

ハイダイナミックレンジ画像のための二階層ロスレス符号化

岩橋 政宏[†] 貴家 仁志[‡][†]長岡技術科学大学 工学部 〒940-2188 新潟県 長岡市 上富岡町 1769-1[‡]首都大学東京 システムデザイン学部 〒191-0065 東京都 日野市 旭が丘 6-6E-mail: [†] iwahashi@nagaokaut.ac.jp

あらまし 本報告では、浮動小数点形式で表現された高ダイナミックレンジ (HDR) 画像に対する、二階層のロスレス符号化を提案する。基本階層からは、標準的なビット深度を持つ、固定小数点形式で表現された低ダイナミックレンジ (LDR) 画像を復号できる。更に、拡張階層と併せることで、元の HDR 画像をロスレスに復号できる。提案法では、ロスレスな階調写像が活用されているため、拡張階層のビット深度を大幅に低減できる。更に、LDR 画像については、HDR 画像が対数変換されており、視覚的にも相応しい画像となっている。

キーワード 画像, 符号化, 圧縮, 可逆

Two Layered Lossless Coding of High Dynamic Range Images

Masahiro IWAHASHI[†] and Hitoshi KIYA[‡][†] Nagaoka University of Technology Niigata, 980-2188, Japan[‡] Tokyo Metropolitan University Tokyo, 191-0065, Japan

Abstract A lossless layered coding for high dynamic range (HDR) images expressed in floating point format is proposed. It outputs two kinds of bit streams - a basic layer and an enhance layer. A low dynamic range (LDR) image represented in fixed point format with a standard bit depth is decoded from the basic layer bit stream. The original HDR image is decoded without any loss from both of the layers. Utilizing the lossless logarithmic tone mapping, the proposed method can reduce the bit depth of signal values in the enhance layer.

Keyword Image, Coding, Compression, Lossless

1. まえがき

画像や映像のデジタル・データ圧縮において、圧縮結果が複数の階層に分離される符号化方法は、スケラビリティ機能と呼ばれ[1]、これまでに多くの研究がなされてきた。ITU-T や ISO による国際標準では、「画素数」や「画質」に関する階層化が古くから実施されているが、近年では「ビット深度」の階層化が盛んに研究されている[2-8]。

ビット深度とは、画素値を表す階調数を意味し、汎用的な 8 ビット深度の装置では、 $2^8=256$ 種類の階調を表現できる。近年では、これまでの固定小数点形式に加え、浮動小数点形式を用いることで、より広い値域 (ハイダイナミックレンジ) を表現できる装置が開発されている[9-10]。

本報告では、浮動小数点形式で表現された高ダイナミックレンジ (HDR) 画像に対する、二階層のロスレス符号化を提案する。とくに、基本階層からは、標準的なビット深度を持つ、固定小数点形式で表現された低ダイナミックレンジ (LDR) 画像が復号され、また、拡張階層と併せることで、元の HDR 画像がロスレス

に復号される方法について検討する。

ビット深度の階層化については、これまでに多くの報告がある。文献[2]では、階調写像により HDR 画像から LDR 画像を作成し、JPEG 方式で圧縮して基本階層としている。また、LDR 画像を逆階調写像し、HDR 原画像からの差分をとり拡張階層としている。文献[3-5]では、MPEG や H.264 方式を用いて動画像に応用されている。文献[6,7]では、離散ウェーブレット変換 (DWT) に基づき、3 次元医用画像の圧縮に適用された。しかし、いずれの方法においても、拡張階層は残差であるため画素間相関が弱く、高い圧縮率は実現できない問題がある[8]。

これに対し本報告では、ロスレスな階調写像を活用することで、拡張階層のビット深度を大幅に低減できる方法を提案し、その効果を実験的に示す。但し、入力信号としては、RGBE 形式や OpenEXR 形式のような[9,10]、浮動小数点表現された HDR 画像を仮定する。更に、基本階層から復号される LDR 画像が、HDR 画像が、文献[12]の対数関数により階調写像されたものと、ほぼ同一となることを示す。

