Optical Engineering

SPIEDigitalLibrary.org/oe

Context-dependent JPEG backward-compatible high-dynamic range image compression

Pavel Korshunov
Touradj Ebrahimi

SPIE

Context-dependent JPEG backward-compatible high-dynamic range image compression

Pavel Korshunov
Touradj Ebrahimi
Multimedia Signal Processing Group
École Polytechnique Fédérale de Lausanne
CH-1015 Lausanne, Switzerland
E-mail: pavel.korshunov@ep.ch

Abstract. High-dynamic range (HDR) imaging is expected, together with ultrahigh definition and high-frame rate video, to become a technology that may change photo, TV, and film industries. Many cameras and displays capable of capturing and rendering both HDR images and video are already available in the market. The popularity and full-public adoption of HDR content is, however, hindered by the lack of standards in evaluation of quality, file formats, and compression, as well as large legacy base of low-dynamic range (LDR) displays that are unable to render HDR. To facilitate the wide spread of HDR usage, the backward compatibility of HDR with commonly used legacy technologies for storage, rendering, and compression of video and images are necessary. Although many tone-mapping algorithms are developed for generating viewable LDR content from HDR, there is no consensus of which algorithm to use and under which conditions. We, via a series of subjective evaluations, demonstrate the dependency of the perceptual quality of the tone-mapped LDR images on the context: environmental factors, display parameters, and image content itself. Based on the results of subjective tests, it proposes to extend JPEG file format, the most popular image format, in a backward compatible manner to deal with HDR images also. An architecture to achieve such backward compatibility with JPEG is proposed. A simple implementation of lossy compression demonstrates the efficiency of the proposed architecture compared with the state-of-the-art HDR image compression. [DOI: 10.1117/1.OE.52.10.102006]

Subject terms: high-dynamic range; backward compatibility; JPEG; subjective tests; tone mapping; compression.

Paper 130320SSP received Mar. 1, 2013; revised manuscript received Jul. 17, 2013; accepted for publication Jul. 18, 2013; published online Aug. 19, 2013.

1 Introduction

High-dynamic range (HDR) imaging is an increasingly popular topic and has been the focus of attention in both scientific and artistic communities for several years. Several proprietary image and video compression algorithms have been proposed in the literature for the compression of HDR content. Among them, the most popular are those that are backward compatible with existing low-dynamic range (LDR) image and video compression standards.[1,2] More recently, some standardization committees have also rolled-out compression algorithms that can handle both HDR and LDR content, such as JPEG 2000 and JPEG XR for images and MPEG-4 and H.264/AVC for video, providing support for HDR. There is also an ongoing work on HDR compression standard by JPEG standardization body. However, despite the recent developments, there is a lack of standards in the evaluation of quality and compression of HDR images and video. Most such solutions are ad hoc patches to existing algorithms, where care has been mainly devoted to extend existing approaches to compression and storage and adapting them to cope with HDR content without an explicit exploitation of knowledge of the human visual system, statistical properties of HDR content, or application-specific parameters. Such an approach leads to suboptimal compression solutions for HDR imaging.

Many different subjective evaluations have been previously performed to compare different tone-mapping operators for HDR images and video. The main focus of these studies was either to determine a more superior approach to tone mapping or establish an evaluation methodology for subjective evaluation of HDR content. As different evaluations result in different sets of best tone-mapping algorithms, it demonstrates that other factors may also affect perceptual quality of the resulting LDR images. To achieve the best possible viewing experience, it is necessary for an HDR compression algorithm to accustom the factors affecting perception of resulted tone-mapped images.

In order to take human perception into account, we analyze the impact of contextual and environmental parameters on perception of quality in HDR image and video such as display type, size, contrast, and brightness characteristics, as well as the type of content in different surrounding lighting conditions. To understand how differently the HDR tone-mapping operators influence the perception under different conditions, we conducted a comprehensive subjective evaluation with 20 human subjects participating in the study. Five commonly used and cited in research literature tone-mapping operators were selected for the subjective study: (1) proposed by Drago et al.,[3] (2) proposed by Mantiuk et al.,[4] (3) proposed by Reinhard and Devlin,[5] (4) proposed by Fairchild and

0091-3286/2013/$25.00 © 2013 SPIE

Johnson,[6] and (5) a simple logarithm-based operator as a baseline. We used the test bed and infrastructure consisting of displays with different sizes and characteristics such as mobile phones, tablets, and large monitors. Environmental conditions and contextual information included the amount of environmental lighting, the way subjects viewed the images, and the backlit light of displays, as well as their size and contrast. One main novelty of this approach is evident when considering that the majority of current evaluation work in HDR ignores the context and environmental factors. By varying different environmental parameters, one could see how these factors affect the perceptual quality of the content and use it when designing a backward-compatible HDR compression.

The results of the subjective evaluation motivated us to develop an HDR compression algorithm that takes into account the statistical properties of the environment, the content, and used tone-mapping algorithm. We consider JPEG format for backward compatibility as it is the most popular format for images. We propose a generic compression scheme accommodating the above-mentioned contextual and environmental parameters into the encoding and decoding design. We then implement a prototype version of the compression scheme using only simple tone-mapping operators such as linear, gamma, and logarithm-based operators. Even this trivial implementation resulted in a more efficient encoding of HDR images in terms of size of compressed data and PSNR values, demonstrating the advantage of utilizing the knowledge about the tone-mapping algorithm that was used.

This paper is an extended version of the work by Korshunov and Ebrahimi.[7] In this extension, rate-distortion curves for several possible implementations of the proposed compression scheme are presented and compared with rate-distortion curves of current state-of-the-art approaches such as JPEG-HDR, JPEG backward-compatible lossy HDR compression, as well as adaptation of JPEG 2000 for HDR images and JPEG XR, both of which are not JPEG-backward compatible.

2 Related Work

Several subjective evaluation studies have been conducted in literature to compare different tone-mapping methods for HDR images. Subjective evaluations of HDR images were first performed by Ledda et al.[8] The authors used paired comparison to evaluate the perceptual quality of six different tone-mapping algorithms. An HDR display was used as a reference display for subjects. The focus of this work was on the evaluation methodology for the subjective comparison of HDR images in a controlled environment. The evaluations provided the performance ranking of different tone-mapping algorithms leading to different perceptual qualities in color and gray images.

Yoshida et al.[9] evaluated seven different tone-mapping algorithms via subjective tests to rate the resulted images with regards to their naturalness, contrast, brightness, and details of the reproduction in bright and dark regions. A prior study was conducted to find out the best tuning parameters for each tone mapping. The overall goal was to see if different tone-mapped images are perceived differently. The results show that brightness was the largest differentiator, and local-based tone mapping is better for details in brighter regions. Three methods were concluded as producing the most natural LDR from HDR images.

The focus in the work by Park and Montag[10] was on the scientific images (astronomic, medical, infrared, and radar), where nine tone-mapping algorithms were evaluated. Subjective tests with 25 subjects were conducted in a typical controlled environment, where only default parameters were used. Paired comparison was used to (1) find out which tone mapping is preferable perceptually and (2) determine which operator is more "scientifically useful." Results concluded that there were no correlations between perceptual preference and scientific usefulness. One method was selected as performing best in both criteria.

