HDR Image Compression with Optimized JPEG Coding

Azza OULED ZAID
SysCom Laboratory, National Engineering School of Tunis
University of Tunis El Manar
Le belvedere, 1002 Tunisia
Email: azza.ouledzaid@ieee.org

Amira Houimli
Higher Institute of Computer Science
University of Tunis El Manar
2080, Ariana, Tunisia

***Abstract*—This paper presents an efficient compression system adapted to High Dynamic Range (HDR) images. First a Tone Mapping Operator (TMO) generates the Low Dynamic Range (LDR) version of the HDR content together with its extra information. The obtained LDR image is encoded using an optimized JPEG coding scheme, whereas the extra information is encoded as side data. Specifically, the optimized JPEG based algorithm constructs near-optimal rate-distortion quantization tables using DCT coefficient distribution statistics and Lagrangian optimization approach. To ensure accurate HDR reconstruction, the extra information is compressed with conventional JPEG encoder using the highest quality level. The aim of the proposed HDR coding system is twofold. First, it performs a bit allocation mechanism, to achieve near-optimal rate control. Second, it maintains the backward compatibility with the conventional JPEG. Experiments show that the compression performance of the proposed HDR coder outperforms that of the reference method.**

I. Introduction

High dynamic range (HDR) imaging is an attractive way of capturing real world appearance. It allows to preserve the accurate luminance values that can be found in real scenes. Due to the higher luminance range, each pixel is coded as a triple of floating point values, which can range from 10^5 to 10^{10}. Such floating point representation induces huge memory and storage requirements. In particular, the size of HDR content is one of the major barrier to be overcome for low space storage and high speed transmission. This critical problem can be resolved by designing efficient compression solutions for HDR images. To promote the wide spread use of HDR images, the backward compatibility with commonly used compression and display technologies is necessary. Since JPEG is currently the most commonly used imaging format, it is obvious that HDR image coding format should be backward compatible with JPEG format to facilitate its inclusion in current imaging ecosystems [1].

Several algorithms have been developed in the literature for the compression of HDR images. Among them, JPEG HDR [2] and HDR JPEG 2000 [3], that are extensions to JPEG and JPEG 2000 compression schemes, respectively. The JPEG HDR algorithm starts with the TM of the HDR content to obtain its 8-bits LDR version. Thereafter, the original HDR image is divided by the LDR one obtaining the ratio image (RI) which is stored as a sub-band. The RI is generally down-sampled to reduce the its size. During the HDR reconstruction, this down-sampling causes halos and/or glare around edges. To alleviate this problem, the authors proposed to introduce corrections in the tone mapped image. Despite their effectiveness, theses corrections produce artifacts in the LDR image for the backward compatibility and this cannot be tolerable in many applications.

In HDR JPEG 2000 coder [3] the dynamic range of HDR images is reduced using a logarithm with base 2. The obtained values are then remapped in unsigned (16-bit) integers, that are supported by JPEG 2000 standard. JPEG 2000 and JPEG XR also support HDR representations. However their adoption requires a certain investment not always affordable in existing imaging ecosystems as they are not backward compatible with the widely popular JPEG image format [1]. More recently, JPEG working group proposed the JPEG XT standard [4] which extends the JPEG functionalities by enabling support for high dynamic range imaging. Particularly, JPEG XT part 7 divides the image data in LDR JPEG legacy compliant codestream and an extra information. The latter is smartly inserted in the JPEG legacy codestream such that it does not jeopardize backward and/or forward compatibility.

Despite the recent developments in the field of HDR image coding and representation, there is a lack of rate-distortion optimization for HDR image compression. To date, no work has been made to develop an HDR compression method that tackles the problem of rate adaptation under communication constraints.