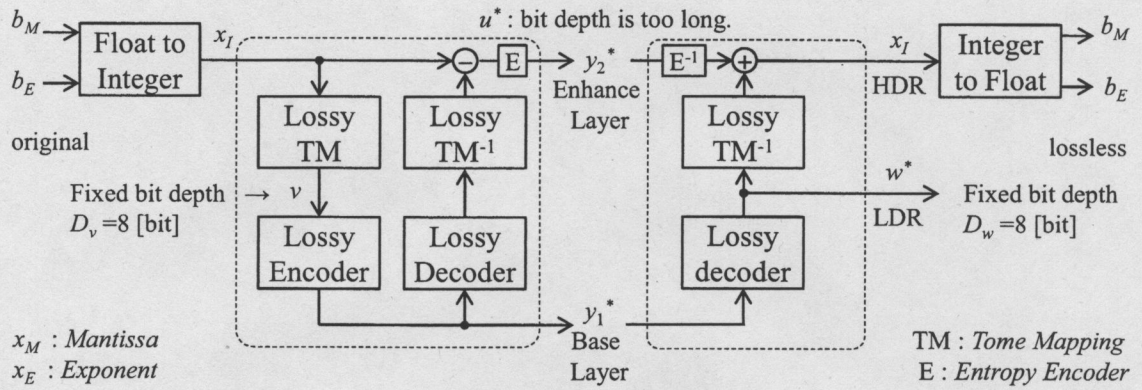


図1 ロッシー・エンコーダに基づく方法

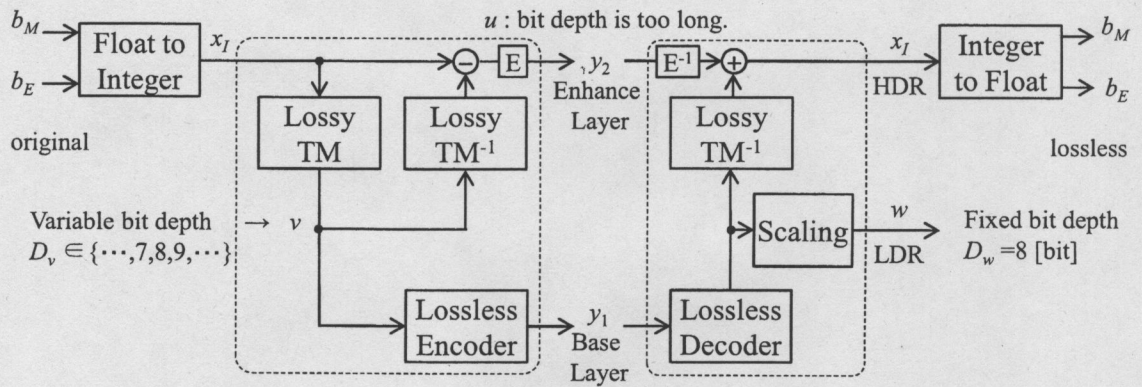


図2 ロスレス・エンコーダに基づく方法 (本報告における従来法)

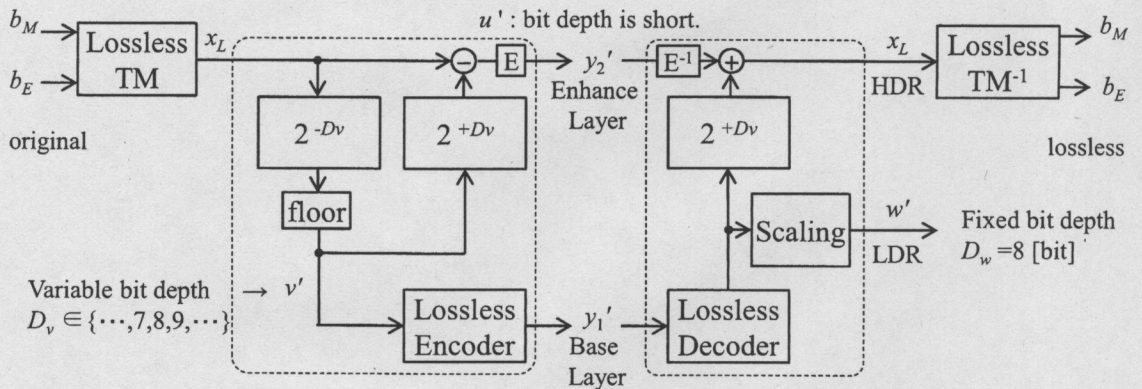


図3 ロスレス・エンコーダに基づく方法 (本報告における提案法)