Kuang et al.[11] studied the overall preference of tone-mapping algorithms (nine in total) and their accuracy when compared with actual world scenes from which they were captured. The focus was on the evaluation methodology. Three subjective studies (paired comparison, rating-scale method, and real-world scenes method) were performed with fixed environmental parameters (lighting, luminance, screen sizes, etc.) and with 19 to 23 subjects. In the rating-scale method, the following features were investigated: details, shadow detail, overall contrast, sharpness, colorfulness, and lack of artifacts. Different evaluation methods were compared with each other. Bilateral filter was shown to outperform other tone-mapping algorithms, which seems to indicate a different conclusion when compared with other papers mentioned above.

Čadík et al.[12] studied how tone-mapping algorithms (14 in total) affect the comprehensive set of image attributes, namely and mainly, brightness, contrast, color reproduction, reproduction of details, and artifacts. They proposed an overall quality metric as a combination of the image attributes. Two subjective tests with 20 subjects in total were performed: (1) with real-world reference and (2) without a reference. Authors did not find any significant differences between the two testing methodologies in terms of results. The global tone-mapping algorithm was shown to be superior.

Annighöfer et al.[13] evaluated eight tone-mapping operators against linear tone mapping with 51 subjects. The focus of their work was on objective metrics that match subjective results. In addition, HDR images with restricted bit resolution were used (with different base exposures). Three tone-mapping algorithms were found to perform well, consistent with previous studies. However, the performance of each operator was found to be content dependent. The Naka-Rushton RMS objective metric demonstrated to be the most robust.

The study by Mai et al.[14] extended the subjective evaluations to three-dimensional (3-D) HDR content on 3-D stereoscopic displays with active shutters. The main objective of this work was to study which image attributes contribute to a good 3-D representation, and what are the differences with the two-dimensional (2-D) case. A pilot study was conducted to find the best possible parameters for the tone-mapping algorithms. 3-D effect and overall quality of the tone-mapped images were evaluated including contrast, naturalness, sharpness, and detail reproduction image attributes. Paired comparison and rating scales were used. Global tone-mapping operators were found to outperform others (same as in 2-D case). In addition, a high correlation between brightness and 3-D effect was found in the results.

As opposed to the above work, the goal of our study is not to find the best tone-mapping algorithm, but to demonstrate the importance of context (display size, ambient lighting, and image-viewing environment) and image content on the perceived quality of the resulted tone-mapped images. Then, by taking into account these factors, one can develop more efficient compression algorithms for HDR images and can design better file formats for HDR. Currently existing HDR compression schemes do not offer the desired flexibility, simplicity, or efficiency that would be comparable with the JPEG and JPEG 2000 standards that exist for LDR images. Similarly, file formats suffer from the lack of support for both lossless and lossy compression, backward compatibility with legacy LDR images, or standardization support.

One HDR file format currently more popular than others is Radiance, which was first proposed by Ward.[15] This format encodes floating-point HDR pixels represented in 32-bits RGBE format with red, green, and blue mantissas plus a common exponent, each stored as 16-bits integers. This format was supported by released Radiance software,[16] which is a comprehensive set of tools for image manipulation and rendering, implemented under public license and significantly boosting the popularity of the format. This format uses simple run length encoding to losslessly compress an image.

The same Ward together with Simmons also proposed a JPEG backward-compression algorithm and file format called JPEG-HDR,[1] which is one of the few lossy HDR compression schemes and is the main JPEG backward-compatible scheme widely used. JPEG-HDR compresses an image into a JPEG tone-mapped version of that HDR image (any tone-mapping operator is allowed to be used), and a residual data (also compressed by JPEG) stored as an extension of JPEG. This approach allows any conventional JPEG decoder to render a tone-mapped version, and for JPEG-HDR aware software to reconstruct the original HDR image. In this article, we agree with the idea of having JPEG backward-compatible HDR compression and format, but propose a more general and more adaptive architecture of the compression algorithm that would allow one to have both lossless and lossy compression, as well as to achieve a more efficient compression and flexibility, given how the research on compression has advanced since JPEG-HDR was first introduced.

This idea of using JPEG compression for both tone-mapped version (to ensure JPEG backward compatibility) of an HDR image and the residual data is further elaborated by Richter.[17] This variant of JPEG backward-compatible compression is designed as an extension (with the aim to have the least possible modifications) to the original JPEG standard and can be viewed as a modification of JPEG-HDR.

Another attempt on designing HDR compression scheme that is backward compatible with JPEG was done by Chen et al.[18] The authors also propose to store a tone-mapped version of the original HDR image as a JPEG file and residual (the difference between original HDR and reconstructed HDR from tone-mapped version) in an extension of the same JPEG file. The main difference of our approach is the ability to accommodate different tone-mapping operators and such a decoding scheme that allows displaying the same HDR image differently, depending on its content and given context. Also, the authors restrict residual to be under 64 kb of size only, which is not acceptable for images with large resolution.

JPEG XR and JPEG 2000 are two standards that are capable of storing and compressing HDR content. The main problem with these formats, however, is that they are not very popular and have little software support for HDR content. We believe that, since JPEG is currently a *de facto* and the most popular imaging format, any next generation format and compression for HDR content is bound to offer backward compatibility with JPEG in order to increase chances of use by the public.

3 Subjective Evaluations

The goal of this section is to analyze the suitability of the most common image and video quality evaluation methods for the subjective evaluation of HDR content and to adapt/extend these methods to take into account the contextual and environmental information. We study the effect of the environmental conditions, display characteristics, and content types on the perceptual quality of HDR images. We have designed a comprehensive methodology for subjective evaluation of quality and conducted a supporting set of subjective tests to build a model of the perception of quality of HDR content by human subjects in various contexts and environments.

3.1 *Test Material and Tone-Mapping Algorithms*

As the aim of the subjective study was not to find the best tone-mapping algorithm, but to understand how they perform in different conditions, we selected the following five tone-mapping operators (shortened accordingly): "dg" by Drago et al.,[3] "mt" by Mantiuk et al.,[4] "rh" by Reinhard and Devlin,[5] iCAM ("ic") by Fairchild and Johnson,[6] and "lg" as a simple logarithm-based operator. The implementation provided in library "pfstools" was used for the first three operators, and code provided in the book by Reinhard et al.[19] was used for the last two. These algorithms were selected to have representation of different approaches to tone mapping such as global operators ("lg" and "dg"), based on local information ("rh" and "ic"), and operators utilizing properties of human visual system ("mt").

To test the selected tone-mapping operators, we used four images (see Fig. 1) from the collection provided in the book by Reinhard et al.,[19] as the images typically used for evaluation of tone-mapping operators have very low resolution for today's monitors. The smallest resolution of the chosen images is 1084×1224. Images are stored in Radiance file format[15] with 32-bits per pixel (bpp). Therefore, tone-mapping operators map floating-point values of HDR content in RGBE representation into 24-bits RGB. The images were also selected in such a way to have different representation in terms of content type (indoor or outdoor) and luminance range (night and day shots).

