IMAGE DE-QUANTIZATION VIA SPATIALLY VARYING SPARSITY PRIOR

Pengfei Wan, Oscar C. Au, Ketan Tang, Yuanfang Guo

The Hong Kong University of Science and Technology
Clear Water Bay, Kowloon, Hong Kong
Email: {leoman, eeau, tkt, eeandyguo}@ust.hk

ABSTRACT

We address the problem of image de-quantization, which is also known as bit-depth expansion if the reconstructed 2D signal is re-quantized into higher bit-precision. In this paper, a novel image de-quantization method based on convex optimization theory is proposed, which exploits the spatially varying characteristics of image surface. We test our method on image bit-depth expansion problems, and the experimental results show that proposed method can achieve superior PSNR and SSIM performance.

Index Terms— De-quantization, Image bit-depth expansion, L_1-L_2 optimization

1. INTRODUCTION

De-quantization, as its name shows, is the inverse process of quantization. Image de-quantization aims to recover the original 2D analog signal from an digital image. The most important application of image de-quantization is bit-depth expansion, in which the reconstructed 2D signal is re-quantized into an image with higher bit-depth. Bit-depth is the number of bits to represent a quantized pixel value. Nowadays most images are stored in 8-bit precision.

In fact, image bit-depth expansion from low bit-depth (LBD) to high bit-depth (HBD) is often desirable. For example, incompatibility occurs when 6-bit DVD videos are displayed on 8-bit LCD displays [1], or when 8-bit images are displayed on high dynamic range (HDR) displays. In the forthcoming video coding standard HEVC, an internal bit-depth increase (IDBI) scheme is incorporated to improve the precision of data processing [2].

To reduce the bit-depth, least significant bits (LSBs) truncation is sufficient, however bit-depth expansion is ill-posed. In this paper we study the spatially varying characteristics of image intensity surface and formulate the bit-depth expansion problem into a clean L_1-L_2 optimization [3] problem.

The paper is organized as follows. In Section 2 we briefly review the existing bit-depth expansion methods. Section 3 elaborates the proposed method and In Section 4 we provide numerical results to show the effectiveness of our method. Section 5 summarizes our work.

This work has been supported in part by the Research Grants Council (RGC) of the Hong Kong Special Administrative Region, China.

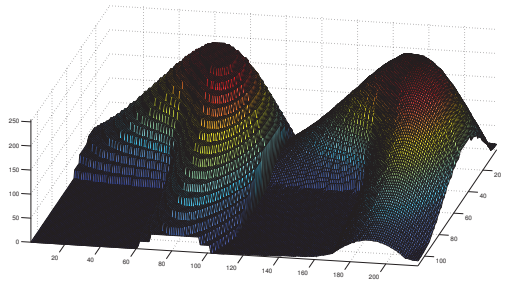


Fig. 1. Image bit-depth expansion from 4-bit to 8-bit for MATLAB® membrane. Image surface on the left is the zero-padded output, right is the proposed output.

2. PREVIOUS WORK

Plenty of work on bit-depth expansion exist in literature [4–6]. To begin with, let $\mathbf{I_h^{gt}}$ be the ground-truth HBD image with bit-depth q and $\mathbf{I_l}$ be the input LBD image with bit depth p ($p < q$). $\hat{\mathbf{I}}_\mathbf{h}$ denotes the output HBD image.

The most straightforward method for bit-depth expansion is zero padding, which has a very simple form:

$$\hat{\mathbf{I}}_\mathbf{h}^{\mathbf{zp}} = 2^{q-p}\mathbf{I_l} = \mathbf{I_l} \ll (q-p) \tag{1}$$

Because of uniform rescaling of image intensities, severe contouring artifacts exist in $\hat{\mathbf{I}}_\mathbf{h}^{\mathbf{zp}}$, see Fig. 1 and Fig. 2.

In [6], Daly et al. proposed the predictive de-contouring method as follows:

$$\hat{\mathbf{I}}_\mathbf{h}^{\mathbf{dl}} = \hat{\mathbf{I}}_\mathbf{h}^{\mathbf{zp}} + \tilde{\mathbf{I}}_\mathbf{h}^{\mathbf{zp}} - (\tilde{\mathbf{I}}_\mathbf{h}^{\mathbf{zp}} \gg d) \ll d \tag{2}$$

Where $d = q - p$, $\tilde{\mathbf{I}}_\mathbf{h}^{\mathbf{zp}} = \hat{\mathbf{I}}_\mathbf{h}^{\mathbf{zp}} * \mathbf{F}$, and $\mathbf{F}$ is a low-pass filter with fixed spatial size. So Daly's method has intrinsic problems in robustness and adaptiveness.