JPEG coding scheme is difficult to optimize because of its use of zero runlength coding, which combines zero coefficients from different frequency bands into one symbol. The resulting coupling of different frequency bands prevents the use of classical bit allocation methods, which assume independently coded frequency bands. The key to good compression (in the rate-distortion sense) when using Discrete Cosine Transform (DCT) lies in the quantizer step size selection. The simplest and most commonly used approach is to use a default table and scale it up or down until the desired target rate (or distortion) is reached. Other methods include psycho-visual model based quantization [8], and stochastic optimization techniques [9]. In both approaches, a particular quantization table Q is evalu-

ISBN 978-0-9928626-7-1 © EURASIP 2017

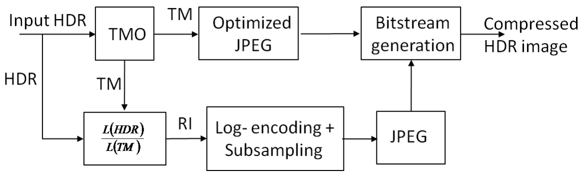


Fig. 1. Proposed scheme for JPEG backward compatible encoding.

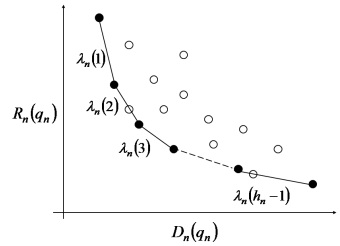


Fig. 2. The convex hull of rate-distortion points that is retained by the optimisation algorithm for Lagrangian minimization.

ated by going through the entire compression/decompression cycle. Later, Ratnakar proposed the RD-OPT algorithm [10] which uses DCT statistics for the given image to determine rate/distortion specific quantization tables with nearly optimal tradeoffs.

In the present work, we address the task for building an HDR image coding scheme aiming at offering high rate/distortion performance, while remaining faithful to the JPEG syntax. To achieve backward compatibility with JPEG, a TM operation is applied on the original HDR image to obtain its LDR version together with a restorative information, needed to reconstruct the HDR content. The tone-mapped image is encoded in the usual 8×8, 8-bit blocks using JPEG compression. To improve the rate-distortion (R-D) performance, we propose to integrate a bit allocation mechanism to JPEG coding chain. The main idea consists in using DCT coefficient statistics and Lagrangian minimisation strategy to construct near optimal quantization tables over a wide range of bit rates and distortions. At the decoder side, a naive JPEG decoder provides the LDR image, while an HDR-enabled decoder can decode the HDR image after performing an inverse TM on the decompressed LDR image. Unlike existing JPEG XT compliant encoder, our encoder can achieve rate control which is currently a very desirable feature for several multimedia applications like streaming on limited capacity channels.

The rest of the paper is organized as follows. In Section 2, we introduce our compression system for HDR images. In Section 3, we report the compression results and compare them to those of JPEG XT. Conclusions are presented in Section 4.

II. Description of the proposed framework

A. Global coding scheme

In the present work, we propose an HDR coding system that takes advantage of the simplicity of JPEG coding architecture and the efficiency of near optimal scalar quantization strategy. The global coding structure is depicted in Fig. 1. The HDR image is mapped onto an LDR image using a tone-mapping operator (TMO). In our implementation we use the the global version of Reinhard et al [7] TMO which is the most commonly used in many applications. The LDR version of the original HDR image is accompanied by a ratio image (RI) which is obtained by dividing the HDR original luminance ($L(HDR)$) pixel values by the tone-mapped luminance ($L(TM)$) ones. The RI image is then log-encoded and compressed using JPEG coding system with the highest quality setting, assuming that HDR reconstruction is more sensitive to ratio image noise. Note that the RI will be ignored by LDR image applications, but it will be used to reconstruct the HDR image. On the other hand, the tone-mapped image is efficiently encoded, as the foreground one, using optimized JPEG coder. This output referred signal is stored as an 8-bit greyscale image using JPEG/JFIF standard wrapper format. After being compressed, the encoded RI is inserted as metadata, in the tone-mapped image bitstream, using an appropriate application marker (APP). It is worth outlining that JPEG/JFIF provides a limited set of application markers. In the JFIF specification, there are sixteen application markers. A single marker holds a maximum of 64 Kbytes of data, according to the image dimensions. To overcome the 64 Kb limit, we proposed to store the compressed RI in multiple markers by reusing the same identifier.