3.2 *Test Environment and Methodology*

We have conducted subjective tests as pairwise comparison of different tone-mapping operators. Pairwise comparison of five operators with four test images comprises 40 comparison pairs in total.

(a) (b) (c) (d)

Fig. 1 Four high-dynamic range (HDR) images: (a) CraterLake1, (b) Duomo1, (c) BridgeStudios2, and (d) Room were used in the experiments.

The goal of the subjective evaluations of different tone-mapping operators was to find how different the resulted perceptual visual quality is under different environmental conditions, on different devices, and with different types of content. The tests were performed in a laboratory for professional subjective tests when using a monitor and in a typical office environment when tablets and mobile phones were used.

The test room was equipped with a 30-in. LCD Eizo monitor with resolution 2560×1600. The ambient lighting was obtained with neon lamps of 6500-K color temperature, and the walls were painted a mid gray 128, as recommended in Ref. 20. The luminance of the Eizo screen was set with EyeOne Display 2 calibration tool to 120 $\mathrm{cd/m^2}$. We used first generation of iPad with resolution 1024×768 in experiments with tablets and Samsung Galaxy S with resolution 800×480 for tests with mobile phones. The office for testing with tablets and mobile phones had a typical lighting of about 500 lux, and the brightness was turned to maximum in both types of devices.

In practical usage scenario, the environmental ambient light can be determined by either frontal camera of the device (in case of mobile phone and tablet) or a separate web camera (in case of monitor).

Subjective tests with the monitor were performed in a passive mode. Each subject was sitting in front of the monitor at a distance two to three times the height of the stimuli (presented images). A pair of test stimuli in the same comparison set were played one-after-another. All possible pairs in each comparison set were used for comparison in order to obtain complete winning frequency matrices. Since the monitor used had a native resolution of 2560×1600, images could fit in the horizontal space of the display. Each subject was asked to choose which stimulus had better quality between the two presented stimuli and to mark the answer between "first," "same," and "second" on the score sheet. Each stimuli of the pair was displayed for 7 s, and 5 s were given to vote after the pair. Each subject had a training session followed by two separate test sessions, each of which contained 20 pairs of stimuli. Twenty subjects (12 males and 8 females) having normal or corrected-to-normal vision participated in the tests having a median age of 25 and age range from 20 to 61 years old.

The experiments with mobile phones and tablets were performed in a similar way as with the monitors, except users were allowed to scroll through the pairs of images by themselves, enabling a more realistic active mode of subjective evaluations in such contexts.

The test images (see Fig. 1) in original "Radiance" format were resized to fit different resolutions of monitors, tablets, and mobile phones. Then, all the five selected tone-mapping operators were run on each image with default settings to produce LDR versions in JPEG format, which were used in the pairwise comparisons.

3.3 *Evaluation Results*

To have a better understanding of the subjective tests results, we first computed the number of subjective votes for each compared pair of tone-mapping operators, as presented in Figs. 2–5. All possible pairs of tone-mapping algorithms under question are displayed on vertical axis including

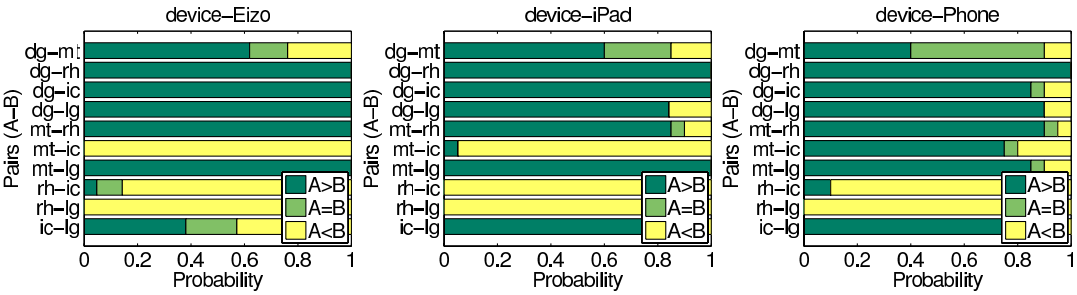

Fig. 2 Scores for each compared pair of tone-mapping operators for "CraterLake1" image.

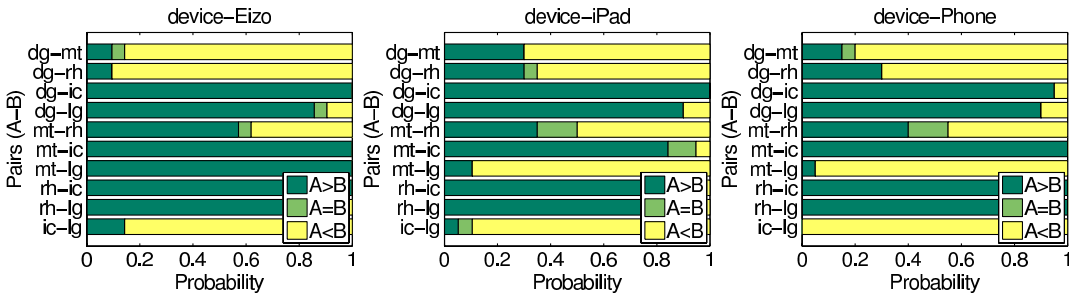

Fig. 3 Scores for each compared pair of tone-mapping operators for "Duomo1" image.

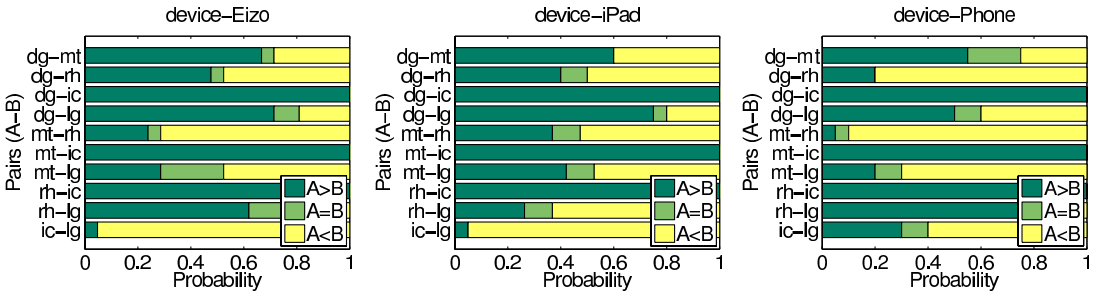

Fig. 4 Scores for each compared pair of tone-mapping operators for "BridgeStudios2" image.

