Two-Layer Coding Algorithm For High Dynamic Range Images based on Luminance Compensation

Masahiro Okuda, Nicola Adami

The University of Kitakyushu, 1-1 Hibikino, Wakamatu, Kitakyushu, Fukuoka, 808-0135, Japan
University of Brescia, Via Branze, 38, 25123 Brescia, (BS) Italy

Abstract

A two layer coding algorithm for high dynamic range images is discussed. In the first layer, a low dynamic range image is encoded by a conventional codec, and then the residual information that represents the difference between an original and the decoded images in the first layer is encoded in the second layer, which realizes compatibility with conventional image file formats. Our method utilizes the approximation of an inverse tone mapping function that reduces the high dynamic range to a displayable range. Our algorithm significantly improves a compression performance, compared to conventional methods.

Key words: High Dynamic Range Images, Two-layer Encoding, Inverse Tone Mapping Function

1 Introduction

Photographic films, negatives and positives, typically have a higher dynamic range and larger color gamut than typical display devices such as CRT, LCD and printers (1). Some of today 's analog film scanners captures digital images with larger dynamic range than standard 24 bit. CCD or CMOS sensors in most of today's cameras also can capture wider dynamic range of scene and

Email address: okuda-m@env.kitakyu-u.ac.jp, nicola.adami@ing.unibs.it (Masahiro Okuda, Nicola Adami).
URL: vig.is.env.kitakyu-u.ac.jp/english (Masahiro Okuda, Nicola Adami).

additionally many cameras can output raw images as well as 24 bit images compressed by JPEG, which have 12 to 16 bits for each RGB color component.

By adapting lights in any viewing condition, the human visual system can perceive a wider range of radiance (about 14 $\log_{10}$ units) than the one that can be captured by these sensors. In the last decade, to capture the high dynamic range of natural scene brightness, techniques have been proposed based on the multi-exposure image principle (2) - (6). A work done by Mann et al. (2) has proposed a method for merging multiple photographs shot with different exposures, which realizes very high dynamic ranges. Devebec et al. (3) has also proposed a method to create High Dynamic Range (HDR) Images (e.g. contrast ratio of $10^{10} : 1$) and applied to high quality image based lighting.

Instead of merging low dynamic range images off line, some sensors that can directly capture the full dynamic range of a scene are being developed (such as Viper, SMal) (7). In addition, similar researches has been done for the development of vision chips, (8), (9), which realizes 120-200 [dB] of dynamic range by controlling shutter speed on line.

The HDR imaging inspires many applications such as high quality CG rendering, in-vehicle sensors, camera surveillance, digital negative developments, etc, but it is not displayable for conventional output devices. Many researches have proposed "tone mapping operations", in which the dynamic ranges of the HDR images are reduced to displayable ranges (10)-(17). These operations aim at reducing the high dynamic range without loss of detail. Thus although these discard a significant amount of information, one may estimate the loss by utilizing the obtained low dynamic range images.

Since the sizes of the HDR images are often huge, development for functional compression is one of the research topics. The image coding standard, JPEG 2000, provides seamless compression from 1 to 16 bits per color channel (18). Ruifeng et al.'s scheme (20) verifies the validity of JPEG2000 for the high dynamic range images. An HDR video compression scheme which uses MPEG2 has also been proposed (22). Spaulding (24), (25) proposes a two layer encoding for gamut extended images. In the first layer, an image with clipped gamut is encoded. And then in the second layer, the residual information that represents the difference between the original gamut extended image and the decoded image in the first layer is encoded. The main advantage of this approach is that the format is backward compatible to existing file formats, and no extra efforts are needed to extract the 24 bit image. In the field of Computer Graphics, similar concepts are adopted for the high dynamic range image compression (26), (27).

In this paper, we extend and elaborate on the conventional two-layer coding methods (24) - (27). We approximate a tone mapping function that reduces

dynamic ranges of the HDR images to displayable ranges by a single function. We encode the difference between the HDR image and its tone mapped version, indicated as Low Dynamic Range (LDR) Image. This approximation significantly improves coding efficiency.

In Section 2, previous work related to our method is summarized. The outline of the proposed algorithm is explained in Section 3. The approximation of the tone mapping function, that is a key of our method, is discussed in Section 4, and some details of the adopted compression scheme is introduced in Section 5. Finally we evaluate the validity of our method with some experimental results.

2 Previous Work

There are some types of file formats designed for the HDR images. The well-known RGBE format spends one byte each for mantissa of RGB colors and one common exponent (E) component, that is $E = \lceil \log_2(\max(R_W, G_W, B_W)) + 128 \rceil$. LogLuv format(28) converts RGB colors to the logarithm of luminance and the u, v channels, and then uses 16 bits for the luminance, 8 bit each for the u and v channels. In these two formats, each channel is then losslessly compressed by the run length coding. OpenEXR(29) represents each of the RGB channels by the sign bit, mantissa, and exponent with 16 bits, similar to IEEE floating point. This achieves "almost lossless" representation and have options for some compression schemes, such as run-length coding and wavelet-based lossless coding. These three formats intend to preserve as much information as possible, and do not exploit correlation and human visual system very much.

The HDR images often contain a large amount of information and hence it is time-consuming to retrieve them from a storage device and/or to download them through the network. Thus the demand for lossy compression that achieves higher rate-distortion performance increases. The conventional HDR lossy compression schemes are categorized in two types: one layer and two layer compressions. The JPEG-HDR(26) proposed by Ward and Simmons is a first attempt for the two layer compression, which has a compatibility with the conventional 24 bit image formats such as JPEG. That is, a tone mapped 24 bit image is first encoded by JPEG and then the ratio between the image and the original HDR is encoded and stored in a user-available buffer of the JPEG wrapper. The decoder can not only extract the conventional JPEG image, but it can also recover the HDR image from the side information. Since the operations to create the LDR from the HDR images consists of the tone mapping, gamut clipping, quantization, and these operations are often complicated and may not be open for the public, the compatibility is preferable in many applications. A similar idea has been previously proposed to encode

raw images of digital cameras by Spaulding (24) and (25). They extend the dynamic range and color gamut by applying an inverse nonlinearity to the LDR images, and then encode the difference between the extended images and the original HDR images. In the conventional two layer algorithms, they pay little attention to tone mapping operation that transforms the HDR image to the 24 bit version. To our knowledge, only (27) tackles this problem. They compute a reconstruction function implemented with Look Up Table (LUT) that represents the inverse tone mapping and then compensate the difference of the two images. Although the LUT-based method approximates the tone mapping curve well, the approximation is not smooth. The lack of smoothness may yield extra energy in high frequency, which is not desirable in a sense of compression efficiency.

