i=1}^{3}\sum_{b=0}^{B-1}\eta_b\, R_{bi} - R\right), \tag{4}$$

where λ is the Lagrange multiplier. By minimizing (4), one can derive the optimal subband rates R_{bi}^* and the target function for the optimal CCT. The optimal solution for this target function is image adaptive. A sub-optimal solution is the DCT as a CCT [2], as used in this work. It is of interest to note that this result (without proof, however) was also found to be most efficient in [6].

In practice, some coding systems such as JPEG perform down-sampling on part of the color components prior to coding. For such systems the down-sampling can be taken into account by introducing down-sampling factors α_i, so that the global rate constraint for the image is

$$\sum_{i=1}^{3}\alpha_i\sum_{b=0}^{B-1}\eta_b\, R_{bi} = R, \tag{5}$$

where the down-sampling could be, for example, by a factor of 2 horizontally and vertically and then

$$\alpha_i = \begin{cases} 1 & \text{full component} \\ 0.25 & \text{down-sampled component} \end{cases}.$$

We thus wish to minimize the MSE of (3) under the rate constraint of (5). Additional constraints are the non-negativity of the rates: $R_{bi} \geq 0$. Using the Lagrange multipliers method, we have to minimize:

$$L(\{R_{bi}\}, \boldsymbol{M}, \lambda, \{\mu_{bi}\}) = \frac{1}{3}\sum_{i=1}^{3}\sum_{b=0}^{B-1}\eta_b\, G_b \sigma_{bi}^2 \varepsilon_i^2 e^{-aR_{bi}} \cdot ((\boldsymbol{M}\boldsymbol{M}^T)^{-1})_{ii} + \lambda\left(\sum_{i=1}^{3}\alpha_i\sum_{b=0}^{B-1}\eta_b\, R_{bi} - R\right) - \sum_{i=1}^{3}\sum_{b=0}^{B-1}\mu_{bi} R_{bi}, \tag{6}$$

where λ and μ_{bi} are the Lagrange multipliers for the rate constraint and the non-negativity constraints, respectively.

Minimizing the Lagrangian $L()$ for the rates R_{bi} requires knowing the rates that are positive and those that are zero. We denote by Act_i the set of all the active subbands in the color component i, that is, those subbands with positive rates:

$$Act_i \triangleq \{b \in [0, B-1] \quad | \quad R_{bi} > 0\}. \tag{7}$$

We also define the following:

$$\xi_i \triangleq \sum_{b\in Act_i}\eta_b, \quad GMA_i \triangleq \prod_{b\in Act_i}(G_b\sigma_{bi}^2)^{\frac{\eta_b}{\xi_i}}, \tag{8}$$

i.e., the relative part of the coefficients in the active subbands from the total signal length (ξ_i) and the weighted geometric mean of their variances (corrected by the energy gains G_b) GMA_i. It can be shown that the solution for $b \in Act_i$ becomes:

$$R_{bi} = \frac{1}{a} ln \left[\frac{\frac{\varepsilon_i^2 G_b \sigma_{bi}^2 ((MM^T)^{-1})_{ii}}{\alpha_i}}{\prod_{k=1}^{3} \left(\frac{((MM^T)^{-1})_{kk} \varepsilon_k^2 GMA_k}{\alpha_k} \right)^{\frac{\alpha_k \xi_k}{\sum_{j=1}^{3} \alpha_j \xi_i}}} \right] + \frac{R}{\sum_{j=1}^{3} \alpha_j \xi_j} \quad (9)$$

The determination of the active subbands can be done according to the algorithm given in [3].

II. IMPROVING THE COMPRESSION ALGORITHM

Based on the Rate-Distortion theory, a DCT-based algorithm for color image compression is proposed. This algorithm employs the DCT as the color components transform and utilizes the optimal rates expression of (9). It consists of the following stages:

1. Apply the CCT (DCT) to the RGB color components of a given image to obtain new color components $C1, C2, C3$.
2. Decimate the 2 color components with minimal energy (variance) by a factor of 2 in each direction.
3. Apply the two-dimensional block DCT to each color component Ci.
4. Quantize each subband of each color component independently using uniform scalar quantizers. The quantization step-sizes are chosen so that optimal subband rates are achieved according to Subsection 1.2. The rates are calculated using the Laplacian distribution model for the coefficients of the DCT subband transform.
5. Apply lossless coding to the quantized DCT coefficients. Coding techniques similar to JPEG [20] can be used: differential Huffman coding for the DC coefficients and zigzag scan, run-length coding and Huffman coding (combined with variable-length integer codes) for the AC coefficients.

2.1. The Laplacian distribution

We say that a stochastic variable X has Laplacian distribution if its probability distribution function is

$$p_X(x) = \frac{\mu}{2} e^{-\mu|x|} \quad (10)$$

for some positive constant μ. For such a variable, we can derive the variance σ_X^2 and entropy $h(X)$ as functions of μ:

$$\sigma_X^2 = \frac{2}{\mu^2}, \quad h(X) = \log_2 \frac{2e}{\mu}. \quad (11)$$

Thus the following relationship holds:

$$2^{h(X)} = \sqrt{2} e \sigma_X. \quad (12)$$

Assume that X is quantized by a uniform scalar quantizer with a step size Δ to the discrete variable $\hat{X}$. The probability distribution of $\hat{X}$ is