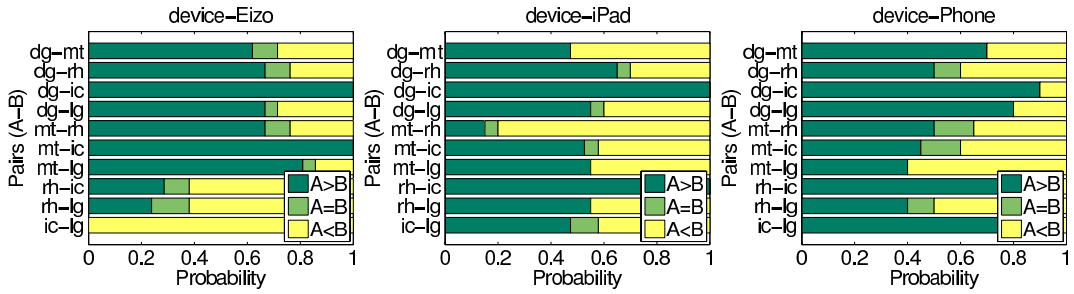

Fig. 5 Scores for each compared pair of tone-mapping operators for "Room" image.

"dg" (by Drago et al.[3]), "mt" (by Mantiuk et al.[4]), "rh" (by Reinhard and Devlin[5]), "ic" (iCAM by Fairchild and Johnson[6]), and logarithmic as "lg." Values representing horizontal bars are computed as follows (in the order from left to right): the number of subjects that favored first algorithm of the compared pair (denoted as "$A > B$" in the figure), the number of subjects voting that the algorithms were the same (denoted as "$A = B$"), and the number of subjects that favored the second algorithm (denoted as "$A < B$").

All these values are divided by the total number of subjects that participated in tests for each device resulting in a probability value. Each figure presents results for each image with three subfigures corresponding to three devices (in order of appearance): Eizo monitor, iPad tablet, and Samsung mobile phone on which the evaluations were performed.

As can be noted from these figures, there is a variety of scores across both devices and different images with variety

being more significant across images. It means that content of the images (the luminance range, whether it is a day or night shot, variety of details, etc.) plays a more significant role in determining which is better suited by tone-mapping.

To clearly illustrate which tone-mapping outperforms in which conditions, we have computed judging probabilities from the subjective tests using BTL (Bradley–Terry–Luce) model,[21] which is a commonly applied model for comparison of pairwise data. The judging probabilities for each tone-mapping operator are presented in Figs. 6–9. From these figures, one can identify rather easily the most and the least favorite tone-mapping algorithms for each device and for each image.

From the selected set of tone-mapping operators, the one by Drago et al. performs the best in most scenarios, with exceptions in a few cases only including "Duomo1" image (for all devices) and "BridgeStudio2" image (only for the Phone device, as per the right graph in Fig. 8). Such declines in performance of this tone-mapping operator could be because the default input parameters were used, and no tuning of the algorithm for darker images was done.

4 Proposed HDR Image Compression

The subjective evaluations' results reported in the previous section can guide the design of an efficient JPEG backward-compatible compression beyond the state-of-the-art. The

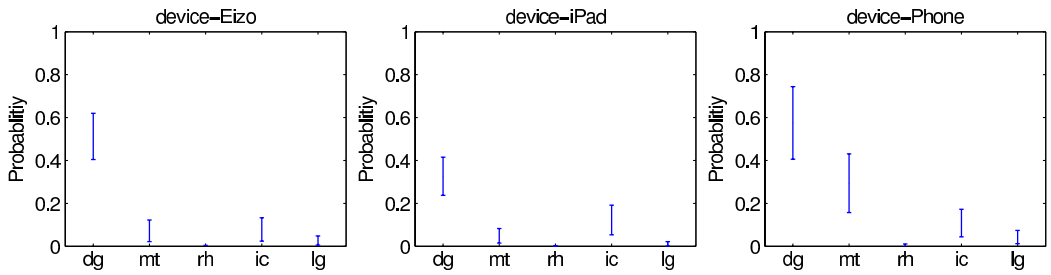

Fig. 6 Overall subjective scores of tone-mapping operators for "CraterLake1" image.

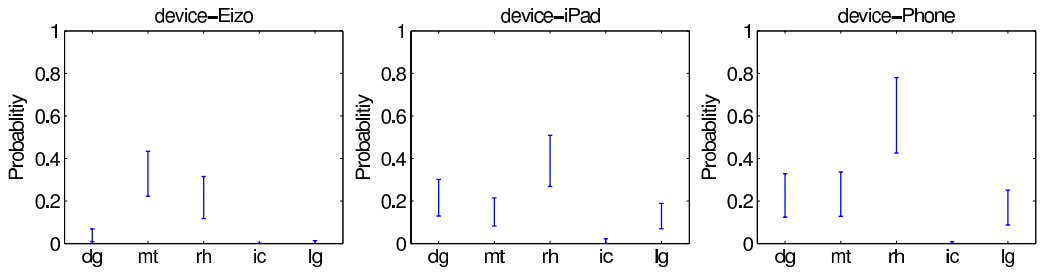

Fig. 7 Overall subjective scores of tone-mapping operators for "Duomo1" image.

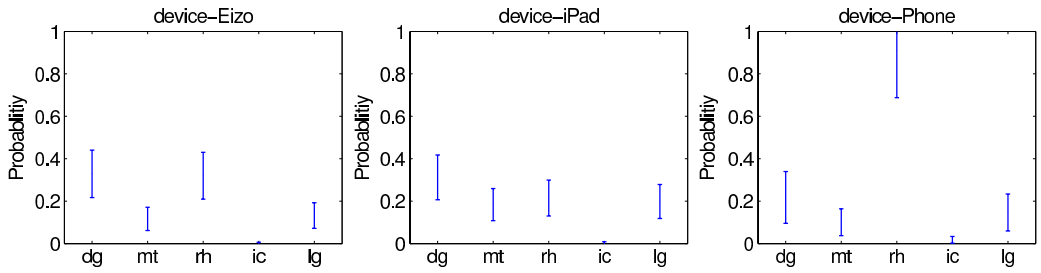

Fig. 8 Overall subjective scores of tone-mapping operators for "BridgeStudios2" image.

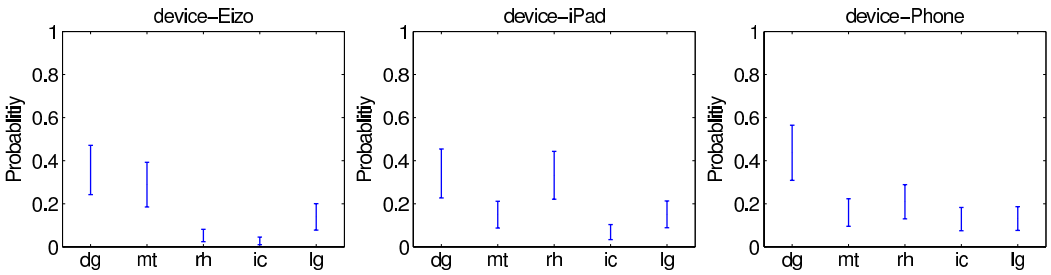

Fig. 9 Overall subjective scores of tone-mapping operators for "Room" image.

subjective scores indicate that to be optimal, compression needs to be adaptive to different environments, devices, and contents. That means a certain mechanism is required for the selection of an appropriate tone-mapping operation when converting an HDR image to LDR for display on different devices and under different conditions.