Although the one layer coding is superior in a sense of compression efficiency, these two layer coding has an advantageous property of the compatibility with the 24 bit image formats, that is, 24 bit image is encoded by the existing method such as JPEG in the first layer. Mantiuk et al propose a HDR video compression scheme based on the MPEG codec (22). They derive an optimal quantization strategy using the threshold versus intensity functions of the human visual systems, and quantize the full dynamic ranges to 8 bit before MPEG encoding. Xu et al. (20) apply JPEG2000 to the HDR images. They transform an image to logarithm domain, quantize to 12-bit integers, and then input it to the JPEG2000 encoder. Although their method does not have the compatibility with the 24 bit images, its compression efficiency outperforms others.

Our framework is a similar to (26) and (27), that is, the proposed algorithm also consists of the two layers with the 24-bit image compatibility. The rest of this paper describes our algorithm based on the inverse tone mapping and shows our method outperforms the conventional methods.

3 Summary of the Algorithm

3.1 Brightness v.s. Luminance Nonlinearity

Camera raw images, in which current values obtained from CMOS or CCD sensors are stored in pixels without any image processing, have 12 to 16 bits for each color channel. The pixel values in the raw image are proportional to actual luminance. In the field of CG, they propose several methods to create higher dynamic range (HDR) images by combining the series of multi-exposure photos(3)-(4). For each pixel, radiance is calculated by estimating the camera response curve and compensating the in-camera nonlinear operations. Thus

these pixel values are also proportional to luminance. Thus the operation that maps the HDR image to a 24 bit 'output referred' image more or less resembles the work of the human that perceives brightness.

It is widely agreed that there is a nonlinear relationship between an amount of sensation and intensity of lights, or in other words, brightness human perceives and actual luminance are not linearly related. There are many experimental results for the approximation of the nonlinearity. The well known Stevens' power law indicates that the relationship between them is modeled by

$$y = k(x - C)^n \tag{1}$$

where y and x are brightness that human retina perceives and input luminance, respectively. This model provides accurate approximation of the visual response in a limited range, but is not reliable outside the range where the response is saturated. To model the response of the rod and three types (LMS) of cones on Retina in such a wider range, a s-shape function is utilized (Fig.1).

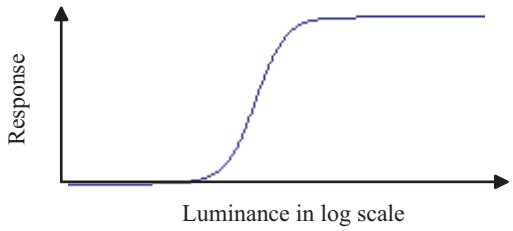


Fig. 1. Response of cones/rods in an adaptation luminance: The curve shifts horizontally depending on different adaptation luminance.

Most of the tone mapping operators (and camera film developments) take into account the nonlinearity of the human perception models. Thus, if the tone mapping operations are well estimated by a simple function, the HDR images can be reconstructed by the tone mapped version. This is the key of our method.

3.2 Outline of Proposed Method

The proposed method realizes the two-layer compression similar to (26). That is, starting from the original HDR, a tone mapped image (LDR) is generated or given as input, and then encoded by using JPEG. A residue information is also encoded in order to enable the reconstruction of the HDR from the knowledge of its LDR representation. The most significant difference from the conventional methods is the use of an Inverse Tone Mapping Function (ITMF). Fig.2 depicts the encoding and decoding steps. Our method assumes that an

original HDR and its tone mapped LDR image are given as inputs. The LDR image is encoded by JPEG, followed by JPEG decoding. The intensity Y is transformed by the ITMF to expand the reduced dynamic range. Then we compare it with the intensity of the HDR image and calculate a residue. The residue is encoded by the wavelet-based image compression. The ITMF used here is image-dependent function. Thus the parameters to describe the ITMF are calculated using the given two images on line and then they are sent to the decoder as well as the encoded LDR and the residue. At low bit rates, only the intensity of the HDR image is encoded and the chroma is estimated by the LDR image. At higher bit rate, color difference between the HDR and the LDR with color compensation is compressed by the wavelet. Each step is explained in more detail in the following sections.

4 Image Dependant Inverse Tone Mapping Function

4.1 Inverse Tone Mapping Function

We call the procedure for converting an HDR to LDR image, the Tone Mapping (TM), and the inverse procedure of the TM, the Inverse Tone Mapping (ITM). In this section, we treat the problem of finding the Inverse Tone Mapping Functions (ITMF), given an image pair of HDR and tone mapped LDR. Our attempt here is to closely approximate the relationship between HDR and tone mapped LDR by ITMF, which will improve a coding efficiency.

Many tone mapping algorithms have already been proposed. In some methods, global operations are applied to the HDR image. Some other algorithms locally map each pixel of HDR to LDR. (A comprehensive review of the existing tone mapping operations can be found in (7).) Although the local tone mapping operations can not inherently described by a function, many of them can be approximated by the s-shape. Thus regardless of which type of mapping is actually used to create the LDR image, in this paper, we approximate it only by a single function.

As is pointed out in Fig.1, the response of the retina is well approximated by the s-shape function. To fairly approximate the s-shape, we have adopted the Hill function after testing other functions such as polynomials, sigmoid functions. The formula of the Hill function f is given by