$$\begin{aligned} P_n &\triangleq Prob(\hat{X} = n) = Prob((n - 0.5)\Delta \le X \le (n + 0.5)\Delta) \\ &= \int_{(n-0.5)\Delta}^{(n+0.5)\Delta} p_X(x) dx = \int_{(n-0.5)\Delta}^{(n+0.5)\Delta} \frac{\mu}{2} e^{-\mu|x|} dx. \end{aligned} \quad (13)$$

and hence:

$$P_n = \begin{cases} 1 - e^{-0.5\mu\Delta}, & n = 0 \\ \frac{1}{2} e^{-\mu n\Delta} \left(e^{0.5\mu\Delta} - e^{-0.5\mu\Delta} \right), & n > 0 \end{cases} \quad (14)$$

$$P_{-n} = P_n$$

Defining $k \triangleq e^{0.5\mu\Delta} - e^{-0.5\mu\Delta}$ and using (9), we can derive the entropy of $\hat{X}$:

$$\begin{aligned} H(\hat{X}) &= -\sum_{n=-\infty}^{\infty} P_n \log_2(P_n) = -P_0 \log_2(P_0) - 2\sum_{n=1}^{\infty} P_n \log_2(P_n) \\ &= -\left(1 - e^{-0.5\mu\Delta}\right) \log_2\left(1 - e^{-0.5\mu\Delta}\right) - e^{-0.5\mu\Delta} \left(\log_2 k - 1 \right) + \frac{\mu\Delta}{k} \log_2(e). \end{aligned} \quad (15)$$

Note that the μ parameter in (15) can be expressed by the standard deviation of X using (11) as $\mu = \frac{\sqrt{2}}{\sigma_X}$.

2.2. DCT coefficients distribution

When examining the coefficients' distribution of the 2D block DCT, it can be concluded that the distribution of all the subbands, except for the DC subband, can be modeled by the Laplacian distribution [9]. Using this model we can benefit in 2 ways:

1. An algorithm for the calculation of the quantization step sizes for optimal rates can be introduced based on the approximate connection [19]:

$$\Delta_{bi} = 2^{h_{bi} - R_{bi}}, \tag{16}$$

2. where h_{bi} is the entropy of subband b of the color component i prior to quantization and Δ_{bi} is its quantization step. The quantization steps initialization can be chosen according to (11) directly by substituting the right hand side of (7) for $2^{h(X)}$, i.e., we get an expression that can be easily calculated:

$$(\Delta_{bi})_0 = \sqrt{2}e\sigma_{bi}2^{-R^*_{bi}}, \tag{17}$$

3. where $(\Delta_{bi})_0$ is the initial quantization step and R^*_{bi} is the optimal rate of subband b of color component i. Following the initialization, the optimal quantization steps can be calculated iteratively using the update rule:

$$\Delta^{new}_{bi} = \Delta_{bi}2^{-(R^*_{bi} - R_{bi})}, \tag{18}$$

4. based on (11), where Δ_{bi} and R_{bi} are the current quantization steps and rates respectively, and Δ^{new}_{bi} are the updated steps. This update rule can be repeated until the optimal rates R^*_{bi} are sufficiently close, i.e., $E(|R^*_{bi} - R_{bi}|) < \varepsilon$ for some small constant ε. Note that the rates R_{bi} are measured by the entropies of the subbands. Those can be calculated as follows.
5. The entropies of the quantized DCT coefficients can be approximately calculated according to (10) without the need to calculate the subband histograms. We use the Laplacian distribution assumption also for the DC subband, although its distribution is usually not Laplacian. The use of the approximated entropies reduces the number of the calculations required, thus reducing the run time of the algorithm - as discussed in Subsection 2.4.

To assess the performance of the proposed approach, simulation results of the new algorithm with estimated rates (entropies) according to (10) are presented and compared to JPEG.

2.3. Simulations and Comparison

2.3.1 Comparison to JPEG

Similar to the PSNR (Peak signal to Noise Ratio): $PSNR = 10log_{10}\left(\frac{255^2}{MSE}\right)$, we use the PSPNR (Peak Signal to Perceptible Noise Ratio), defined as

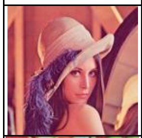
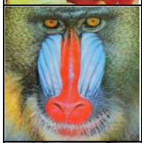
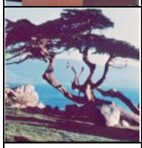
$$PSPNR = 10\log_{10}\frac{255^2}{WMSE}, \tag{19}$$

where $WMSE$ (Weighted Mean Square Error) for each color component, is calculated as:

$$WMSE = \sum_{b=0}^{B-1} \eta_b W_b G_b d_b. \tag{20}$$

Here W_b denotes the visual perception weight of subband b and d_b is its MSE distortion. We have taken the WMSE suggested in [19] for JPEG2000, so that the subbands in (20) are of the Discrete Wavelet Transform (DWT). We consider 256x256 size images displayed on a screen as 12cm x 12cm size images and a viewing distance of approximately 50 cm. The PSPNR used here is the mean PSPNR of the three color components. Simulation results for several images are summarized in Table 1. It can be seen that the new algorithm outperforms JPEG by 1.5dB PSNR and 1.85dB PSPNR on average. These results are not limited just to the compression ratios of Table 1. Fig. 3 shows the algorithm's mean performance gain for a range of compression ratios corresponding to the major range of PSNRs.

Table 1. PSNR and PSPNR results for the DCT-based compression algorithm with estimated rates (New Alg.) and JPEG at the same compression ratio (CR).