JPEG is selected to represent LDR version of the HDR image because of the wide-spread use of this format and availability of image viewers that are able to handle images in JPEG. Although a few proprietary JPEG backward-compatible HDR compression formats have already been proposed in the state-of-the-art, they often do not explicitly take into account the perceptual quality, nor context, content, and environmental parameters in their design. Our subjective evaluations described above indicate that such parameters should be considered when selecting a tone-mapping operation to convert an HDR to an LDR image.

4.1 *Architecture of Encoder and Decoder*

Figure 10 describes the generic block diagram of the encoder architecture proposed in this article for HDR image compression, with the feature of being JPEG backward compatible while offering a more optimal solution when compared with the state-of-the-art. By JPEG backward compatibility, we mean that when the resulting bitstream is fed into a conventional JPEG decoder, the latter can decode it into an image and display the content as an LDR version of the original HDR image.

In this diagram, TMO_{c1} refers to an appropriate tone-mapping operation which converts the input HDR image into a format suitable for JPEG compression and decompression. In particular, the most widely used JPEG compression format relies on a YUV color image representation with all three components coded with eight unsigned integer bits, and where U and V components are subsampled by a factor of 2 in both horizontal and vertical directions when compared to Y component. The selection of TMO_{c1} is made by the encoder based on any consideration, and for the rest of

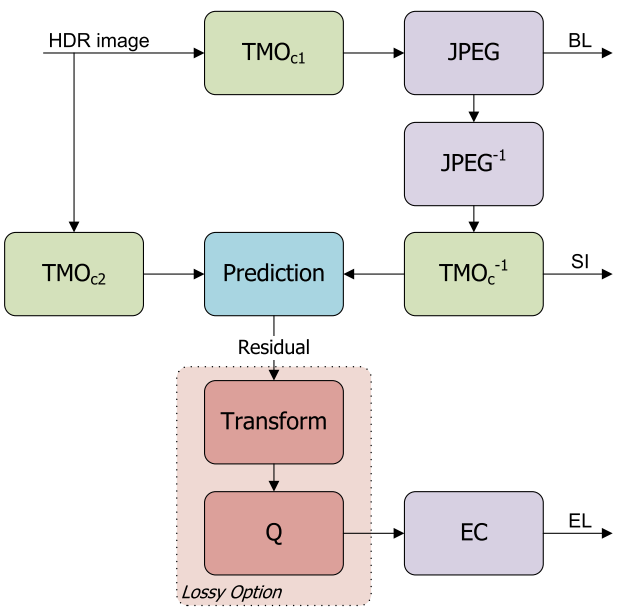

Fig. 10 Scheme of JPEG backward-compatible encoding process.

this article, we simply assume that it can be any tone-mapping algorithm. In this sense, the proposed approach is designed to cope with any tone-mapping algorithm, as is also the case with some existing solutions in the state-of-the-art such as JPEG-HDR. TMO_{c2} indicates an optional tone-mapping operation which could exist in the encoder, in cases where the representation of the input HDR image is different from the internal HDR content representation in the codec or when the application requires a different HDR representation than that of the input image. In particular, the color components and bit-depth representations of the input HDR image may be different from that used internally in the codec. This component brings an additional flexibility to the approach proposed and allows one to cope with a wide variety of HDR images and situations. JPEG and JPEG^{-1} indicate a conventional JPEG compression and decompression, respectively. In this context, no further assumption is made and the compression ratio, or any other JPEG compression parameters such as the choice of quality factor, quantization, and entropy coding tables, or any pre- and post-processing for the purpose of JPEG compression and decompression, are left to the encoder with the largest degree of flexibility. The BL indicates the baseline portion of the resulting bitstream and consists of a fully compatible JPEG format, readable by any compliant JPEG decoder. As in many extensions of JPEG, the additional bit stream necessary for decoding the HDR image will be included in the baseline JPEG format, thanks to an appropriate APPn application marker as proposed in JPEG standard.

One of the essential components of the proposed algorithm is the TMO_c^{-1}, which refers to an operation which will convert the JPEG-decoded LDR image to an appropriate HDR version. In principle, this operation can be anything in a generic situation, but the efficiency of the proposed algorithm largely depends on the choice of this operation. For instance, TMO_c^{-1} could be an algorithm aiming to minimize the residual (see Sec. 4.2). In this case, a well-designed TMO_c^{-1} will lead to very efficient compression performance. The proposed approach leaves the choice of this operation to the encoder and provides a mechanism to inform the decoder by means of a side information (SI) how the operation should be reproduced in the decoder. One can imagine that efficient TMO_c^{-1} operations, depending on the content on one hand and the TMO_{c1} on the other hand, would lead to more efficient solutions. This is a major difference between the proposed solution in this paper and other state-of-the-art solutions for HDR image compression. The output of TMO_c^{-1} is then fed into a prediction component, which will calculate the residual HDR portion of the input image. The prediction can be either in the form of a differential, a ratio, or a more complex mechanism. The residual image can then be compressed either in a lossless fashion using an appropriate entropy coding, if a lossless solution is desired, or transformed, quantized, and then entropy coded, if a lossy compression is used. The enhancement layer (EL) containing the residual HDR portion of the input image along with the SI is then embedded inside the JPEG-compressed file format using APPn marker as indicated earlier.

The proposed decoder is depicted in Fig. 11 and essentially provides the dual operations of the encoder. In addition and in order to be able to cope with context and

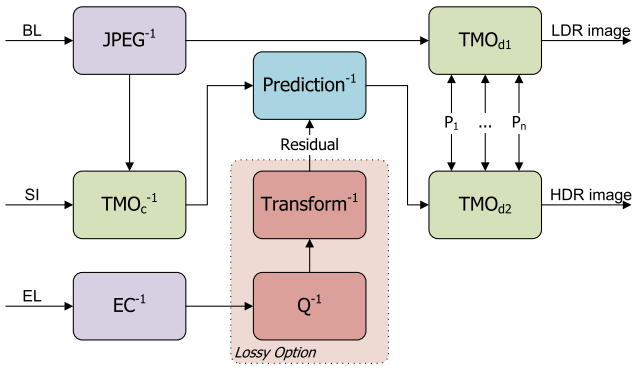

Fig. 11 Scheme of JPEG backward-compatible decoding process.

environmental parameters (denoted as $p_1 \ldots, p_n$), additional operations denoted as TMO_{d1} and TMO_{d2} can be applied as post-processing taking into account also other encoding parameters aimed at maximizing the visual quality of LDR and/or HDR-decoded images.

4.2 *Illustrative Implementation Examples and Results*

To illustrate the compression scheme presented in Figs. 10 and 11, we implemented three variants of a simple codec based on the proposed compression architecture. The following simple tone-mapping algorithms were used in each variant: (1) linear-based operator, (2) logarithm-based operator, and (3) gamma correction (with a typical gamma value 2.2). We have selected some of the simplest tone-mapping operators, because the purpose of this implementation is to illustrate the general JPEG backward-compatible compression scheme, as well as to demonstrate that knowledge of the tone-mapping used to produce the JPEG image from the original HDR can help achieving efficient compression even via simple means.