$$y = f(x) = \frac{x^n}{x^n + k^n}. \tag{2}$$

where the function is normalized by 1. This function is also known as the

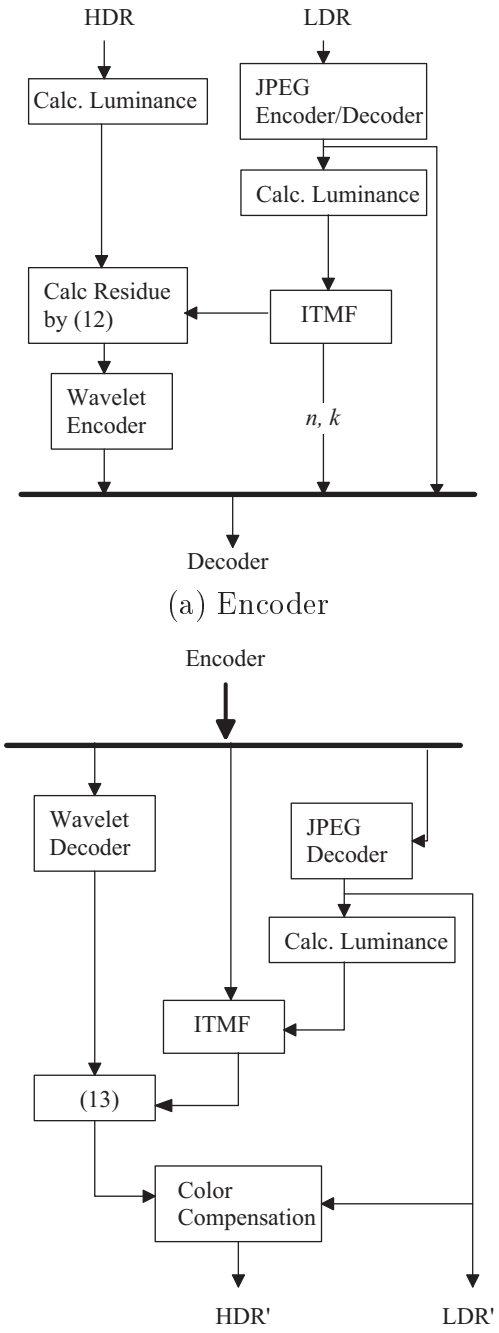


Fig. 2. Outline of Our Method

Naka-Rushton equation or the Michaelis-Menten equation.

The function has some desirable properties:

(1) By using the function one can easily express the s-shape with long saturated regions
(2) For $x > 0$ and $k > 0$, the function monotonically increases (see Ap-

pendix).
(3) Response is controlled by only two parameters, n and k.
(4) In the log domain, the inverse of the function is linear to the two parameters $\log(k)$ and $1/n$, which will be further discussed later.

The first two properties are preferable for the approximation of the response of the retina, while the others make it easy to design and handle the inverse tone mapping functions. Further discussion on this function in a viewpoint of human perception can be found in (30) and (31).

4.2 Closed Form Solution

Here we derive the ITMF to convert the LDR to HDR, which is optimal in least squares sense. The inverse of f in (2), $g = f^{-1}$ is given by

$$x = g(y) = k\left(\frac{y}{1-y}\right)^{\frac{1}{n}}, \tag{3}$$

Then the conversion from LDR to HDR image by the ITMF is expressed as

$$\mathcal{H}'_y(i,j) = g(\mathcal{L}_y(i,j)) = k\left(\frac{\mathcal{L}_y(i,j)}{1-\mathcal{L}_y(i,j)}\right)^{\frac{1}{n}}, \tag{4}$$

where $\mathcal{H}'_y(i,j)$ and $\mathcal{L}_y(i,j)$ is the luminance Y of the reconstructed HDR image and the LDR image, respectively. Then we define the error between $\mathcal{H}'_y(i,j)$ and the original HDR image $\mathcal{H}_y(i,j)$ by

$$e(i,j) = \log(\mathcal{H}_y(i,j)) - \log(\mathcal{H}'_y(i,j)) \tag{5}$$

The error in the log domain put more weights on low luminance. This metric matches human perceptions more and suppress the overestimate of high luminance. Moreover it linearizes the function w.r.t. the parameters $\log(k)$ and $1/n$ as follows.

$$\log(\mathcal{H}'_y(i,j)) = \log(k) + \frac{1}{n}\log\left(\frac{\mathcal{L}_y(i,j)}{1-\mathcal{L}(i,j)}\right) \tag{6}$$

Then the design problem of ITMF is stated as the minimization of the mean squared error of (5),

$$
\begin{aligned}
E &= \frac{1}{M}\sum_{i,j} e(i,j)^2 \\
&= \frac{1}{M}\sum_{i,j}\left\{\log(\mathcal{H}_y(i,j)) - \log(k) - \frac{1}{n}\log\left(\frac{\mathcal{L}_y(i,j)}{1-\mathcal{L}_y(i,j)}\right)\right\}^2,
\end{aligned} \tag{7}
$$

where M is the number of pixels. Since (7) is a quadratic form with respect to $\log(k)$ and $1/n$, the optimum solution that minimizes (7) is uniquely determined. In the end, solving the two equations

$$
\begin{aligned}
\frac{\partial E}{\partial(\log(k))} &= 0, \\
\frac{\partial E}{\partial(1/n)} &= 0
\end{aligned} \tag{8}
$$

yields a closed form solution

$$
\begin{aligned}
k &= \exp\left\{\frac{(\sum_{i,j} Y_{i,j}^2)(\sum_{i,j} X_{i,j}) - (\sum_{i,j} Y_{i,j})(\sum_{i,j} X_{i,j}Y_{i,j})}{M(\sum_{i,j} Y_{i,j}^2) - (\sum_{i,j} Y_{i,j})^2}\right\} \\
n &= \frac{M(\sum_{i,j} Y_{i,j}^2) - (\sum_{i,j} Y_{i,j})^2}{M(\sum_{i,j} X_{i,j}Y_{i,j}) - (\sum_{i,j} Y_{i,j})(\sum_{i,j} X_{i,j})},
\end{aligned} \tag{9}
$$

where

$$
X_{i,j} = \log(\mathcal{H}_y(i,j)), \quad Y_{i,j} = \log\frac{\mathcal{L}_y(i,j)}{1-\mathcal{L}_y(i,j)}. \tag{10}
$$

In the tone mapping procedures, to reduce the dynamic range of the HDR images, special attention should be paid for bright area (luminance more than a few thousands in cd/m^2), where human sensitivity is saturated. The simplest operation is to clip the values higher than a threshold and they are mapped to 255. To improve the quality, more sophisticated operations have been proposed(7), in which the local operators determine the mapping depending on the neighbors of each pixel. Due to these, the approximation described in this section often fails in the bright regions, especially in the case that the LDR images are coarsely quantized by the JPEG. Fig.3(a) shows the plot of the LDR pixel values versus the HDR pixel values in corresponding positions. One can see from this figure that the variance of the HDR pixel values increases as the luminance gets high. Due to this, in the bright area the optimal approximation in least squares sense does not necessarily give an optimal compression performance. Thus, in our algorithm we simply approximate the bright area by a first order polynomial. Fig.3(b) shows the obtained ITMF. In this example, the region $\mathcal{L}_y(i,j) \in [0,\ a]$ is approximated by the hill function and the linear mapping is done in $\mathcal{L}_y(i,j) \in (a,\ 1]$ (in this paper, we suppose

that LDR images are normalized to 1). The boundary a is determined as follows. Fist we divide the region of the LDR intensity $[0,\ 1]$ to Q bins. For each bin, the variance of HDR intensity in corresponding positions, $\sigma_q^2 = var(H_q)$, $q = 1, 2, \cdots, Q$, where

$$H_q = \{\mathcal{H}_y(i,j)|\ \ i,j,\ \ (q-1)/Q \leq \mathcal{L}_y(i,j) < q/Q\} \qquad q = 1, 2, \cdots, Q.$$