		PSNR		PSPNR		
	Image	**New Alg.**	**JPEG**	**New Alg.**	**JPEG**	**CR**
	Lena	30.030	29.486	39.048	37.192	44.84
	Peppers	29.971	28.277	37.818	35.439	33.77
	Baboon	30.024	26.306	38.273	36.023	16.62
	Fruit	30.024	29.268	39.048	36.974	46.53
	Girl	29.987	28.719	38.407	36.871	52.56
	House	30.006	28.573	38.788	37.073	54.36
	Tree	29.987	28.798	39.210	38.072	14.19
	Mean	**30.004**	**28.490**	**38.656**	**36.806**	

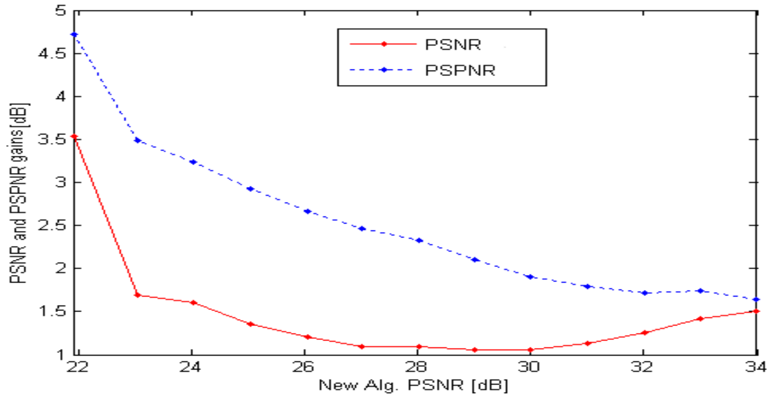


Figure 3. PSNR and PSPNR gains of the new algorithm (New Alg.) with estimated rates on JPEG for various values of PSNR.

Visual results for the House and the Tree images are shown in Fig. 4. It can be seen that JPEG introduces color artifacts in both images. These are significantly less visible in our new algorithm at the same rate. Furthermore, quantitatively, there is a gain of 1.25dB PSNR and 1.6dB PSPNR by the new algorithm for the Tree image. For the House image the gain is even larger: 1.55dB PSNR and 2.1dB PSPNR.

2.3.2. Comparison to the algorithm in [3]

Another comparison is to the algorithm in [3]. The compression results for several images are displayed in Table 2. It can be seen that the algorithm proposed in this work always outperforms the algorithm in [3] with a mean gain of 0.522dB PSNR and 0.278dB PSPNR. The gain for an individual image can be as much as 2dB PSNR as for Baboon.

The new algorithm is superior to the one in [3] also with respect to the execution times. This is elaborated in Subsection 2.4.

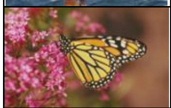
2.4. Run time of the algorithm

The run-times for the new algorithm, the algorithm in [3] as well as JPEG are shown in Table 3 for several image sizes. As expected, the use of estimated rates improves the run-time of the compression algorithm, especially for smaller images (15.7% improvement for 256x256 images, 3.5% improvement for 512x768 images). The new algorithm is also comparable with JPEG (14% slower for 256x256 images, 6.9% slower for 512x768 images).

III. RATE CONTROL

Having introduced a DCT-based color compression method that outperforms the baseline JPEG algorithm, it should be noted that this application does not target rate control. Since we can derive the expression for the optimal rates for some given image rate R as in Equation (9) and we can calculate the quantization tables to achieve these rates, an application for rate control can be designed. Next we describe such an application, based on the DCT block transform, aimed to achieve a given rate with performance higher or equal to JPEG. Although this algorithm employs the DCT as a CCT as well and designs the quantization tables for optimal rates, it codes the subbands differently: each subband (DC or AC) is coded independently using the size/value representation of the baseline JPEG [20]. The sizes are coded using adaptive arithmetic coding and

Table 2. PSNR and PSPNR results for the new algorithm (New Alg.) and the algorithm in [3] at the same compression ratio (CR).

		PSNR		PSPNR		
	Image	**New Alg.**	**Alg. in [3]**	**New Alg.**	**Alg. in [3]**	**CR**
	Lena	30.011	29.765	39.038	38.671	45.07
	Peppers	29.971	29.770	37.818	37.489	33.77
	Baboon	30.024	28.056	38.273	37.859	16.92
	Cat	30.019	29.172	40.066	39.603	21.97
	Sails	29.990	29.923	39.018	39.012	14.61
	Monarch	29.975	29.721	38.221	37.871	27.08
	Goldhill	29.999	29.928	40.519	40.499	13.23
	Mean	**29.998**	**29.476**	**38.993**	**38.715**	