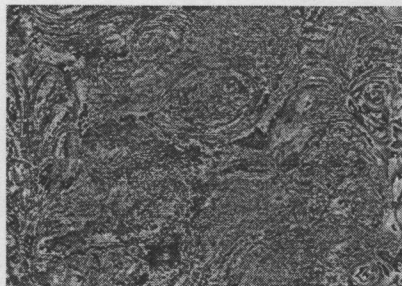


図4 原画像の仮数部 \$b_M\$



図5 原画像の指数部 \$b_E\$



図6 LDR 復号画像 \$w'\$

2. 浮動小数点表現された HDR 画像

本稿では、入力画像のデータ形式として、浮動小数点表現を前提としている。これは、正負符号部 b_S 、仮数部 b_M 、指数部 b_E から構成され、

$$b_S \in [0, 2^{D_S}), \quad b_M \in [0, 2^{D_M}), \quad b_E \in [0, 2^{D_E}) \quad (1)$$

であり、それぞれのビット深度が D_S, D_M, D_E [bit] の整数である。例えば、OpenEXR 形式では R,G,B の各色が $[D_S \ D_M \ D_E]=[1 \ 10 \ 5]$ なるビット深度であり全体では 48 [bpp] となる。また、RGBE 形式では符号部はなく、仮数部は Red, Green, Blue のそれぞれが $D_M=8$ であるが、指数部は R,G,B で共通に使われ $D_E=8$ であり、全体では 32 [bpp] となる。

図 1~3 に示す符号化方法は全て、このように浮動小数点表現された値を入力とする。例えば、OpenEXR 形式の場合、 $0 < b_E < 31$ に対して画素値は、

$$x = x_M \cdot 2^{x_E} \quad (2)$$

$$\begin{cases} x_M = (-1)^{b_S} \cdot (1 + b_M \cdot 2^{-D_M}) \\ x_E = b_E - 15 \end{cases}$$

と表され実数となる。ここで、 x_M は仮数部、 x_E は指数部である。図 4,5 にはこれらの信号 (Red のみ) を例示する。但し、最小を 0、最大を 255 に正規化してある。以下、紙面の都合上、説明を $b_S=0$ の場合に限定する。

3. HDR 画像の二階層符号化

図 1 は、文献[2]の従来法に、浮動小数点表現から整数への一対一な写像 (Float to Integer)

$$x_I = (2^{D_M} + b_M) \cdot 2^{b_E - \text{Min}(b_E)} \quad (3)$$

と、その逆変換が追加されている。図中のロッシェン・エンコーダには、例えば、JPEG、MPEG、H.264 などが利用される。これに対する入力は、HDR 画像が階調写像 (TM) され、ビット深度が D_v に縮小されている。本稿では、階調写像として文献[12]の、

$$v^* = \log_e x_I \quad (4)$$

$$v = \left\lfloor (2^{D_v} - 1) \cdot \frac{v^* - \text{Min}(v^*)}{\text{Max}(v^*) - \text{Min}(v^*)} \right\rfloor \quad (5)$$

を用いている。また、ビット深度は $D_v=8$ とした。

本稿では、HDR 画像に対するロスレス符号化を考えているため、図 1 におけるロッシェン・エンコーダは、図 2 のように、ロスレス・エンコーダに置き換えられる。更に、LDR 画像 v のビット深度 D_v を適宜変えることで、基本階層のビットレートを制御できる。この方法を、本稿における従来法とする。

従来法の問題は、拡張階層の圧縮率が極端に低いことである[8]。第 1 の理由は、 x_I のビット深度が、

$$D_I = D_M + 2^{D_E} \quad (6)$$

程度と非常に長くなってしまふ。第 2 の理由は、信号 u が残差 (乃至は剰余) であるため、画素間の相関が弱くなる。このため、DCT や DWT のような変換処理では、高い圧縮効果が望めない。