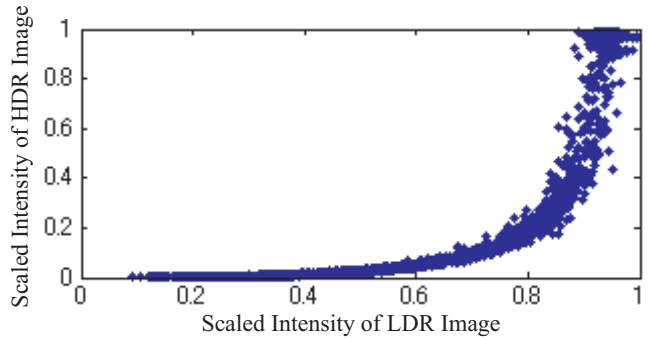
Next we find the smallest value q^* that satisfies

$$\sigma_{q^*}^2 > c \cdot \sigma_{hdr}^2,$$

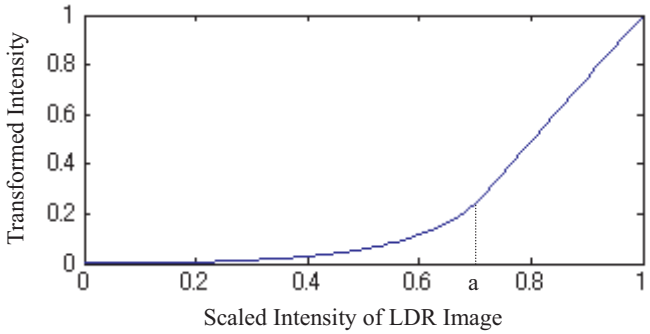
and then we set

$$a = (q^* - 1)/Q,$$

where σ_{hdr}^2 is a variance of the HDR image and c is a constant. For most examples, it works well with $c = 0.01$.



(a) LDR-HDR plots



(b) Inverse Tone Mapping Function

Fig. 3. Approximation by ITMF

5 HDR coding algorithm

5.1 Residue Compression

As explained in the previous section, the luminance of the LDR image is transformed by the ITMF in (4). Then the residual image $\mathcal{R}$ is created by using the original HDR image, $\mathcal{H}_y(i,j)$ and the transformed image, $\mathcal{H}'_y(i,j)$:

$$\mathcal{R}(i,j) = \left(\frac{\mathcal{H}_y(i,j)}{\mathcal{H}'_y(i,j) + \epsilon} \right)^{\gamma}, \tag{11}$$

where γ is a constant less than 1. When both of the two images, $\mathcal{H}_y(i,j)$ and, $\mathcal{H}'_y(i,j)$ have very small values, slight difference between them will be overvalued by (11) more than the case that they have large values. Thus this division process implicitly weights the small pixel values more. γ in (11) plays a role in relieving this weight effect. ϵ is introduced to prevent the values from being too large when $\mathcal{H}'_y(i,j)$ approaches to zero.

Then the image is reconstructed in the decoding steps by

$$\bar{\mathcal{H}}_y(i,j) = \mathcal{R}(i,j)^{\frac{1}{\gamma}} (\mathcal{H}'_y(i,j) + \epsilon), \tag{12}$$

where $\bar{\mathcal{H}}_y$ is the decoded image.

5.2 Color Compensation

In (26), Ward et al. point out that if the TM operation preserves hue of an HDR image, and its saturation change is estimated, then the chroma of the original HDR can be fairly recovered from its LDR by the saturation compensation. The Ward et al's compensation method works well. However it cannot be directly applied to our method since we use the inverse tone mapping operation. Instead our algorithm adopts different two color compensation schemes: one is modified Ward et al's method and the other is the proposed color compensation based on polynomial approximation. We first introduce the modified Ward's scheme for comparison.

5.2.1 *Modified Ward et al's Compensation Scheme*

In (26), the compensated saturation is given by

$$S_c(i,j) = \alpha \left\{ S_{\mathcal{L}}(i,j) \right\}^{\beta},$$

where $S_{\mathcal{L}}$ is the saturation of the LDR image. For the definition of the saturation, we follow (26):

$$S \equiv 1 - \min\{R,G,B\}/Y.$$

In this framework, we basically adopt the Ward et al's method described in (26). Only difference from (26) is that we calculate the optimal parameters α and β that minimizes

$$\begin{aligned} E_s &= \textstyle\sum_{i,j} \{ \log S_{\mathcal{H}}(i,j) - \log S_c(i,j) \}^2 \\ &= \textstyle\sum_{i,j} \{ \log S_{\mathcal{H}}(i,j) - \log \alpha - \beta \log S_{\mathcal{L}}(i,j) \}^2. \end{aligned} \tag{13}$$

Since this error is a simple quadratic form of the two parameters $\log \alpha$ and β, then the minimization of (13) yields the closed form solution

$$\begin{aligned} \alpha &= \exp \left\{ \frac{(\sum_{i,j} Y_{i,j}^2)(\sum_{i,j} X_{i,j}) - (\sum_{i,j} Y_{i,j})(\sum_{i,j} X_{i,j} Y_{i,j})}{M(\sum_{i,j} Y_{i,j}^2) - (\sum_{i,j} Y_{i,j})^2} \right\} \\ \beta &= \frac{M(\sum_{i,j} X_{i,j} Y_{i,j}) - (\sum_{i,j} Y_{i,j})(\sum_{i,j} X_{i,j})}{M(\sum_{i,j} Y_{i,j}^2) - (\sum_{i,j} Y_{i,j})^2}, \end{aligned} \tag{14}$$

where

$$X_{i,j} = \log S_{\mathcal{H}}(i,j), \quad Y_{i,j} = \log S_{\mathcal{L}}(i,j). \tag{15}$$