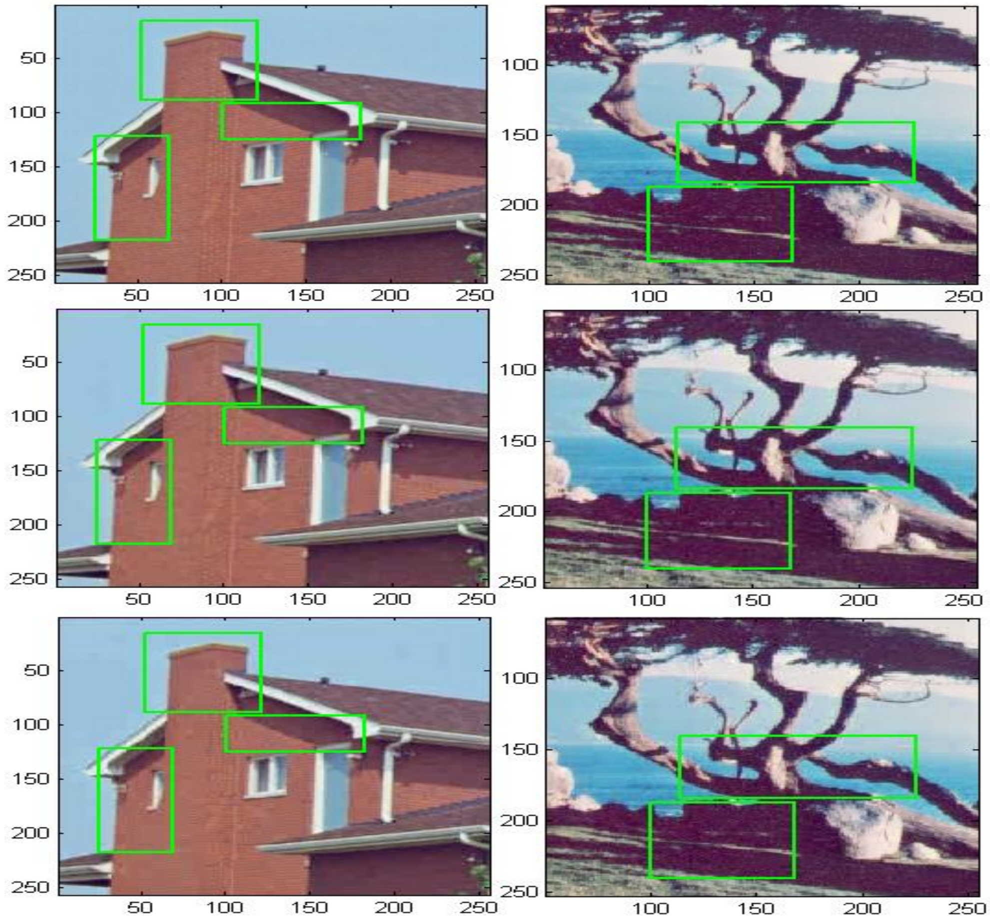


Figure 4. The House and the Tree images - from top to bottom: original, compressed by JPEG and compressed by the new algorithm.
PSNR for the House image: 29.45dB for JPEG and 30.98dB for the new algorithm. PSPNR: 38.01dB and 40.10dB, respectively, at CR=45.15.
PSNR for the Tree image: 28.73dB (JPEG) and 29.98dB (new algorithm). PSPNR: 37.99 and 39.56dB, respectively, at CR=15.36.

Table 3. Run-times in sec. for the algorithm in [3], the new algorithm and JPEG.

Image Size	The alg. in [3]	New Alg.	JPEG
256×256	1.53	1.29	1.13
512×768	7.73	7.46	6.98

the values as variable length integer (VLI) codes similar to JPEG.For the decoder to be able to reconstruct the sizes, the numbers of appearances (or counts) of each size are sent with the compressed image data to the decoder. The number of bits for these counts as well as the number of bits for the quantization tables (adaptively designed for the image) have to be taken into account. This bits overhead can be considerably reduced, especially at high compression ratios by sending the quantization steps and counts only for the active subbands (with optimal rates $R_{bi}^{*} > 0$) in addition to bit vectors signifying which subbands are active. This requires, however, calculating the optimal rates as described below.
The stages of the rate control algorithm are:

1. Apply the CCT (here DCT) to the RGB color components of a given image and obtain new color components $C1, C2, C3$.
2. Apply the 2D block-DCT to each color component Ci (using block size of 8×8 if comparison to JPEG is

of interest).

3. Calculate the optimal rates corresponding to the given total rate R (using (9)). Assume here that no bits are sent for the quantization steps and the counts (of the arithmetic coding) to find the rate for the actual compressed image data.
4. Using the optimal rates of the previous stage, find the number of active subbands in each color component and calculate the number of bits needed both for the quantization tables and the counts. Then recalculate the rate for the actual image data and find the optimal subband rates accordingly.
5. Quantize each subband of each color component independently using uniform scalar quantizers with quantization steps designed as in Subsection 2.2.
6. Encode the quantized DCT coefficients in each subband similarly to JPEG's coding of the DC subband, but without using delta modulation. Delta modulation can be used, but for the DC subband only (see below). Each coefficient is split into size and value representation when the sizes are coded arithmetically and contain the information of the number of bits in the coefficient while the following values are coded by VLIs (1's complement). For applications where rate control precision is of primary concern, we allow optional non differential coding of the DC subband of the maximal energy color component (denoted $C1$). Such coding requires more bits, however, it allows achieving the required total rate more accurately, especially at high compression ratios. The decision whether to code the DC subband differentially or not is made according to a comparison between the sum of the real rates R_{b1} allocated to the subbands of $C1$ and the sum of optimal rates R^*_{b1}. The decision can be made according to

$$\begin{cases} \text{If} \quad \sum_{b=0}^{B-1} R_{b1} < \varepsilon \sum_{b=0}^{B-1} R^*_{b1} & \text{use non differential coding} \\ \text{Else} & \text{use differential coding} \end{cases}, \tag{21}$$

where $0 < \varepsilon < 1$ is a constant close to 1, e.g., 0.98.

The image data bitstream at the end of the algorithm consists of the bit vectors signifying the active subbands, the quantization tables, the arithmetic coding counts and the encoded DCT coefficients (in addition to potential headers, etc.). Note that optional down-sampling of some of the color components (usually $C2$ and $C3$) can be performed between Stages 1 and 2.