B. Rate distortion optimization

To improve the coding performance for the tone-mapped image, we proposed to use an optimized bit allocation process inspired from the RD-OPT algorithm [10]. The latter determines the quantization table that achieves the minimum Lagrangian RD cost, without going through the entire compression-decompression cycle. An $H \times W$ image pixels is expressed by a set I of $\frac{H \times W}{64}$ blocks. By applying a 8×8 DCT transform to each block B of the original image, we obtain the DCT block C. Furthermore, the quantization table Q is specified by 64 step sizes that quantize the DCT coefficients at spatial indices $ij, (i, j = 0, ..., 7)$, in each 8×8 image block. To simplify the notation, we order the 8×8 set of spatial frequencies into a 1-D array of 64 coefficients indexed by n using the zig-zag scan, $n = 8 \times i + j$. Hence, denoting the quantized block DCT by $\hat{C}$, the quantized coefficient $\hat{C}[n]$ is calculated as:

$$\hat{C}[n] = Round(\frac{C[n]}{Q[n]}), \qquad (1)$$

where $Round(x)$ designates the closest integer to x. RD-OPT estimates the compression bit rate as the sum over the 64 elements of the measured entropies of the quantized DCT coefficients. For each possible quantizer step size q, $1 < q < 255$, the contribution to total bit rate $R[n][q]$ of

ISBN 978-0-9928626-7-1

the n^{th} component $C[n]$ is given by:

$$R[n][q] = \frac{1}{64} Entropy(Round(\frac{C[n]}{q})), \quad (2)$$

where the entropy is calculated over all the blocks. If $Round(\frac{C[n]}{q})$ takes the value v in a fraction $p_v > 0$ of all the blocks C, then, this entropy is $-\sum_v p_v \log_2 p_v$. Note that p_v values are approximated by using histograms of DCT coefficient statistics. Thanks to the orthonormal basis property, the total pixel distortion, in terms of mean-squared-error sense, in an image block equals the sum of the coefficient distortion in the corresponding DCT block. Define $D[n][q]$, the contribution to total distortion from the n^{th} coefficient as

$$D[n][q] = \frac{1}{64} Mean((C[n] - Round(\frac{C[n]}{q}))^2), \quad (3)$$

with the "Mean" is taken over all the non-overlapping 8×8 blocks in the image data. Afterwards, for a specific quantization table Q, the modelled bit rate is given by:

$$R(Q) = \sum_{n=0}^{63} R[n][Q[n]], \quad (4)$$

and the distortion is defined by:

$$D(Q) = \sum_{n=0}^{63} D[n][Q[n]]. \quad (5)$$

Hence, by using histograms of DCT coefficient distributions, RD-OPT builds tables of $D[n][Q(n)]$ and $R[n][Q(n)]$. In practice, this could be very accurate while being conceptually and computationally simple. The rate-distortion optimisation problem is formulated by Lagrangian minimization, $min(D(Q) + \lambda R(Q))$, in such a way that solutions of the latter for any no-negative λ are solutions to the former for a target bit rate R_{target} or a target quality D_{target}. Using both formulas 4 and 5, the Lagrangian minimisation problem reduces to minimising

$$\sum_{n=0}^{63} D[n][Q[n]] + \lambda R[n][Q[n]]. \quad (6)$$

For each quantized DCT coefficient indexed by n, we obtain a subset of the operating points by sorting, such that $D[n][.]$ is strictly increasing and $R[n][.]$ is strictly decreasing. Then the Graham algorithm [11] is performed to get the lower half of convex hull that includes $(D[n][.], R[n][.])$ points. Fig. 2 illustrates the convex hull of rate-distortion points that is selected by RD-OPT for Lagrangian minimization. Let the values q that give the h_n points on the convex-hull be represented by $q_n(1)$ through $q_n(h_n)$. The slopes of the rate-distortion curve for the n^{th} coefficient at these h_n points are given by:

$$\lambda_n(l) = \frac{R[n][q_n(l)] - R[n][q_n(l+1)]}{D[n][q_n(l+1)] - D[n][q_n(l)]} \quad (7)$$

where $1 \leq k \leq h_n - 1$. The optimal slope value λ^* is unknown in advance and varies with the target bit rate

TABLE I
EVALUATION OF THE COMPRESSED LDR IMAGES USING TMQI METRIC [14].