5.2.2 *Proposed Color Compensation*

Assume the relation between that RGB values of the HDR and LDR images is approximated by a polynomial,

$$y = p(x) = \sum_{n=0}^{N} c_n x^n, \tag{16}$$

then we simply apply this to the LDR image, that is,

$$\mathcal{L}_r^c(i,j) = p(\mathcal{L}_r(i,j)), \quad \text{for channel R} \tag{17}$$

The function p can be determined by a simple least squares method that minimizes the following cost function, that is, total summation of the weighted suqared errors for the red, green and blue channels:

$$\begin{aligned} E_p &= \textstyle\sum_{x\in\{r,g,b\}} \sum_{i,j} \{w(i,j)\left(\mathcal{H}_x(i,j) - \mathcal{L}_x(i,j)\right)\}^2 \\ &= \boldsymbol{c}_n^T \boldsymbol{L}^T \boldsymbol{L} \boldsymbol{c}_n - 2\boldsymbol{c}_n^T \boldsymbol{L}^T \boldsymbol{H} + \boldsymbol{H}^T \boldsymbol{H}, \end{aligned} \tag{18}$$

where w is a weight function defined on the pixels. The matrices of (18) are defined by

$$\boldsymbol{c} = \begin{bmatrix} c_0 & c_1 & \cdots & c_N \end{bmatrix}^T,$$

$$\boldsymbol{L} = \begin{bmatrix} \boldsymbol{L}_r^0 & \boldsymbol{L}_r^1 & \cdots & \boldsymbol{L}_r^N \\ \boldsymbol{L}_g^0 & \boldsymbol{L}_g^1 & \cdots & \boldsymbol{L}_g^N \\ \boldsymbol{L}_b^0 & \boldsymbol{L}_b^1 & \cdots & \boldsymbol{L}_b^N \end{bmatrix},$$

$$\boldsymbol{H} = \begin{bmatrix} \boldsymbol{H}_r^T & \boldsymbol{H}_g^T & \boldsymbol{H}_b^T \end{bmatrix}^T,$$

whre $\boldsymbol{L}_r^n$, $\boldsymbol{L}_g^n$, $\boldsymbol{L}_b^n$, $\boldsymbol{H}_r$, $\boldsymbol{H}_g$, and $\boldsymbol{H}_b$ are

$$\boldsymbol{L}_{r,g,b}^n = \begin{bmatrix} w(i,j)\mathcal{L}_{r,g,b}^n(0,0) & w(i,j)\mathcal{L}_{r,g,b}^n(0,1) & w(i,j)\mathcal{L}_{r,g,b}^n(0,2) & \cdots \end{bmatrix}^T,$$

$$\boldsymbol{H}_{r,g,b} = \begin{bmatrix} w(i,j)\mathcal{H}_{r,g,b}(0,0) & w(i,j)\mathcal{H}_{r,g,b}(0,1) & w(i,j)\mathcal{H}_{r,g,b}(0,2) & \cdots \end{bmatrix}^T.$$

In the end the coefficients are determined by solving a linear equation

$$\left(\boldsymbol{L}^T \boldsymbol{L}\right) \boldsymbol{c}_n = \boldsymbol{L}^T \boldsymbol{H}$$

We adopt the Gaussian function for the weight w, that is, small pixel values are weighted more than large values. Note that its coefficients are transformed to the decoder as a side information.

After all the colors are compensated by (17), we decompress the dynamic range as follows

$$\bar{\mathcal{H}}_r'(i,j) = \mathcal{L}_r^c(i,j) \frac{\bar{\mathcal{H}}_y'(i,j)}{\mathcal{L}_y^c(i,j)}, \quad \text{for channel R,} \tag{19}$$

where $\bar{\mathcal{H}}_y'(i,j)$ is the decoded luminance Y of the HDR image and $\mathcal{L}_y^c(i,j)$ is the luminance of the LDR image after the compensation (17). The same

procedure is done for G and B channels as well. This procedure maps the color of the LDR to the HDR image.

In a low bit rate, one can spend all bits for the intensity $\mathcal{H}_y$, and the chroma is compensated by (19). In high bit rates, however, it often gives better results to encode chroma as well. The color encoding is required when the tone mapping operator drastically changes the colors. Thus in the high bit rates, our scheme encodes the chroma difference between the original HDR and color compensated HDR given by (19) as well.

6 Experimental Results

We have tested dozens of HDR images collected from some web sites, many of which are frequently used as sample images. Since the HDR images may have very bright and dark areas in that human do not even perceive luminance change, conventional error metrics (e.g, mean squared error, snr) are not reliable. Fig. 4 illustrates two HDR images with a same snr. One can see much degraded area in the right image. This often happens since errors in highlight area are overestimated by the snr. Thus we instead use two other metrics, the Mean Distance (MD) in CIELAB color space to evaluate color appearance differences, and the Mean Squared Error (MSE) with Daly's nonlinearity. The

Fig. 4. Reconstructed raw images with a same PSNR (To display, the images are gamma corrected).

CIELAB (also called CIE 1976 L*a*b*) is a well known uniform color space, which is one of the most reliable evaluation for color images. We use the mean of the absolute difference in this space for color difference evaluation, that is,

$$E_{lab} = \frac{1}{M} \sum_{i,j} \sqrt{(L_1(i,j) - L_2(i,j))^2 + (a_1(i,j) - a_2(i,j))^2 + (b_1(i,j) - b_2(i,j))^2}. \quad (20)$$

To evaluate only the luminance, we first transform the intensity with Daly's nonlinearity (32)

$$DN(x) = \frac{x}{x + 12.6x^{0.63}} \tag{21}$$

Then we calculate the mean squared error

$$E_{mse} = \frac{1}{M}\sqrt{\sum_{i,j}\{DN(Y_1(i,j)) - DN(Y_2(i,j))\}^2}. \tag{22}$$

First we test the performance on several images (see Fig.5), and compare it with Ward et al's JPEG-HDR (26). This method encodes the luminance residue, while colors are recovered from saturation information of the LDR image. For fair comparison, we use exactly the same jpeg images as the LDR images for input. We spend all the bits for luminance compression, and we do not encode other color channels. Fig.6 shows comparisons for some images [1]. At very low bit rates, our method is comparable or even worse, but at high bit rates our method gains significant improvements.