3.1 Rate-control Results

When considering, for example, images of size 256x256 and a target compression ratio (CR) of 30 (or total rate of 0.8bpp), the results of the algorithm are shown in Table 4. Here the accuracy of the algorithm is measured and a comparison of its performance to JPEG is given in the PSNR and PSPNR sense. The rate-control algorithm succeeds to achieve the desired compression ratio with a relative absolute error of 1.464% and mean of 29.995. The standard deviation is 0.524 or 1.747% of the target CR. Its performance is superior to JPEG with a gain of 0.360 dB PSNR and 1.129 dB PSPNR on average.

Note that all the relative measures are relative to the target CR, i.e., the relative absolute error for example is the ratio of L1 norm of the error and the CR value.

3.1.1. Accuracy of the algorithm

When applying the algorithm to the same images for other compression ratios, the results are given graphically in Figs. 5 and 6. It should be noted that the algorithm is more accurate at medium compression ratios: 20-110 where the mean relative error is approximately 1% or less, the mean relative absolute error is in the range 1-2% and the relative standard deviation (STD) is below 2% for compression ratios below 100, and grows up to 2.6% above it. At high compression ratios (above 110) and at low ratios (below 20) the accuracy decreases. Note that for each value of target CR the mean relative error describes the shift of the mean CR relative to the target value (in % of this value) while the relative STD describes the width of the data distribution around the mean CR. Finally, the absolute relative error is a measure of the actual mean error when all errors are absolute value averaged.

The main difficulty at low rates is that there are many active subbands and since the size/value coding of each subband usually achieves slightly greater rates than plain entropy coding, the results are that the more subbands are coded the greater the error in the bit rate. At high rates the number of active subbands becomes small and therefore the decision how the C1 DC subband is coded becomes of greater importance: if the DC subband is coded differentially, the resulting compression will be greater than the target CR since the DC rate will be smaller than the subband's (non-differential) entropy. On the other hand, coding the DC non-differentially results in a greater number of bits than needed for the DC information and often in a compression ratio that is too low. If the algorithm's complexity is of lower priority than the rate-control accuracy, both options for the DC subband coding can be always tested and the more precise one chosen. Then the precision at high rates improves as demonstrated in Table 5 - it is higher in terms of (lower) mean relative absolute error and also the CRs distribution is narrower in terms of standard deviation. Note that the original algorithm usually reduces the complexity (in terms of run-time and storage capacity) by running the differential DC coding and checking the condition in (21).

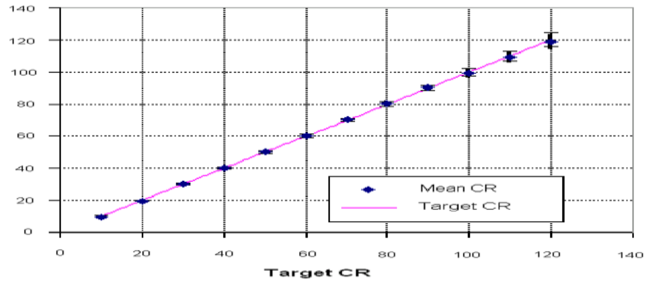


Figure 5. The mean CR for the set of images in Table 4 vs. the target CR. The solid line describes the target CR, while the points describe the achieved mean CR and the error bars are of standard deviation (STD) size.

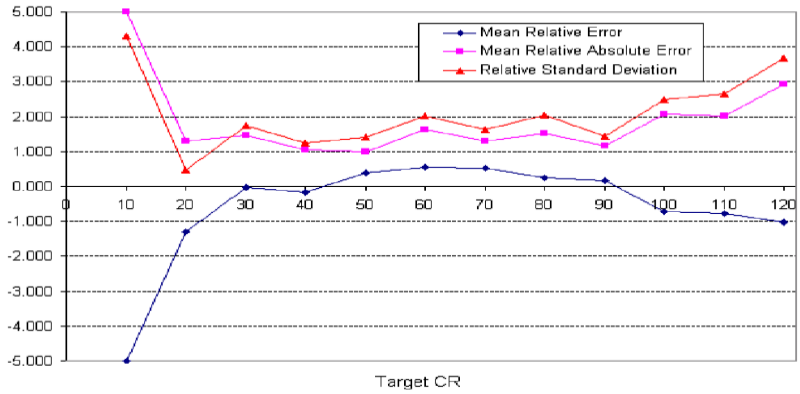


Figure 6. Accuracy measures for the rate-control algorithm: mean relative error (%), mean relative absolute error (%) and relative standard deviation (%).

Table 4. Results for the rate-control algorithm on medium size images for target CR=30. From left to right: The obtained CR, Relative Error, PSNR for the new algorithm, PSNR for JPEG at the same CR and same columns for PSPNR.

	Image	CR	Rel. Err.(%)	PSNR		PSPNR	
				New Alg.	JPEG	New Alg.	JPEG
	Lena	30.586	1.952	31.921	31.925	40.824	39.950
	Peppers	29.785	-0.718	29.982	29.271	37.696	36.515
	Baboon	30.860	2.868	24.553	24.462	35.226	33.766
	Cat	30.291	0.970	27.052	26.881	37.749	36.662
	Sails	29.500	-1.667	26.311	25.595	35.879	33.923

	Tulips	29.482	-1.727	25.826	24.987	33.784	32.431
	Monarch	29.581	-1.395	28.287	28.228	36.494	36.176
	Goldhill	29.876	-0.413	24.716	24.422	36.289	35.487
	Mean	**29.995**	**-0.016**	**27.331**	**26.971**	**36.743**	**35.614**

Table 5. Accuracy measures for the rate-control algorithm at high values of target CR: original algorithm and with both DC coding options tested.

