# A Companding Based Two-Layer codec for HDR images

Asif Siddiq
Faculty of Engineering and Applied Sciences Riphah International University Islamabad, Pakistan
asif.siddik@gmail.com

Jameel Ahmed
Faculty of Engineering and Applied Sciences Riphah International University Islamabad, Pakistan
jameel.ahmed@riphah.edu.pk

Ishtiaq Rasool Khan
College of Computer Science and Engineering, University of Jeddah, Jeddah, Saudi Arabia
irkhan@uj.edu.sa

**ACM Reference Format:**
Asif Siddiq, Jameel Ahmed, and Ishtiaq Rasool Khan. 2020. A Companding Based Two-Layer codec for HDR images. In *Special Interest Group on Computer Graphics and Interactive Techniques Conference Posters (SIGGRAPH '20 Posters), August 17, 2020, Virtual Event, USA.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/3388770.3407438

## 1 INTRODUCTION

TIFF and OpenEXR formats store HDR images as floating-point data. Therefore, these raw data HDR image formats occupy large storage space and require very high bandwidth for transmission. RGBE format transforms three floating-point HDR channels to four 8-bit channels by extracting a common exponent from RGB values. However, the algorithm requires lossless encoding, and therefore reduction in file size is not significant. RGBE also encodes raw data and is not compatible with existing codecs and displays.

To resolve the problem of backward compatibility, two-layer solutions were proposed. In these formats, the first layer is a tone-mapped version of the HDR image suitable for display and can be encoded with the existing codecs. The second layer, known as residue layer or enhancement layer, stores the information lost in the process. JPEG released an extended JPEG-XT standard, which implements some of these algorithms as various profiles. These two-layer structures are generally called backward compatible formats, as they allow HDR contents to be visualized on traditional LDR displays and encoded with existing codecs.

As the HDR contents became more prevalent and market share of HDR displays increased, coding inefficiency of these formats due to requirement of backward compatibility started being a bottleneck. Therefore, some two-layer structures dropped the condition of compatibility with LDR displays. ([MAI et al., 2010]) optimized the tone-mapping curve to maximize the information contained in the first layer, thus not giving consideration only to the visual quality of the first layer on the LDR displays. More recently, new standards like HDR10 and Dolby Vision altogether dropped the idea of two layers and focus on other aspects of encoding and display. However, the idea of two-layer codecs is still being pursued actively. Researchers from Dolby [6] filed a patent application for their two-layer structure, in which first layer is constructed using a linear or non-linear mapping function whereas the second layer contains the non-linearly quantized residual data. The focus is on maximizing the encoded information through quantization. The authors refer to their structure as non-backward-compatible, however it is still compatible with the existing codecs. Other works have used smooth mapping functions to improve the compression of the second layer ([KHAN, 2015]) and non-linear quantization to enhance the information content encoded in the second layer ([KHAN, 2016]). Recently, ([KOBAYASHI et al., 2019]) presented a near lossless two-layer encoding.

In this work, we propose a two-layer structure comprising three channels of the first layer and one channel of the enhancement layer. Like ([SU et al., 2016]), the first layer of our design is not meant to be viewed on LDR displays but can be encoded using existing lossy or lossless codecs. All the existing two-layer structures generate some metadata for the decoder to reconstruct the HDR image by merging the two layers. The metadata can be embedded in the first layer or encoded separately. In either case, it adds to the design complexity. The proposed format does not require metadata for prediction of HDR from LDR contents. Rather it uses a continuous mapping functions in closed form for reshaping the contents while encoding and decoding, which distinguishes it from other two-layer structures. Moreover, it encodes the data with more precision than other designs, as demonstrated by the experimental results reported later in section 3.

## 2 OUR APPROACH

Companding refers to a technique of compressing and then expanding a signal. It has been standardized as ITU-T G.711 and is widely used for audio signals. The compression and expansion are carried out with non-linear log functions as given below by equations (1) and (2) respectively:

$$y = F(x) = \operatorname{sgn}(x)\,\frac{ln\,(1+\mu\,|x|)}{ln\,(1+\mu)},\ -1 \leq x \leq 1 \tag{1}$$

$$x = F^{-1}(y) = \operatorname{sgn}(y)\,\frac{(1+\mu)^{|y|}-1}{\mu},\ -1 \leq y \leq 1 \tag{2}$$

where $\mu$ is a constant that determines the slope of transformation curves. In a typical streaming scenario, compression function raises the amplitude of weak signals at transmitter and expansion function at receiver converts them back to the original values. A non-linear quantization using companding can reduce overall quantization error by better preserving low amplitude signals which suffer higher percentage loss under linear quantization. We leverage this property of companding function to design the quantizer for our two-layer HDR encoding structure.

In the proposed approach, HDR image is converted into YCbCr color space. The Y channel is compressed using (1) and quantized to 8 bits, while chroma channels are linearly quantized. All channels are encoded using an existing lossy or lossless format such as JPEG