Image	Method	Bitrate				
		0.3	0.6	0.8	1	1.5
Chairs	JPEG XT [1]	0.8	0.83	0.842	0.844	0.848
	Our method	0.9	0.94	0.94	0.95	0.95
Room	JPEG XT [1]	0.695	0.724	0.73	0.735	0.74
	Our method	0.75	0.79	0.8	0.8	0.81
Doll small	JPEG XT [1]	0.749	0.766	0.769	0.77	0.77
	Our method	0.8	0.81	0.83	0.83	0.83
Memorial	JPEG XT [1]	0.736	0.767	0.777	0.78	0.787
	Our method	0.8	0.82	0.83	0.83	0.83

($R_{target} = R(Q)$) or quality ($D_{target} = D(Q)$) constraint. But it can be determined according to a fast convex search using the bisection algorithm [12]. The optimized bit allocation mechanism can then be summarized as follows:

Input: An image I of size $H \times W$,
Output: Near optimal DCT quantization tables Q,
Step 1: Gather the statistics of DCT coefficients for the processed image,
Step 2: Use the statistics to determine $R[n][q]$ and $D[n][q]$ for the tested q value,
Step 3: Use Lagrangian minimization to optimize the bit rate, $R(Q)$, against the distortion, $D(Q)$.

III. PERFORMANCE RESULTS

The performance of the developed compresison method is assessed in this section. Particularly, we analyse the distortions of LDR and HDR images, respectively, at various bit rates. Four HDR test images are used: Memorial of size 512×768, Room of size 1840×1224, Chairs of size 343×231, and Doll-small of size 461×450. These images are in pfm format (96 bits/pixel). Our rate-distortion results have been compared to the reference method JPEG XT [1]. It is unsuitable to compare the proposed method with HDR-JPEG 2000 [3]. Indeed, the latter reduces the range of the HDR pixels in unsigned short integers (16-bit), and consequently does not support the 8-bit images neither backward compatibility with JPEG. Concerning the JPEG XT reference software, which was accessible from the Internet public domain (www.jpeg.org/jpegxt/software.html), we used profile A. The performance evaluation presented in [1] showed that compared to profiles B and C, profile A exhibits higher performance at low bit rates and appears to be more independent on the used TMOs for the base layer. To offer a fair comparison, the highest quality setting for the residual layer is selected, and the optimized Huffman coding is enabled, for both coding methods. Furthermore, no constraints were fixed on the quality of the base layer of the JPEG XT compressed image. This means that the best rate-distortion constrain was obtained by varying the quality parameter until the desired target bit rate is reached. Before evaluating the performance of our coding scheme, in terms of HDR reconstruction, we first assess the quality of the decompressed tone-mapped images for a wide