Although, the proposed solution can cope with both lossless and lossy compression, in our illustration we focus only on lossy scenario. For implementation of various components, we relied on IJG's version 6b implementation of JPEG (http://libjpeg.sourceforge.net/) and JasPer (http://www.ece.uvic.ca/frodo/jasper/) implementation of JPEG 2000 for encoding of the residual (see below for details).

The detailed process of the proposed prototype HDR image compression in a JPEG backward-compatible manner consists of the following steps:

1. Apply TMO_{c1} (simple linear-, logarithm-based operator, and gamma correction) on the HDR input image and compress with JPEG. Prior to compression, the tone-mapped image is scaled to fit into an 8-bit JPEG representation, and maximum values of each color channel are also stored for future decoding. Also, for logarithm-based TMO, an additional maximal luminance value is stored as well.
2. Application of an inverse TMO_c^{-1} to the decoded JPEG image, in order to obtain an approximation of the original HDR image denoted by HDR_j. Before applying the inverse, the decoded image is scaled back using stored maximum values for each color channel. Inverse logarithm-based TMO also uses stored maximal luminance value.
3. Conversion of both the original HDR image and HDR_j to YCbCr format illustrating TMO_{c2} and the final processing stage in TMO_c^{-1}.
4. Formation of residual of the luminance component Y as the difference (chosen for simplicity) between Y component of the original HDR image and Y component of predicted HDR_j obtained from JPEG decoding and TMO_c^{-1}.
5. Compression of the residual Y component as a gray image in 16-bits integer format using JPEG 2000 compression algorithm.

The decoding process follows a similar but dual path when compared with encoding. To make it simple, during decoding the residual for Y component of the HDR image (represented in YCbCr format) is decoded only in an enhanced resolution, while chroma components are obtained directly from the JPEG version of the image.

Since we only focus on lossy case, we compare the performance of our prototype HDR compression implementation with the state-of-the-art lossy compression algorithms capable of compressing HDR images. First is JPEG-HDR algorithm,[1] which is a well known and popular lossy JPEG backward-compatible HDR compression that uses JPEG for compressing both residual and tone-mapped versions of the HDR image. JPEG-HDR uses Reinhard's tone-mapping[5] for the LDR image, and the residual is computed as the ratio between LDR and original HDR image. At this moment, at least, until the currently under developed new JPEG standard for HDR imaging is finalized, JPEG-HDR seems to be the only publicly available JPEG backward-compatible HDR encoder and decoder. Another compression which we compare with our prototype is an adaptation of the JPEG 2000 for HDR content. Pixels in HDR image, which are in float representation, are scaled to 16-bits integers (and maximum values for each color channel are stored separately) and encoded with JPEG 2000, which is capable of compressing 16-bits images. Finally, we also use JPEG XR compression standard implemented in XnView software (http://www.xnview.com/), which supports RGBE format of HDR images directly. It is important to note, however, that neither JPEG 2000 adaptation nor JPEG XR are JPEG backward compatible.

We computed rate-distortion pairs for each variant of the proposed implementation (linear-, logarithm-based, and gamma correction tone-mapping) and for JPEG-HDR, JPEG 2000, and JPEG XR compression algorithms. For JPEG-HDR, different rates were obtained by changing compression quality value (ranging from 1 to 100) provided in the implementation by the algorithm's author Ward. For JPEG XR compression, we also varied its quantizer value.

For JPEG 2000 adaptation, we changed its compression rate value depending on a given JPEG-HDR quality value. For each JPEG-HDR quality, compression rate for JPEG 2000 was computed as a ratio between the resulted JPEG-HDR file size and a size of uncompressed HDR image. This approach makes bitrates of JPEG-HDR and JPEG 2000 compressed images as close as possible, which allows a clearer comparison of encoders' performances. We followed a similar logic with our implementation of HDR compression. We compressed LDR version of the image with the same JPEG compression quality as used for JPEG-HDR. A

residual was compressed with JPEG 2000 using a rate computed as a ratio between residual of JPEG-HDR and uncompressed grayscale image (both our compression and JPEG-HDR compute residual based on luminance only).

We ran all compression algorithms using the above settings on four images used in our subjective evaluations: "CraterLake1" [Fig. 1(a)], "Duomo1" [Fig. 1(b)], "BridgeStudios2" [Fig. 1(c)], and "Room" [Fig. 1(d)]. For each encoded image, we computed bpp value, representing the rate of rate-distortion curves, and PSNR of luminance component, representing the distortion value.

Figures 12–15 demonstrate the resulted rate-distortion curves for luminance, where JPEG-HDR is denoted as "jpeghdr," JPEG XR as "jpegxr," JPEG 2000 adaptation as "jpeg2000," variant of the proposed implementation based on logarithm as "proposed-log," based on gamma correction as "proposed-gamma," and based on linear tone-mapping as "proposed-linear."

From the figures, we can note that our proposed method outperforms (at least one of the implemented variants) all other compression schemes except for the image "Duomo" (see Fig. 13), where the linear-based implementation of our compression is on par with JPEG 2000. It is especially important to note that the proposed approach well surpasses JPEG-HDR, which is a well-used JPEG backward-compatible approach. Actually, JPEG-HDR shows the worst performance for all images, except "CraterLake1" (see Fig. 12), where it has higher PSNR than JPEG XR for the higher bitrates. The main reason why our approach shows such high efficiency compared with JPEG-HDR is that we rely on JPEG 2000, a state-of-the-art compression, for the residual instead of JPEG, a less effective and old standard used by JPEG-HDR.

To demonstrate the effect of different compression methods on color components of the HDR, we also computed similar rate-distortion curves but with distortion represented as PSNR of chrominance. To not pollute the paper with unnecessary figures, we only show the result for image "Room," as in Figs. 16 and 17, since for other images the trend is very similar. These figures demonstrate the weakness of JPEG backward compatible methods, both the proposed

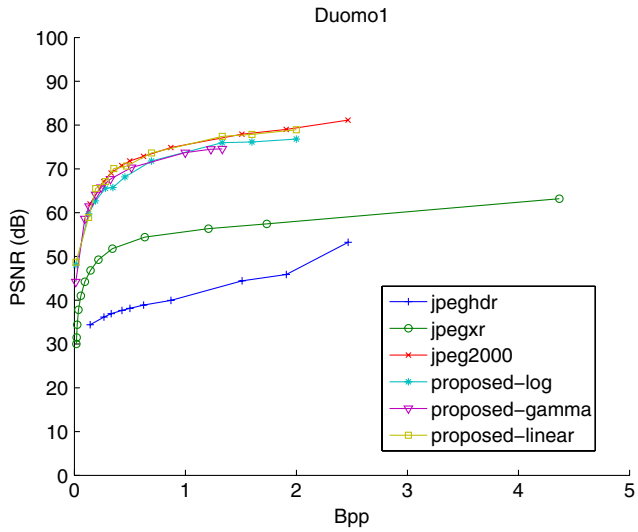

Fig. 13 Luminance rate-distortion curves for "Duomo1" image.