	Original Algorithm			Both DC coding options tested		
Target CR	Mean Rel. Err. (%)	Mean Rel. Abs. Err. (%)	Rel. Std (%)	Mean Rel. Err. (%)	Mean Rel. Abs. Err. (%)	Rel. Std (%)
100	-0.713	2.085	2.483	0.401	1.679	2.149
110	-0.772	2.025	2.663	0.958	1.743	2.191
120	-1.027	2.937	3.665	0.820	2.072	2.410

We should also note that if the application requires a compression ratio not less than a given value, as perhaps can be the case in mobile applications, only differential coding should be used. In any case the algorithm's results are comparable to [18] especially at the middle range of compression ratios.

3.1.2. Performance (PSNR and PSPNR) of the algorithm

Fig. 7 describes the performance gain of the rate-control algorithm vs. JPEG in the PSNR and PSPNR senses. Since the algorithm is based on the optimal rates, it indeed outperforms JPEG. In cases of CR = 10 the reconstructed images for both our algorithm and JPEG are visually identical to the original one (e.g. when the $PSNR > 34dB$) and thus of limited interest.

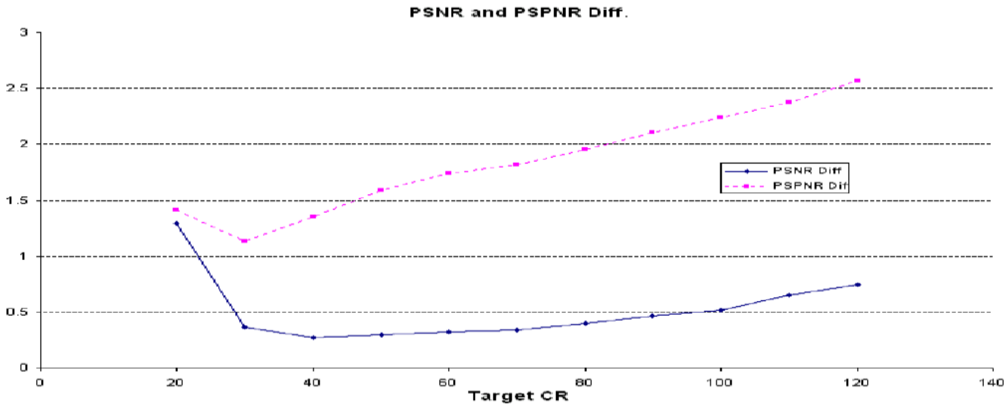


Figure 7. PSNR and PSPNR differences between the rate-control algorithm and JPEG vs. target CR. Along the whole range of values checked the new algorithm outperforms JPEG.

IV. RATE-CONTROL OF VIDEO SEQUENCES

Consider a group of frames (GOF) of a video sequence, to be encoded similarly to the MPEG standard [17]. Here, instead of applying JPEG to the frames or their prediction errors we would like to use the DCT-based compression technique of [3]. We assume for simplicity that the GOF structure is: I,P,P,P,... where I denotes an image coded using intra-frame techniques only and P stands for an image coded using the inter-frame correlation with a previous image. The correlation can be exploited using standard motion estimation. Suppose that the frames are to be coded at a given rate R in bits/sec. We consider k frames of the GOF and assume that these k images are allocated some total number of bits denoted by R_k. Denoting by MSE_j the MSE of the reconstruction of frame j, the average MSE of the frames is simply:

$$MSE = \frac{1}{k}\sum_{j=1}^{k} MSE_j = \frac{1}{3k}\sum_{j=1}^{k}\sum_{i=1}^{3}\sum_{b=0}^{B-1} \eta_b\, G_b\, \varepsilon_i^2 \left(\sigma_{bi}^j\right)^2 e^{-aR_{bi}^j}\left(\left(\boldsymbol{M}\boldsymbol{M}^T\right)^{-1}\right)_{ii}. \quad (22)$$

Here we have used (3) for MSE_j denoting by $\left(\sigma_{bi}^j\right)^2$ and R_{bi}^j the variance and rate, respectively, of subband b of color component i of frame j. If frame j is an I-frame, the variances are simply of its DCT subbands. However, if we consider a P-frame, then the variances are of the DCT subbands of the image of prediction errors. In any case what is important is that the MSE of the reconstruction of the frame is only the result of its DCT-based compression and not the prediction method used (the motion estimation). The prediction technique affects only the magnitude of the variances $\left(\sigma_{bi}^j\right)^2$. Note that $\boldsymbol{M}$ stands here for the CCT as usual (for example, the RGB to YUV transform or the DCT). Substituting $\eta_b = \frac{1}{B}$ and $G_b = 1$ for the DCT subband transform and slightly rewriting (22) we have:

$$MSE = \frac{1}{3kB}\sum_{i=1}^{3} \varepsilon_i^2 \left((\boldsymbol{M}\boldsymbol{M}^T)^{-1}\right)_{ii} \sum_{b=0}^{B-1}\sum_{j=1}^{k} \left(\sigma_{bi}^j\right)^2 e^{-aR_{bi}^j}. \qquad (23)$$