ISBN 978-0-9928626-7-1

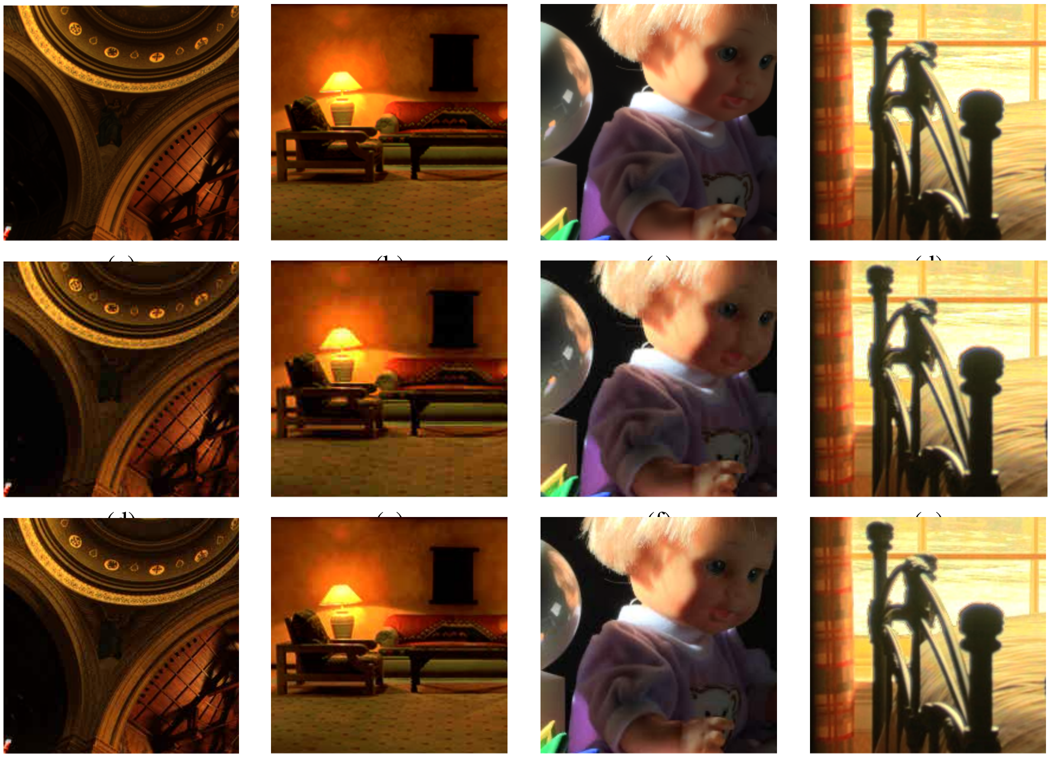


Fig. 3. Subjective comparison of the test HDR compression methods: (a)-(b)-(c)-(d) close-up shots of the tone-mapped images, (e)-(f)-(g)-(h) decoded tone-mapped images using JPEG XT, (i)-(j)-(k)-(l) decoded tone-mapped images using our method.

range of bit rates. To this end, the visual qualities of the tone mapped images are evaluated using two objective metric: the tone-mapped image quality index (TMQI), which assesses the quality of LDR image using its corresponding HDR image as reference. This quality measure is defined by a mixture of a structural fidelity evaluation and a statistical naturalness measure. From the results reported in Table I we can see that when LDR encoding is considered, our method presents the higher TMQI scores for all the tested bit rates. It leads to an averaged compression performance improvement of 12.63% compared to the JPEG XT codec when taking the TMQI to serve as distortion metric. Furthermore, we see that the highest performance is attained for Chairs image.

Fig. 3 depicts close up shots extracted from Memorial, Chairs, Doll small, and Room tone mapped images, compressed at 0.6 bpp using JPEG XT and our method. From this Figure, we can observe that the images compressed by our coding system have a higher quality visual appearance in the homogeneous zones. On the other hand, the JPEG XT images are penalized by perceptually block boundary artifacts. One may also notice that, due to the tone-mapping effect, the close-up shots extracted from Room image present some aliasing effects and artifacts in color transitions. To evaluate the effectiveness of our compression method in the case of HDR reconstruction, we use the Multi-Exposure Peak Signal Noise Ratio (mPSNR) [13], which is a popular objective quality metric for assessing HDR compression methods. For a given HDR image, several tone mapped versions with