Fig.7 shows a comparison with the most recent HDR coding method (27). This method is designed mainly for HDR video compression, but it also gives high performances for still images. The plots indicate the results of "Memorial" images [2]. In these experiments, we use Reinhard et al.'s global tone mapping operator (12) to make LDR images, and we allocate 75 % of bits for the intensity and the rest of them for chroma difference encoding. From these figures, it can be seen that our algorithm performs better than (27).

We compare the two methods described in Section5.2. The solid line in Fig.8 depicts the result of the proposed polynomial-based color compensation. The dotted line is the one of the modified Ward et al's method with the optimal gamma. In this simulation, 70 % of residual bits are used for the intensity and the rest are used for chroma difference. Especially at high bit rates, our method gives better results.

The proposed algorithm was tested on the several tone mapping operators, (12), (13), (14), and (15). Although it can be seen that the compression performance depends on tone mapping operators especially for MD in CIELAB, the algorithm works well for a wide variety of tone mapping operations.

[1] The results are obtained using a software the authors provide
[2] These data are provided by the authors

7 Conclusion

A two layer coding algorithm for high dynamic range images has been proposed. Our approximation of the tone mapping operation using the Hill function significantly improves the coding efficiency compared to other state of the art coding methods.

Acknowledgement

We are grateful for the support of a Grant-in-Aid for Young Sciences (#14750305) of Japan Society for the Promotion of Science, fund from MEXT via Kitakyushu innovative cluster project, and Kitakyushu IT Open Laboratory. We thank Dr. Mantiuk for providing experimental data.

References

[1] J. Nakamura, " Image Sensors And Signal Processing For Digital Still Camera ", Taylor & Francis Group, 2005

[2] S. Mann and R. Picard", "On being 'undigital' with digital cameras: Extending dynamic range by combining differently exposed pictures", In Proceedings of IS&T 46th annual conference (May 1995), pp. 422–428. 1995

[3] P.E. Debevec and J. Malik. " Recovering High Dynamic Range Radiance Maps from Photographs, Proceedings of SIGGRAPH 97, Computer Graphics Proceedings, pp. 369-378 .

[4] T. Mitsunaga, and S. K. Nayer,"Radiometric Self Calibration," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Vol.1, pp.374-380, Jun, 1999.

[5] Y. Tsin, V. Ramesh, and T. Kanade, "Statistical Calibration of CCD Imaging Process, " IEEE International Conference on Computer Vision (ICCV'01), July, 2001.

[6] C. Pal, R. Szeliski, M. Uyttendaele, N. Jojic, "Probability Models for High Dynamic Range Imaging," IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04) - Volume 2 pp. 173-180, 2004

[7] E. Reinhard, S. Pattanaik, G. Ward and P. Debevec, " High Dynamic Range Imaging: Acquisition, Display, and Image-Based Lighting (Morgan Kaufmann Series in Computer Graphics and Geometric Modeling), Morgan Kaufmann Publisher 2005.

[8] X. Liu and A. El Gamal, "Synthesis of High Dynamic Range Motion Blur

Free Image from Multiple Captures," IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications, Vol. 50, No. 4, pp. 530-539, April 2003.

[9] N. Akahane, R. Ryuzaki, S. Sugawa, S. Adachi and K. Mizobuchi, "An Over 200 dB Dynamic Range Image Capture using a CMOS Image Sensor with Lateral Overflow Integration Capacitor and Current Readout Circuit in a Pixel", International Congress of Imaging Science, 2006

[10] C. Schlick, "Quantization Technique for Visualization of High Dynamic Range Pictures", Proceedings 5th Eurographics Workshop on Rendering, pp. 7-20, 1994.

[11] J. Tumblin, G. Turk, "Low Curvature Image Simplifiers (LCIS): A Boundary Hierarchy for Detail-Preserving Contrast Reduction." ACM SIGGRAPH Proceedings, Annual Conference Series, pages 83–90. August 1999.

[12] E. Reinhard, M. Stark, P. Shirley and J. Ferwerda, "Photographic Tone Reproduction for Digital Images", ACM Trans on Graphics. 21, 3, 267–276, 2002

[13] J. Tumblin, J.K. Hodgins, And B.K.. Guenter, "Two Methods for display of high contrast images," ACM Transaction on Graphics, 18 (1) pp. 56-94, 1999.

[14] S. N. Pattanaik, Jack E. Tumblin, Hector Yee, Donald P. Greenberg,"Time-Dependent Visual Adaptation for Realistic Real-Time Image Display", Proceedings of SIGGRAPH 2000, pp. 47-54, New Orleans, 23-28 July, 2000.

[15] R. Fattal, D. Lischinski, and M. Werman. Gradient domain high dynamic range compression. ACM Transactions on Graphics, pp. 249-256, July, 2002.

[16] F. Durand, and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," ACM SIGGRAPH , pp.257-266, 2002.

[17] Y. Li, L. Sharan, E. H. Adelson. "Compressing and Companding High Dynamic Range Images with Subband Architectures, " ACM Transactions on Graphics, 24(3), 2005.

[18] D. S. Taubman and M. W. Marcellin, "JPEG 2000: Image Compression Fundamentals, Standards and Practice," Kluwer International Series in Engineering and Computer Science,

[19] "JasPer Software Reference Manual (Version 1.700.0), ISO/IEC JTC 1/SC 29/WG 1, N2415, 2003

[20] R. Xu, S. N. Pattanaik, C. E. Hughes, "High-Dynamic-Range Still-Image Encoding in JPEG 2000. IEEE Computer Graphics and Applications 25(6): 57-64, 2005.

[21] S. B. Kang, M. Uyttendaele, S. Winder, R. Szaliski, "High Dynamic Range Video", ACM Transactions on Graphics, Volume 22, Issue 3 pp. 319 - 325, 2003.