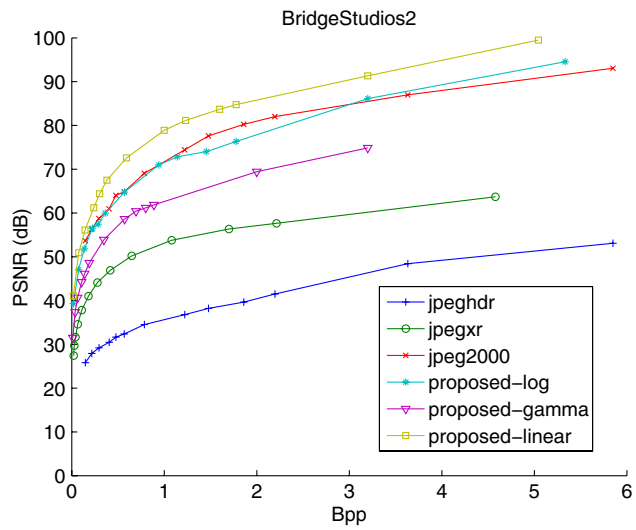

Fig. 14 Luminance rate-distortion curves for "BridgeStudios2" image.

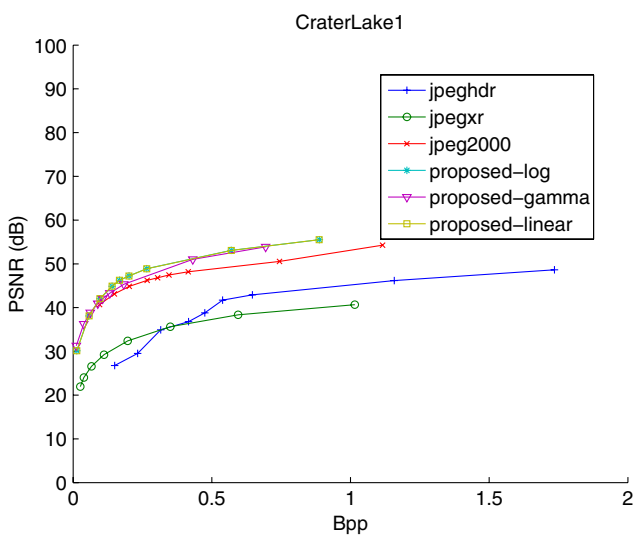

Fig. 12 Luminance rate-distortion curves for "CraterLake1" image.

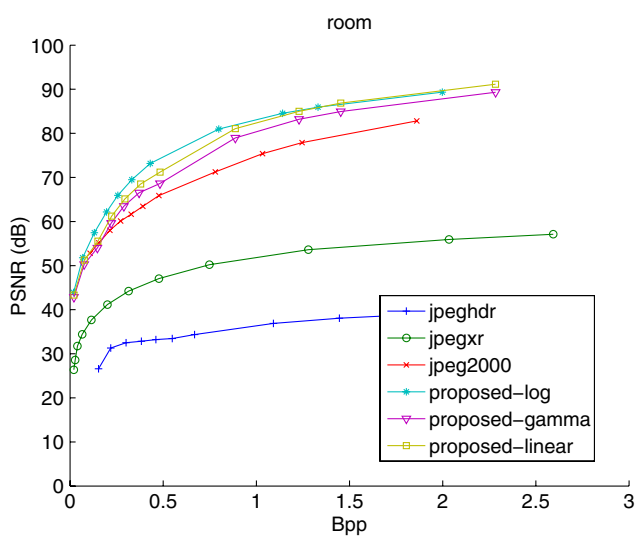

Fig. 15 Luminance rate-distortion curves for "Room" image.

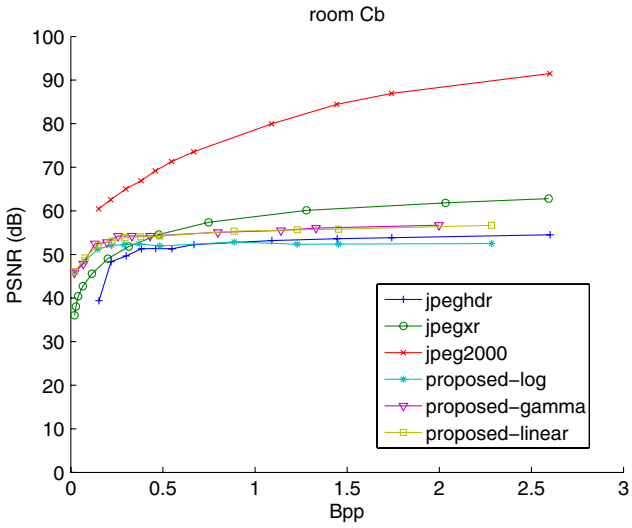

Fig. 16 Chrominance Cb rate-distortion curves for "Room" image.

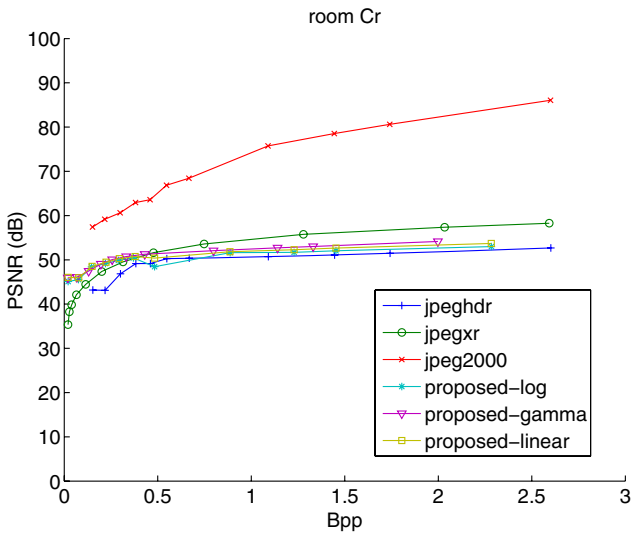

Fig. 17 Chrominance Cr rate-distortion curves for "Room" image.

approach and JPEG-HDR, as chrominance shows significantly lower PSNR compared with JPEG 2000. The reason is that, in order to preserve the dynamic range of the image, only luminance is stored as a residual by JPEG-HDR and by our proposed method, while JPEG 2000 compresses every channel of the image equally. Therefore, the combination of the following factors resulted in degradation of chrominance components encoding: backward compatibility with JPEG, efficient lossy compression, and preservation of the dynamic range of HDR image. One way to improve compression of chrominance in our proposed method is to use a tone-mapping algorithm with focus on better representation of HDR image colors in the resulted LDR version.