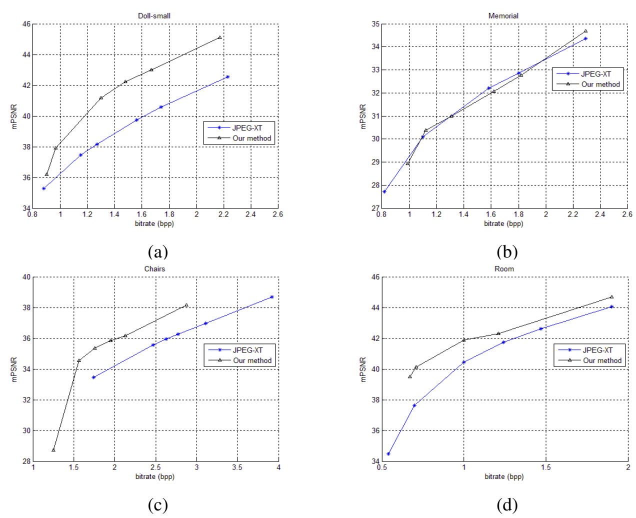
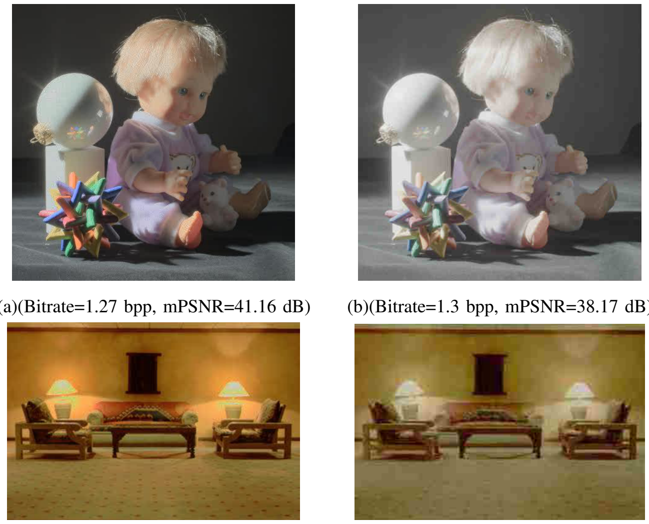


Fig. 4. Rate-Distortion comparison among the tested HDR compression techniques for (a) Doll small, (b) Memorial, (c) Chairs, and (d) Room HDR images.

different exposure parameter values, are created, and then PSNR measures are calculated for these tone-mapped versions. The mPSNR value is then an aggregate of the obtained PSNR measures.

Fig. 4 shows the rate-distortion (mPSNR versus bit rate) curves obtained with our method and JPEG XT for Room, Chairs, Memorial, and Doll small HDR images. From the results reported in this figure, we can plainly observe that our coding scheme outperforms the reference method for the test

ISBN 978-0-9928626-7-1 © EURASIP 2017



(a)(Bitrate=1.27 bpp, mPSNR=41.16 dB) (b)(Bitrate=1.3 bpp, mPSNR=38.17 dB)

(c)(Bitrate=1.76 bpp, mPSNR=35.35 dB) (d)(Bitrate=1.74 bpp, mPSNR=33.47 dB)

Fig. 5. Visual comparison of the compressed HDR images using: (a)-(c) our method, and (b)-(d) JPEG XT. The images were tone-mapped to be printed.

HDR images over all tested bit rates. The average margin of quality gain provided by our method varies from 0.2 dB to more than 2.3 dB. The same figure also shows that the image content has a significant influence on the mPSNR. The highly demanding images, in terms of the relation mPSNR vs. bit rates, are the same for both tested methods. We also note that the test image Memorial exhibits the lowest mPSNR values. This is most likely due to the fact that memorial image is characterized by highly textured appearance and consequently, does not beneficiate from the global rate-distortion mechanism which is based on DCT coefficient histograms through the 64 frequency localizations.

Fig. 5 depicts Doll small and Chairs HDR images compressed using JPEG XT and our coding system. From this figure, we can notice that the 2 sets of images show big differences in terms of luminance and colour not related to coding artifacts. This may be due to the use of the post-scaling nonlinearity block in the expanding stage of JPEG XT.

Examining the Doll small image, which is shown in Fig. 5(a) and Fig. 5(b), it is clear that our compression algorithm improves the visual quality for homogenous regions. The Doll small image compressed by JPEG XT appears very blur and the block artefacts are noticeable, whereas our method produces very clear and sharp image. Regarding Chairs HDR image depicted in Fig. 5(c) and Fig. 5(d), it appears that our compression system improves the visual quality of the compressed image compared to the JPEG XT, especially in the block boundary regions. It is important to note that despite its high efficiency in terms of rate-distortion results, our compression method allows to attain the target bit rate or quality (fixed by the user), whereas the state of the art methods have to be rerun to reach the desired quality or compressed size.