Similarly, the rate constraint $\frac{1}{k}\sum_{j=1}^{k}\sum_{i=1}^{3}\sum_{b=0}^{B-1} \alpha_i\, \eta_b\, R_{bi}^j = R_k$ for the k frames can be written as:

$$\frac{1}{kB}\sum_{i=1}^{3} \alpha_i \sum_{b=0}^{B-1}\sum_{j=1}^{k} R_{bi}^j = R_k, \qquad (24)$$

where here again we have used η_b of the DCT transform. Considering (23) and (24), we conclude that the problem is similar to the case of still images. Assume for a moment that each color component has kB subbands instead of B, and that these subbands have variances $\left(\sigma_{bi}^j\right)^2$ and rates R_{bi}^j ($b \in \{0,1,\ldots,B-1\}$, $j \in \{1,2,\ldots,k\}$). Then finding the optimal rates allocation for these subbands will give us the subband rates allocation for our k frames. Here we can see the significance of the choice of k. If we take a higher value for k we can expect better performance of the rate-control algorithm since we will be able to allocate the bits budget more flexibly. If, for example, one frame is coded with a small number of bits, its 'spare' bits can be allocated to a greater choice of other frames. However, it leads to a more complicated optimization problem since there are $3kB$ subbands to consider. Similarly, a smaller value for k reduces the order of the optimization problem, however, it also reduces the flexibility of the bits allocation.

Using the same algorithm of Section 3 produces the results described in the next subsection.

4.1. Video rate-control results

We consider a CIF (Common Intermediate Format) color video sequence of 352×288 spatial resolution at the frame rate of 25 frames/sec. We would like to encode the video at 1.055 Mbit/sec (compression ratio of 55). Original and reconstructed frames for the video sequence "Hall Monitor" are presented in Fig. 8. We consider here a GOF starting at frame 100 of the sequence and show the first 4 frames in the GOF. The k is chosen to be 5. The rate is 1.073 MBit/sec with an error of 1.73% relative to the desired one. The motion vectors have used 4.17% of the total bits budget.

V. CONCLUSIONS

With the introduction of a Rate-Distortion model for color image compression, it has become possible to optimize subband transform coders both in the preprocessing stage and in the encoding stage itself. The optimization process leads to the use of the DCT as a color component transform in the preprocessing stage (in addition to its role in subband coding) and provides optimal rates allocation for the coding. These rates can be then used to design optimal quantization steps [12]. We have shown that the quantization stage can be further optimized by applying a Laplacian model to the distribution of the DCT coefficients. This model allows for efficient initialization and faster calculation of the optimal steps by estimating the subband entropies, thus reducing the run-time of the algorithm compared to the use of real entropies, as well as improving the compression performance in terms of PSNR and PSPNR. We have also presented an algorithm for optimized image coding with rate-control. This algorithm can be used to achieve a desired compression ratio and for controlling the bit-rate or bandwidth of video transmission. The presented simulations demonstrate that the proposed algorithms outperform JPEG, both visually and quantitatively. Our conclusion is that based on the newly introduced Rate-Distortion model, optimized compression algorithms can be designed with compression results superior to presently available methods.

ACKNOWLEDGEMENT

This research was supported in part by the HASSIP Research Program HPRN-CT-2002-00285 of the European Commission, and by the Ollendorff Minerva Center. Minerva is funded through the BMBF.