[22] R. Mantiuk, G. Krawczyk, K. Myszkowski and H. P. Siedel, "Perception-Motivated High-Dynamic Range Video Encoding, ACM Trans. on

Graphics, col23, no.3 2004, pp.773-741.
[23] R. Xu, S. N. Pattanaik,and C. E. Hughes,"High Dynamic Range Image and Video Data Compression", Technical Report, School of Engineering and Computer Science, University of Central Florida, CS-TR-05-13, 2005.
[24] K. E. Spaulding, R. L. Joshi, G. J. Woolfe, "Using a residual image formed from a clipped limited color gamut digital image to represent an extended color gamut digital image", United States Patent 6301393
[25] K. E. Spaulding, G. J. Woolfe, R. L. Joshi, "Using a Residual Image to Extend the Color Gamut and Dynamic Range of an sRGB Image", The PICS Conference, p. 307-314, 2003
[26] G. Ward, and M. Simmons, "JPEG-HDR: A Backwards-Compatible, High Dynamic Range Extension to JPEG," Proceedings of the Thirteenth Color Imaging Conference, November 2005.
[27] R. Mantiuk, A. Efremov, K. Myszkowski, H.P. Seidel. "Backward Compatible High Dynamic Range MPEG Video Compression". ACM SIGGRAPH 2006
[28] G. W. Larson, "LogLuv Encoding for Full-Gamut, High-Dynamic Range Images," Journal of Graphics Tools, vol.3, 1998, pp.15-31.
[29] F. Kainz, R. Bogart, and D. Hess, "The OpenEXR Image File Format," SIGGRAPH Technical Sketches, 2003.
[30] L. Spillmann, J.S. Wener, "Visual Perception, the Neurphysiological Foundations," Academic Press, San Diego, 1990. .
[31] S. N. Pattanaik, "High Dynamic Range Imaging", in Tutorial notes, ACM-SIGGRAPH 2004, Los Angeles, Aug 2004.
[32] S. Daly, "The Visible Difference Predictor: An algorithm for the assessment of image fidelity," in A.B. Watson, editor, Digital Image and Human Vision, pages 179-206, Cambridge, MA, MIT Press, 1993.

Appendix

The Hill function (2) is rewritten as

$$f(x) = \frac{1}{1 + \left(\frac{k}{x}\right)^n}$$

The derivative of f yields

$$\frac{df(x)}{dx} = (1 + \left(\frac{k}{x}\right)^n)^{-2} \cdot k^n \cdot n \cdot x^{-(n+1)}$$

Thus, when $k > 0$ and $x > 0$, then $\frac{df(x)}{dx} = 0$ holds and the function monotonically increases.

Authors

Masahiro Okuda received the B.E. , M.E. , and Dr.Eng. in 1993, 1995, 1998, respectively from Keio University, Yokohama, Japan. He was with University of California, Santa Barbara and Carnegie Mellon University as a visiting scholar in 1998 and 1999, respectively. He has been on the faculty of The University of Kitakyushu as an associate professor of Environmental Engineering since 2001. His research interests include filter design, vision/geometry coding, and multirate signal processing.

Nicola Adami received Ph.D. in 2002 from Brescia University. Since 2004, he has been Assistant Professor at the Engineering Faculty - University of Brescia. His research interests include multimedia signal processing, adiovisual(AV) indexing, audio visual content description processing, scaling, filtering and integration. .