IV. Conclusion

In this paper, we have presented a compression method, for HDR images, which ensure high compression performance while guaranteeing backward compatibility with JPEG legacy format. The proposed system decomposes an HDR image into its tone-mapped version and a ratio image. An optimized JPEG coder, in the sense of rate-distortion sense, is applied on both image. To this end, an efficient bit allocation algorithm, that determines a sub-optimal rate-quality tradeoff, has been integrated to JPEG coding chain. Simulation results demonstrated that even if using simple tone mapping algorithm and simple approach for calculating residual information, our method can accomplish better compression efficiency compared with JPEG XT standard.

Future work includes the use of sophisticated tone-mapping algorithm. This may leads to better HDR reconstruction results. Besides, we plan to consider an optimal allocation of the bit budget to the LDR and ratio images to increase the HDR compression effectiveness.

References

[1] A. Artusi and R. K. Mantiuk and T. Richter and P. Hanhart and P. Korshunov and M. Agostinelli and A. Ten and T. Ebrahimi, *Overview and evaluation of the JPEG XT HDR image compression standard*, vol. 25, no. 6, pp. 1–16, 2015.

[2] G. Ward and M. Simmons, *JPEG-HDR: A backwards-compatible, high dynamic range extension to JPEG*, ACM SIGGRAPH 2006 Courses, 2006.

[3] R. Xu and S. Pattanaik and C. Hughes, *High-dynamic-range still-image encoding in JPEG 2000*, IEEE Computer Graphics and Applications, vol. 25, no. 6, pp. 57–64, 2005.

[4] T. Richter, *Backwards Compatible Coding of High Dynamic Range Images with JPEG*, Proceedings of Data Compression Conference (DCC), pp. 153–160, 2013.

[5] D. Taubman, *High performance scalable image compression with EBCOT*, IEEE Transactions on Image Processing, vol. 9, no. 7, pp. 1151–1170, 2013.

[6] S. Srinivasan and C. Tu and Zhi Zhou and D. Ray and S. Regunathan and G. Sullivan, *An Introduction to the HDPhoto Technical Design*, JPEG document WG1N4183, 2007.

[7] E. Reinhard and M. Stark and P. Shirley and J. Ferwerda, *Photographic Tone Reproduction for Digital Images*, ACM Transactions on Graphics, vol. 21, no. 3, pp. 267–276, 2002.

[8] A.B. Watson, *DCT quantization matrices visually optimized for individual images*, in Processings of the SPIE, Human Vision, Visual Processing, and Digital Display IV, 1993.

[9] S. Wu and A. Gersho, *Rate-constrained picture-adaptive quantization for JPEG baseline coders*, in Processings of International Conference of Acoustics, Speech and Signal Processing, vol. 5, pp. 389–392, 1993.

[10] V. Ratnakar and M. Livny, *An Efficient Algorithm For Optimizing DCT Quantization*, IEEE Transactions on Image Processing, Speech and Signal Processing, vol. 9, no. 2, pp. 267–270, 2000.

[11] V. Ratnakar and M. Livny, *An Efficient Algorithm for Determining the Convex Hull of a finite Planar Set*, Information Processing Letters, vol. 1, no. 4, pp. 132–133, 1972.

[12] K. Ramchandran and M. Vetterli, *Best Wavelet Packet Bases in a Rate-Distortion Sense*, IEEE Transactions on Image Processing, vol. 2, no. 2, pp. 160–175, 1993.

[13] J. Munkberg and P. Clarberg and J. Hasselgren and T. Mller, *High dynamic range texture compression for graphics hardware*, ACM Transactions on Graphics, vol. 25, no. 3, pp. 698–706, 2006.

[14] H. Yeganeh and Z. Wang, *Objective quality assessment of tone-mapped images*, IEEE Transactions on Image Processing, vol. 22, no. 2, pp. 657–667, 2013.

[15] A. Mittal and R. Soundararajan and A. Bovik, *Making a completely blind image quality analyzer*, IEEE Signal Processing Letters, vol. 20, no. 3, pp. 209–212, 2013.

ISBN 978-0-9928626-7-1 © EURASIP 2017