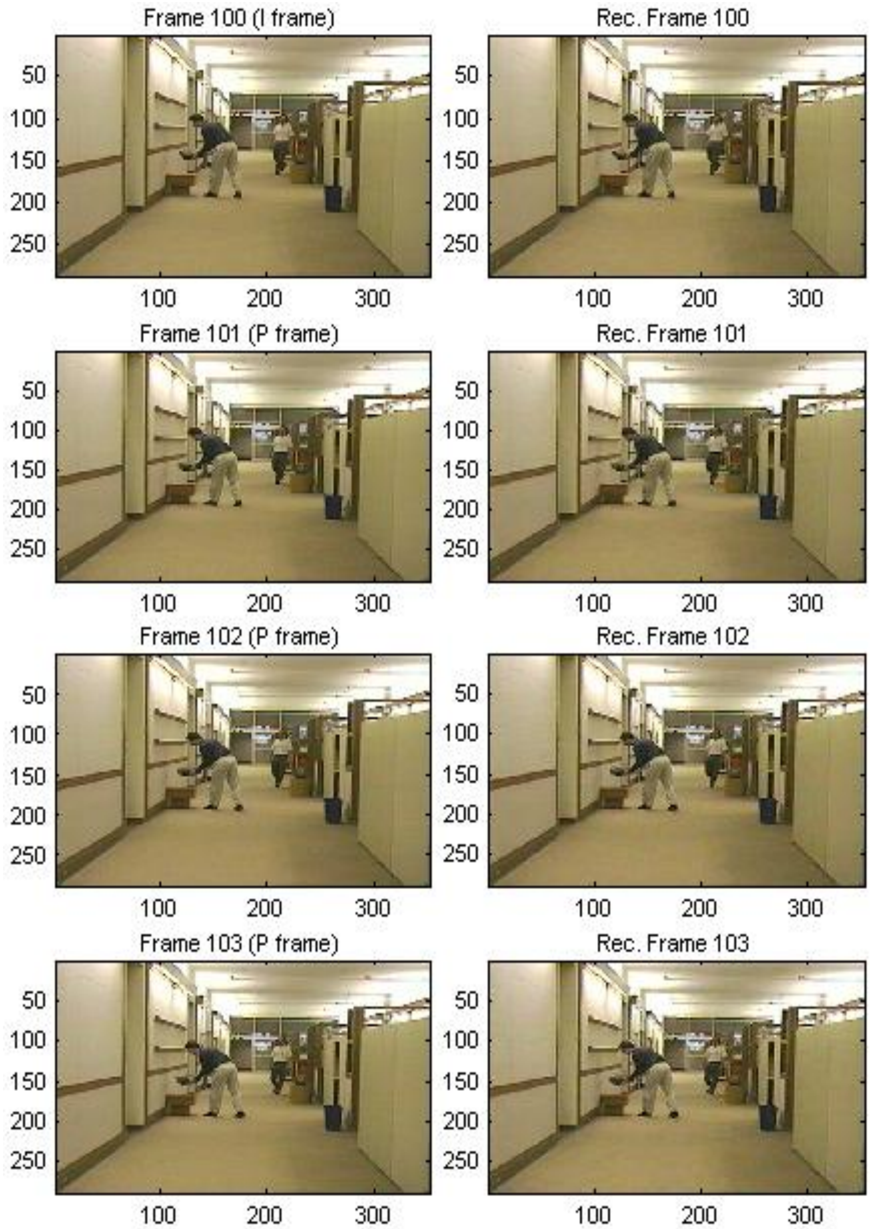


Figure 8. Original and reconstructed frames 100-103 of the Hall Monitor sequence. PSNR = 35.99dB, 34.23dB, 33.90dB and 33.41dB for frames 100, 101, 102 and 103 respectively.

REFERENCES

[1] Gershikov E. and Porat M., "Does Decorrelation Really Improve Color Image Compression?", the International Conference on Systems Theory and Scientific Computation (ISTASC05), Malta, 2005.

[2] Gershikov E. and Porat M., "A Rate-Distortion Approach to Optimal Color Image Compression", EUSIPCO 2006, Florence, Italy, September 2006.

[3] Gershikov E. and Porat M., "On Color Transforms and Bit Allocation for Optimal Subband Image Compression", Signal Processing: Image Communication, 22(1), Jan. 2007, pp. 1-18.

[4] Goffman-Vinopal L. and Porat M., "Color image compression using inter-color correlation", IEEE International Conference on Image Processing, 2, 2002, pp. II-353 - II-356.

[5] Han S.E., Tao B., Cooper T. and Tastle I., "Comparison between Different Color Transformations for the JPEG 2000", PICS 2000, 2000, pp. 259-263.

[6] Hao P. and Shi Q., "Comparative Study of Color Transforms for Image Coding and Derivation of Integer Reversible Color Transform", 15th International Conf. on Pattern Recognition (ICPR'00), 3, 2000, pp. 224-227.

[7] otera H. and Kanamori K., "A Novel Coding Algorithm for Representing Full Color Images by a Single Color Image", J. Imaging Technology, 16, Aug. 1990, pp. 142-152.

[8] Kouassi R.K., Devaux J.C., Gouton P. and Paindavoine M., "Application of the Karhunen-Loeve transform for natural color images analysis", IEEE Conference on Signals, Systems & Computers, 2, 1997, pp. 1740-1744.

[9] Lam E. Y., Goodman J. W., "A mathematical analysis of the DCT coefficient distributions for images", IEEE Trans. on Image Processing, 9(10), 2000, pp. 1661-1666.

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

[11] Mallat S. G., "A theory for multiresolution signal decomposition: the wavelet representation", IEEE Trans. on Pattern Analysis and Machine Intelligence", 11(7), July 1989, pp. 674 - 693.

[12] Porat M. and Shachor G., "Signal Representation in the combined Phase - Spatial space: Reconstruction and Criteria for Uniqueness", IEEE Trans. on Signal Processing, 47(6), 1999, pp. 1701-1707.

[13] Rabbani M. and Joshi R., "An overview of the JPEG2000 still image compression standard", Signal Processing: Image Communication", 17(1), 2002, pp 3-48.

[14] Rao K. R. and Yip P., "Discrete cosine transform: algorithms, advantages, applications", Academic Press, 1990.

[15] Y. Roterman and M. Porat, "Color Image Coding using Regional Correlation of Primary Colors", Elsevier Image and Vision Computing, 25, 2007, pp. 637-651.

[16] Satt A. and Malah D., "Design of Uniform DFT Filter Banks Optimized for Subband Coding of Speech", IEEE Trans. on Acoustics, Speech and signal Processing, 37(11), Nov. 1989, pp. 1672-1679.

[17] Sikora T., "MPEG digital video-coding standards", IEEE Signal Processing, 14(5), 1997, pp. 82-100.

[18] Supangkat S. H. and Murakami K., "Quantity control for JPEG image data compression using fuzzy logic algorithm", IEEE Trans. on Consumer Electronics, 41(1), 1995, pp. 42-48.

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

[20] Wallace G. K., "The JPEG Still Picture Compression Standard", IEEE Transactions on Consumer Electronics , 38, 1992, pp. xviii-xxxiv.

[21] Wolf S. G., Ginosar R. and Zeevi Y. Y., "Spatio-chromatic Model for Colour Image Processing", IEEE Proc. of the IAPR, 1, 1994, pp. 599-601.

[22] Yamaguchi H., "Efficient Encoding of Colored Pictures in R, G, B Components", Trans. on Communications, 32, Nov. 1984, pp. 1201-1209.