(a) Memorial

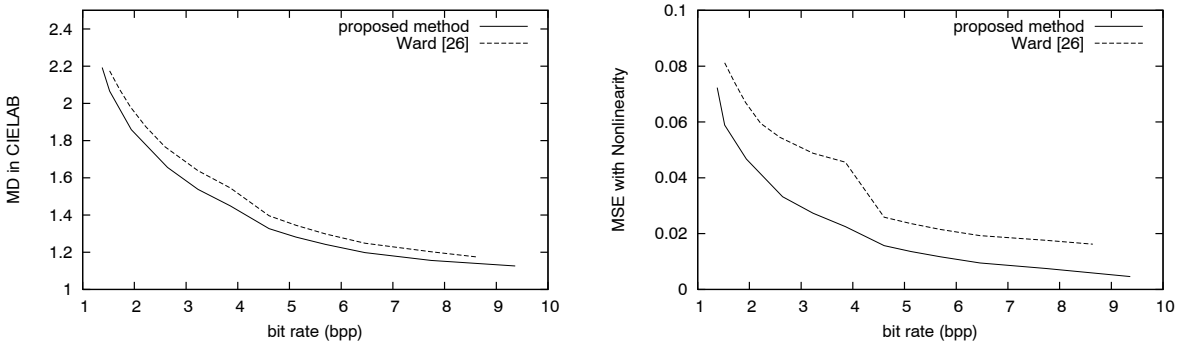
(b) Dyrham Church

(c) Apartment

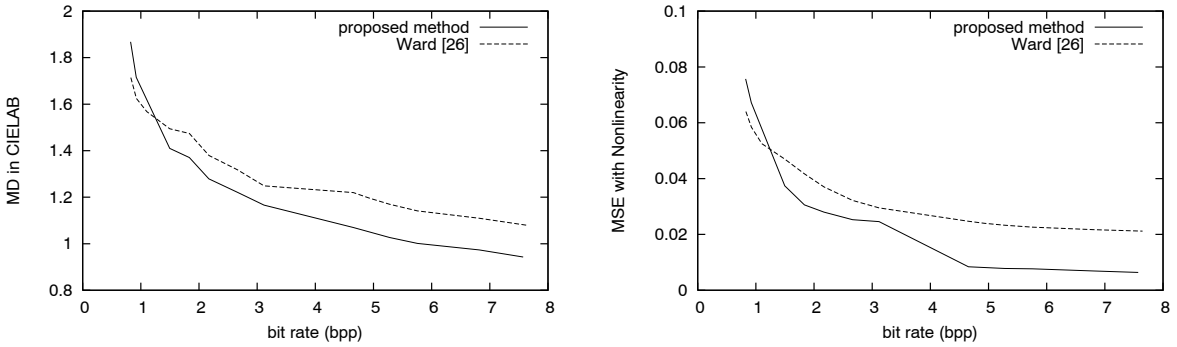
(d) BigFog

(e) Belgium

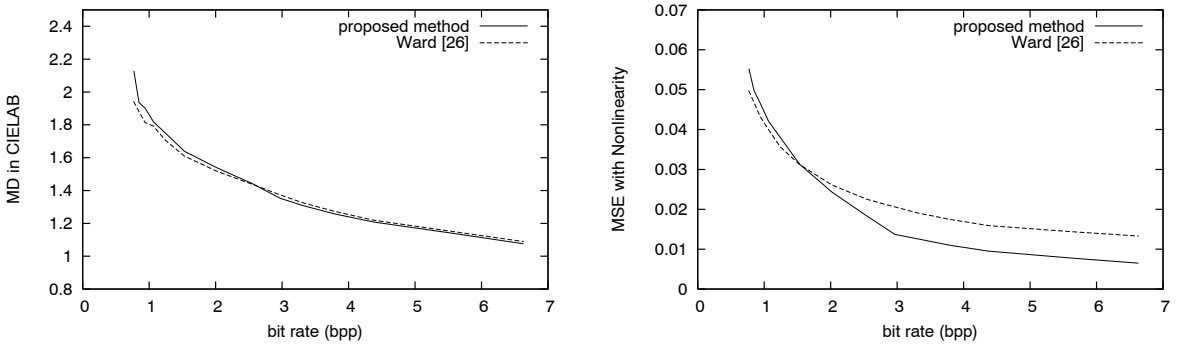
Fig. 5. Sample Images



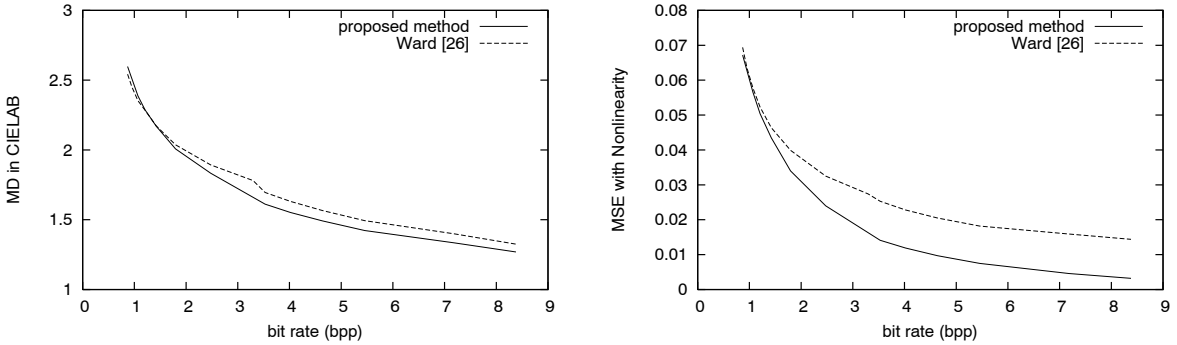
(a) Dyrham Church



(b) Apartment



(c) Big Fog



(d) Belgium

Fig. 6. Comparison with Ward et al.'s method (26)

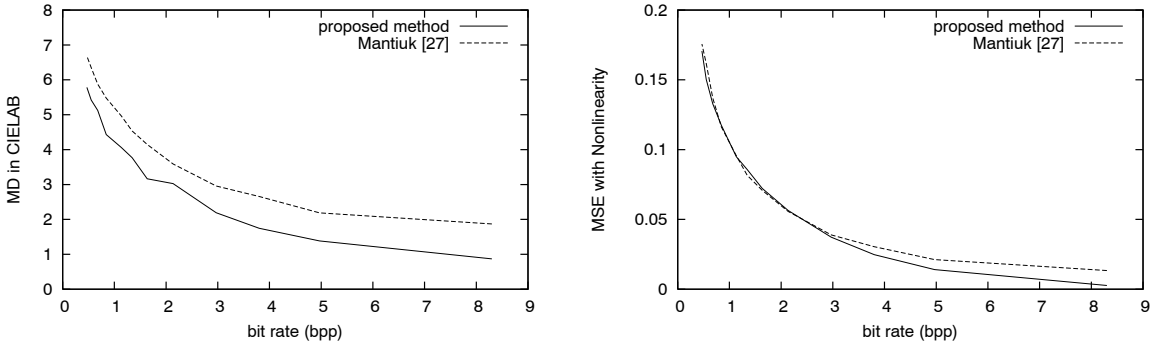


Fig. 7. Comparison with Mantiuk et al.'s method (27)

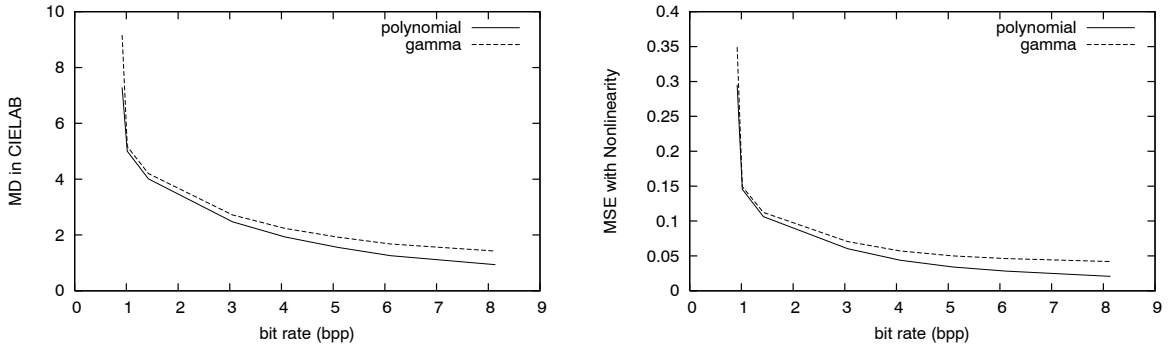


Fig. 8. Effect of Color Compensation ('polynomial': proposed color compensation, 'gamma': modfied Ward et al's compensation)

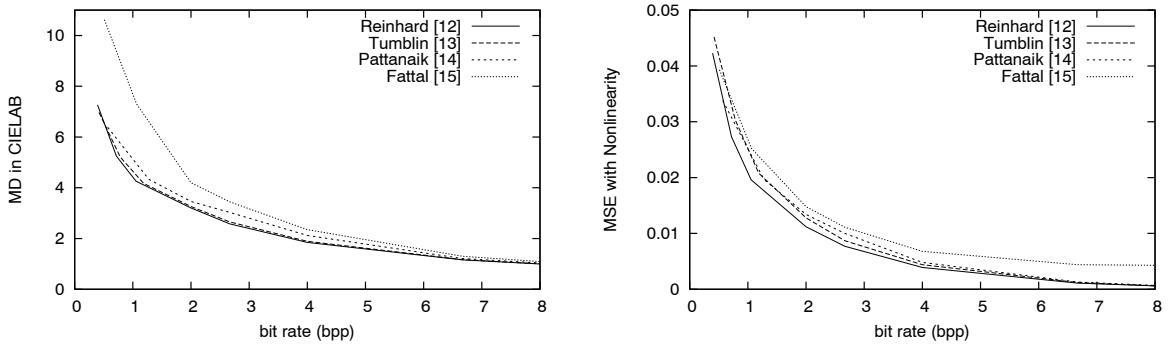


Fig. 9. Compression Performance with some TMO