4. 提案する二階層符号化

図 3 は、本稿における提案法を示す。式(4)に代わり、ロスレスな対数写像、

$$x_L = (b_E - \text{Min}(b_E)) \cdot 2^{D_M} + b_M \quad (7)$$

を活用している[16]。これによりビット深度は、

$$D_L = D_M + D_E \quad (8)$$

となり大幅に低減される。更に、従来法における x_I はヒストグラムの疎性が高いのに対し、提案法における x_L は低いため、圧縮効率の向上に寄与する特長がある[17]。式(7)からは、式(2)を用いることで、

$$x_L = (\log_2 x - \delta - C) \cdot 2^{D_M} \quad (9)$$

$$\delta = (\log_2 x_M + 1) - x_M, \quad C = \text{Min}(b_E) - 15$$

が得られる。式(4)と比べると、対数の底と δ の違いがあるが、両者とも対数関数であるため、図 6 に例示するように、従来法とほぼ同じ LDR 画像を復号できる。これは、今後、文献[13-15]等の階調写像を用いた場合に、若干の補正を施すことで、より視覚的に相応しい LDR 画像を表示できることを示唆している。

5. 実験結果

図 7 には、復号された LDR 画像 w の誤差 ($w-v$ の標準偏差) と基本階層 y_1 の平均符号長 (1 次エントロピ) との関係を示す。提案法では 4 [bpp] 以上であれば、誤差は ± 1 程度と十分小さくなっている。

図 8 には、提案法における x_L と従来法における x_I との関係を、 $D_M=2$ を例に示す。両者には、

$$x_L = (\log_2 x_I - \delta' - C') \cdot 2^{D_M} \quad (10)$$

$$\delta' = (\log_2 x_M + 1) - x_M, \quad C' = D_M$$

なる関係がある。入力に依存する項 δ' を 0 とした場合を点線で示す。この項が僅かであることがわかる。

図 9 には、 D_v を $\{1, 2, \dots, 11\}$ の範囲で変化させたときの、基本階層 v と拡張階層 u それぞれのビット深度 (= $\log_2 [\text{Max} - \text{Min} + 1]$) を示す。提案法により拡張階層のビット深度が 6 [bit] 程度低減されることがわかる。

図 10 には、基本階層 y_1 と拡張階層 y_2 それぞれの平均符号長を示す (今回はロスレス・エンコーダとして DPCM を用いた)。基本階層のビットレートが 3 [bpp] 以上の場合、拡張階層のビットレートが 4 [bpp] 程度低減されることがわかる。

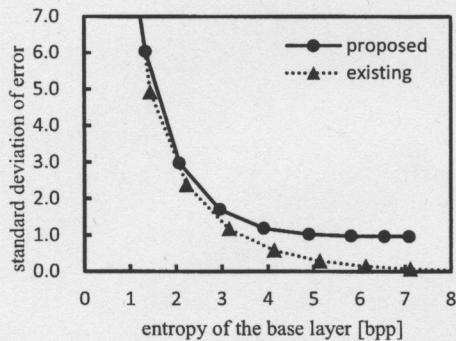


図 7 復号画像 w の画質と基本階層 y_1 の平均符号長

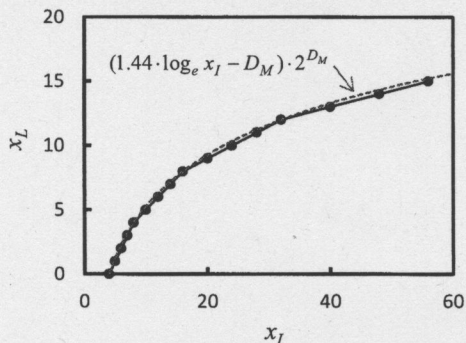


図 8 可逆階調画像の関数形 ($D_M=2$ の例)