These results demonstrate that even by using simple tone-mapping algorithms and simple approaches to compute residuals, we can achieve a better (at least for luminance) compression efficiency compared with the state-of-the-art compression schemes capable of encoding HDR images including popular JPEG backward-compatible HDR compression algorithms JPEG-HDR.

5 Conclusion

In this paper, we proposed a generic HDR compression algorithm backward-compatible with JPEG format. The proposed solution relies on an important observation. By means of rigorous subjective assessments of various tone-mapping algorithms applied to typical HDR images and rendered in various controlled and uncontrolled environments and devices, it is shown that there is no universal tone-mapping algorithm that always stands out when compared with others. The choice of the best algorithm depends not only on the content, but also on the device used, and other environmental parameters such as backlit lighting, display type and size, environment illumination, etc. These parameters are explicitly taken into account in the proposed solution. Illustrative implementations of the proposed solutions using three simple tone-mapping algorithms show that significant compression efficiency can be obtained when compared with state-of-the-art.

The results presented in this paper can be extended in several directions. First and utmost, extension of performance evaluation to include a larger set of typical HDR images and other more sophisticated tone-mapping algorithms. Second, exploring other HDR image prediction strategies from JPEG-decoded LDR image, as well as alternative compression of residual images both in lossy and lossless fashions. Finally, a rigorous subjective evaluation of both LDR and HDR-decoded images were obtained, and comparisons to state-of-the-art from subjective quality point-of-view rather than MSE or PSNR were reported in this paper.

Acknowledgments

This work was conducted in the framework of the EC funded Network of Excellence VideoSense, with partial support from COST Action IC1005–HDRi, and COST Action IC1003–Qualinet. Special thanks to Philippe Hanhart for valuable discussions and suggestions regarding the implementation of HDR compression and computation of rate-distortion values. Help from Shuting Cai in generating some of the HDR and tone-mapped images used in subjective evaluations performed in this paper is also acknowledged.

References

1. G. Ward and M. Simmons, "JPEG-HDR: a backwards-compatible, high dynamic range extension to JPEG," in *ACM SIGGRAPH 2005*, ACM, New York, NY (2005).
2. R. Mantiuk et al., "Backward compatible high dynamic range MPEG video compression," in *ACM SIGGRAPH 2006*, Vol. 25, pp. 713–723, ACM, New York, NY (2006).
3. F. Drago et al., "Adaptive logarithmic mapping for displaying high contrast scenes," *Comput. Graph. Forum* **22**(3), 419–426 (2003).
4. R. Mantiuk, S. Daly, and L. Kerofsky, "Display adaptive tone mapping," *ACM Trans. Graph.* **27**(3), 68:1–68:10 (2008).
5. E. Reinhard and K. Devlin, "Dynamic range reduction inspired by photoreceptor physiology," *IEEE Trans. Visual. Comput. Graph.* **11**, 13–24 (2005).
6. M. D. Fairchild and G. M. Johnson "The ICAM framework for image appearance, image differences, and image quality," *J. Electron. Imag.* **13**(1), 126–138 (2004).
7. P. Korshunov and T. Ebrahimi, "A JPEG backward-compatible HDR image compression," *Proc. SPIE* **8499**, 84990J (2012).
8. P. Ledda et al., "Evaluation of tone mapping operators using a high dynamic range display," in *ACM SIGGRAPH*, ACM, New York, NY (2005).
9. A. Yoshida et al., "Perceptual evaluation of tone mapping operators with real-world scenes," *Proc. SPIE* **5666**, 192–203 (2005).

10. S. H. Park and E. D. Montag, "Evaluating tone mapping algorithms for rendering non-pictorial (scientific) high-dynamic-range images," *J. Vis. Commun. Image Represent.* **18**(5), 415–428 (2007).
11. J. Kuang et al., "Evaluating HDR rendering algorithms," *ACM Trans. Appl. Percept.* **4**(2), 1–9 (2007).
12. M. Čadík et al., "Evaluation of HDR tone mapping methods using essential perceptual attributes," *Comput. Graph.* **32**(3), 330–349 (2008).
13. B. Annighöfer, T. Tajbakhsh, and R. R. Grigat, "Prediction of results from subjective evaluation of real-time-capable tone-mapping operators applied to limited high-dynamic-range images," *J. Electron. Imag.* **19**(1), 011015 (2010).
14. Z. Mai et al., "Subjective evaluation of tone-mapping methods on 3D images," in *Int. Conf. Digital Signal Processing*, pp. 1–6, IEEE, Piscataway, NJ (2011).
15. G. Ward, Chapter "Real pixels," in *Graphics Gems II*, pp. 80–83, Academic Press, Waltham, Massachusetts (1991).
16. G. W. Larson and R. Shakespeare, *Rendering with Radiance: The Art and Science of Lighting Visualization*, Booksurge LLC, North Charleston, SC (2004).
17. T. Richter, "Backwards compatible coding of high dynamic range images with JPEG," in *Data Compression Conference (DCC)*, pp. 153–160, IEEE Computer Society, Washington, DC (2013).
18. M. Chen et al., "JPEG compatible coding of high dynamic range imagery using tone mapping operators," in *Picture Coding Symposium 2006*, pp. 22–28, Curran Associates Inc., Red Hook, NY (2006).
19. E. Reinhard et al., *High Dynamic Range Imaging: Acquisition, Display, and Image-based Lighting*, 2nd ed., The Morgan Kaufmann Series in Computer Graphics, Elsevier (Morgan Kaufmann), Burlington, MA (2010).
20. *Methodology for the Subjective Assessment of the Quality of Television Pictures*, International Telecommunication Union, Geneva, Switzerland (2012).
21. R. A. Bradley and M. E. Terry, "Rank analysis of incomplete block designs I: the method of paired comparisons," *Biometrika* **39**(3/4), 324–345 (1952).

Pavel Korshunov is a postdoctoral researcher in the Multimedia Signal Processing Group at EPFL since 2011. He received his PhD in computer science from National University of Singapore and Dipl. Sci. (MSc) in mathematics from Saint-Petersburg State University (Russia). He is a recipient of ACM TOMCCAP Nicolas D. Georganas Best Paper Award in 2011. He is involved in EC funded Video Sense project (privacy in video surveillance), as well as, COST actions high dynamic range imaging (HDRi) and quality of experience in multimedia (Qualinet). His research interests include computer vision and video analysis, video streaming, video and image quality assessment, crowd sourcing, high dynamic range imaging, and privacy issues in video surveillance systems.

Touradj Ebrahimi is currently a professor at EPFL heading its Multimedia Signal Processing Group. He has been the recipient of various distinctions and awards, such as the IEEE and Swiss National ASE Award, the SNF-PROFILE grant for advanced researchers, seven ISO certificates for key contributions to MPEG-4, JPEG 2000, JP Search, and JPEG XR standards, and the IEEE Transactions on Consumer Electronics Best Paper Award. He is the head of the Swiss delegation to MPEG, JPEG, and SC29 and chair of the Advisory Group on Management in SC29. His research interests include still, moving, and three-dimensional image processing and coding, visual information security, new media, and human computer interaction.