*SIGGRAPH '20 Posters, August 17, 2020, Virtual Event, USA*

ACM ISBN 978-1-4503-7973-1/20/08.
https://doi.org/10.1145/3388770.3407438

**Table 1: Comparison of the proposed coding scheme with the existing methods for the test image Memorial Church.**

| | Ward | Mantiuk | Khan | Proposed |
|---|---|---|---|---|
| MSE | 5.4E-4 | 6.9E-4 | 1.1E-4 | 8.1E-5 |
| PSNR | 32.69 | 31.614 | 39.604 | 40.915 |
| SNR | 42.86 | 41.78 | 49.772 | 51.083 |
| VDP | 84.76 | 84.172 | 98.894 | 81.341 |
| File size | 801.57 | 599.77 | 714.31 | 284.52 |
| Reduction in size | 38.95% | 54.32% | 45.6% | 78.33% |

or PNG. Residue is calculated only for the Y channel. For this, the quantized 8-bit Y channel is expanded using (2) and subtracted from the original Y channel. The residue channel $R$ is quantized to 8 bits using linear quantization. Mathematically, the process of encoding Y channel can be described as:

$$Y_8 = Q(F(Y)), \tag{3}$$

$$R = Y - F^{-1}(Y_8), \tag{4}$$

$$R_8 = Q(R) \tag{5}$$

where $F$ and $F^{-1}$ have been defined in (1)-(2), $Q$ refers to linear quantization of data into 256 values between 0 and 255, and the subscript 8 indicates that the values are 8-bit integers. The reverse procedure to decode the $Y$ channel simply expands 8-bit $Y_8$ and $R_8$ channels and adds them as:

$$Y_d = F^{-1}(Y_8) + Q^{-1}(R_8). \tag{6}$$

It can be noted from above that no metadata is required for reconstruction of HDR image at the decoder. The only information required is the minimum and maximum values of the original HDR signal. Encoding and decoding are done using closed functions which can be implemented in parallel, and decoder does not require interpolation as is the case in most of the existing two-layer formats. These are the added advantage of the proposed structure besides high accuracy as shown in the next section.

## 3 RESULTS

To evaluate performance of the proposed codec, we carried out some comparative studies. A large number of OpenEXR and RGBE images, used as the ground truth, is coded using two-layer methods based on ([WARD and SIMMONS, 2006]), ([MANTIUK et al., 2006]), ([KHAN, 2016]), and the proposed one. The error in the decoded HDR image compared to the ground truth is measured using the objective metrics of mean square error (MSE), peak signal to noise ratio (PSNR), signal to noise ratio (SNR) and the quality score in HDR-VDP2 metric (VDP). For a fair comparison, we encoded HDR images in YCbCr 4:4:4 format for all the methods.

The study was carried out on a dataset of 135 images collected from publicly available resources. The proposed method outperformed the existing codecs in accuracy and achieved higher compression. We show the results for a typical test image, Memorial Church, in Table 1. The metrics' values shown in the table are calculated using all three channels of the reconstructed HDR images in RGB space.

In implementation of our method, we used some best practices on chroma channels to reduce the overall error. We applied the same to the codes of other methods as well to make a fair comparison and to observe the effect of the proposed companding based quantization alone. For the same reason, both layers are encoded losslessly using PNG codec, to exclude the impact of coding loss. Even the layers of ([WARD and SIMMONS, 2006]) are encoded as PNG, although the original method (referred as JPEG-HDR) and its implementation by JPEG-XT profile A use JPEG. Again, the reason is to observe the impact of the algorithms alone excluding the coding loss. The file sizes in Table 1 are for the entire coded data including both layers and the metadata. The reductions shown in file sizes in the last row are relative to the original raw format, RGBE in the case in Table 1

In the results shown in Table 1, the proposed method beats the rest in all metrics except VDP. However, average VDP score over the entire dataset of 135 images remained higher for the proposed method. File size of the proposed method is just 35.5%, 47.4% and 39.8% of the file sizes of ([WARD and SIMMONS, 2006]), ([MANTIUK et al., 2006]) and ([KHAN, 2016]).

## 4 CONCLUSIONS

A two-layer encoding structure for HDR images was presented. The first layer is not optimized for viewing, however data is encoded with much high accuracy and smaller file size than some state of the art existing methods.

## REFERENCES

KHAN, I.R., 2015. Effect of smooth inverse tone-mapping functions on performance of two-layer high dynamic range encoding schemes. *Journal of Electronic Imaging 24* 1(5 Feburary). DOI= http://dx.doi.org/10.1117/1.JEI.24.1.013024.

KHAN, I.R., 2016. A nonlinear quantization scheme for two-layer hdr image encoding. *Signal, Image Video Processing 10*, 5, 921-926. DOI= http://dx.doi.org/10.1007/s11760-015-0841-x.

KOBAYASHI, H., WATANABE, O., and HITOSHI, K., 2019. Two-Layer Near-Lossless HDR Coding with Backward Compatibility to JPEG. In *2019 IEEE International Conference on Image Processing (ICIP)* IEEE, 3547-3551.

MAI, Z., MANSOUR, H., MANTIUK, R., NASIOPOULOS, P., WARD, R., and HEIDRICH, W., 2010. Optimizing a tone curve for backward-compatible high dynamic range image and video compression. *IEEE transactions on image processing 20*, 6, 1558-1571. DOI= http://dx.doi.org/10.1109/TIP.2010.2095866.

MANTIUK, R., EFREMOV, A., MYSZKOWSKI, K., and SEIDEL, H.-P., 2006. Backward compatible high dynamic range MPEG video compression. In *Proceedings of the ACM SIGGRAPH* (July 2006), ACM SIGGRAPH 2006 Papers, 713-723. DOI= http://dx.doi.org/10.1145/1179352.1141946.

SU, G.-M., QU, S., HULYALKAR, S.N., CHEN, T., GISH, W.C., and KOEPFER, H., 2016. Layer decomposition in hierarchical VDR coding Google Patents.

WARD, G. and SIMMONS, M., 2006. JPEG-HDR: A backwards-compatible, high dynamic range extension to JPEG. In *ACM SIGGRAPH 2006 Courses*, 3-11. DOI= http://dx.doi.org/10.1145/1185657.1185685.