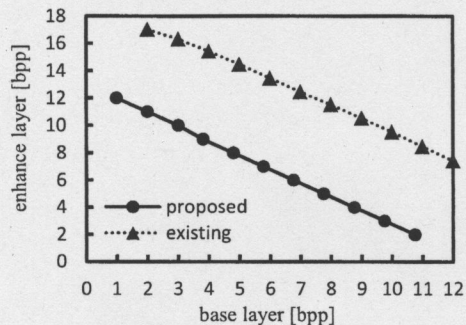


図 9 基本階層 v と拡張階層 u それぞれのビット深度

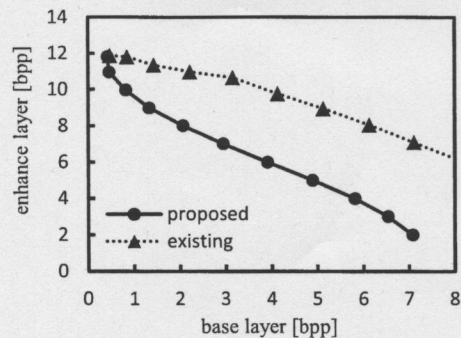


図 10 基本階層 y_1 と拡張階層 y_2 それぞれの平均符号長

6. まとめ

浮動小数点表現された HDR 画像に対し、ロスレス階調画像を活用することで、階層符号化における拡張階層のビット深度を大幅に低減できた。この際、基本階層から復号される LDR 画像を、従来法と同等の画質に維持できた。今後は文献[13-15]を組み込む等、階調画像の視覚特性を向上させる必要がある。

文 献

- [1] H. Schwarz, D. Marpe, T. Wiegand, "Overview of the Scalable H.264 / MPEG4 - AVC Extension," IEEE Proc. ICIP, pp.161-164, Oct. 2006.
- [2] G. Ward, and M. Simmons, "JPEG - HDR: A Backwards - Compatible, High Dynamic Range Extension to JPEG," Proc. 13th Color Imaging Conference, Nov. 2005.
- [3] R. Mantiuk, A. Efremov, K. Myszkowski, H. P. Seidel, "Backward Compatible High Dynamic Range MPEG Video Compression," SIGGRAPH 2006.
- [4] M. Winken, D. Marpe, H. Schwarz and T. Wiegand, "Bit Depth Scalable Video Coding," IEEE Proc. ICIP, vol.1, no.1, pp.5-8, Oct.2007.
- [5] S. Park, K. R. Rao, "Bit Depth Scalable Video Coding based on H.264 / AVC," IEICE Trans. Fund., vol.E91-1A, no.6, pp.1541-1544, June 2008.
- [6] S. Yea, W. A. Pearlman, "A Wavelet - Based Two - Stage Near - Lossless Coder," IEEE Trans. Image Proc., vol. 15, Issue 11, pp.3488-3500, Nov. 2006.
- [7] G. Carvajal, B. Penna, E. Magli, "Unified Lossy and Near-Lossless Hyperspectral Image Compression Based on JPEG 2000," IEEE Geoscience and Remote Sensing, vol.5, Issue 4, pp.593-597, Oct. 2008.
- [8] H. Kikuchi, W. Otake, M. Iwahashi, "Bit Rate Reduction of Enhancement Layer in Bit - Depth Scalable Coding," IEEE ISPACS, pp.264-267, 2009.
- [9] G. Ward, "Real Pixels, Graphics Gems II," edited by James Arvo, Academic Press, 1992.
- [10] E. Reinhard, G. Ward, S. Pattanaik, P. Debevec, W. Heidrich and K. Myszkowski, "High Dynamic Range Imaging - Acquisition, Display and Image based Lighting," Morgan Kaufmann Publisher, 2010.
- [11] R. Mantiuk, G. Krawczyk, K. Myszkowski, and H. P. Seidel "Perception-Motivated High-Dynamic-Range Video Encoding," ACM Trans. Graphics, vol. 23, no. 3, pp.733-741, 2004.
- [12] R. Xu, S. N. Pattanaik and C. E. Hughes, "High Dynamic-Range Still Image Encoding in JPEG2000," IEEE Computer Graphics and Applications, vol.25, no.6, pp.57-64, 2005.
- [13] E. Reinhard, M. Stark, P. Shirley, and J. Ferwerda, "Photographic Tone Reproduction for Digital Images," ACM Tran. Graphics, Vol.21, No.3, pp. 267-276, 2002.
- [14] M. Okuda and N. Adami, "Two Layer Coding Algorithm for High Dynamic Range Images based on Luminance Compensation," Journal Visual Communication and Image Representation, vol.18, Issue 5, pp.377-386, Oct. 2007.
- [15] I. R. Khan, "Two Layer Scheme for Encoding of High Dynamic Range Images, IEEE ICASSP, pp.1169-1172, 2008.
- [16] ISO/IEC, 29199-2, "Information technology - JPEG XR, image coding system - Part 2: Image coding specification," 2010.
- [17] M. Iwahashi, H. Kobayashi, H. Kiya, "Lossy Compression of Sparse Histogram Image", IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), March 2012.