The flooding-based method is first proposed by Cheng et.al [4], which can be summarized as:

$$\hat{\mathbf{I}}_\mathbf{h}^{\mathbf{ch}} = \hat{\mathbf{I}}_\mathbf{h}^{\mathbf{zp}} + \lfloor (2^{q-p}-1)\mathbf{SR} \rfloor \tag{3}$$

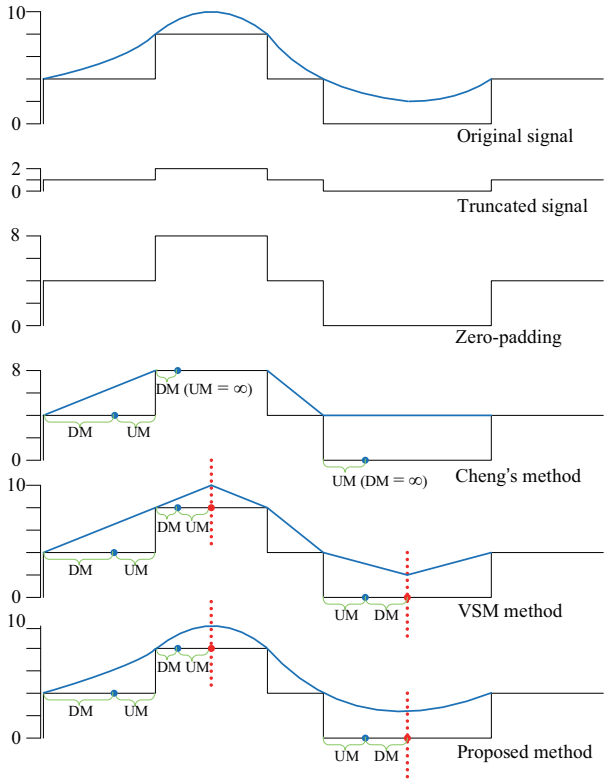


Fig. 2. 1D illustration of selected bit-depth expansion methods. To reconstruct the original 4-bit signal from its truncated 2-bit version, zero padding introduces severe false contours; Cheng's method [4] can avoid contouring but the information in LMM regions is totally lost. Proposed method solves the three major problems in the VMS method [5], see Section 2.

$$\mathbf{SR} = \frac{\mathbf{DM}}{\mathbf{DM} + \mathbf{UM}} \in [0, 1] \tag{4}$$

Where k is pixel index, $\mathbf{SR}$ is step ratio map. $\mathbf{DM/UM}$ is respectively called the downward/upward distance map, where each entry records the distance from pixel k to its nearest contour pixel with smaller/larger intensity, see Fig. 2.

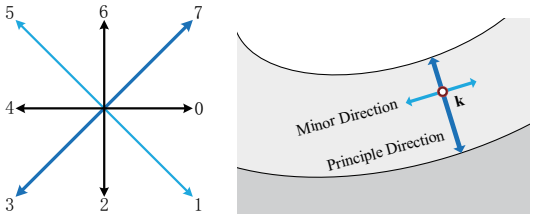
The principle direction of pixel k is then defined as the direction along which the length of line segment across pixel k while connecting k's upward and downward contour pixels is minimal. The minor direction of pixel k is defined as the direction that is orthogonal to its principle direction, see Fig. 3. Note that the concepts of principle and minor directions will be frequently used in this paper.

Although Cheng's method can produce smooth HBD images, it fails to reconstruct the signal in local maxima/minima (LMM) regions, see Fig. 2. Our recent work solves this problem by the virtual skeleton marking (VSM) algorithm [5]. However it still suffers from the nature of linear interpolation along principle directions, which has three major problems:

1. Linear interpolation assumes the second order derivative of image intensity along the principle directions is zero, which fails in very bright and dark regions.

2. First order derivative of image intensity along the minor direction is not specified, which may lead to output HBD image with poor 2D smoothness.

3. Piecewise linear interpolation has discontinuities at the turning points, causing unnatural ringing perceptions.



(a) 8-neighbor model (b) Principle and minor directions

Fig. 3. Neighborhood model and an illustration of the principle direction (dark blue) and minor direction (light blue).

3. PROPOSED METHOD

3.1. Problem Formulation

The proposed method is designed to solve the above problems. We first apply neighborhood flooding [4] on the input image to get the principle and minor directions, then the de-quantized image is obtained by minimizing a regularized convex cost function:

$$\hat{\mathbf{I}}_\mathbf{h} = \arg\min_{\mathbf{I_h}} \ \mathcal{T}_{data} + \lambda_p \mathcal{T}_{pho} + \lambda_s \mathcal{T}_{smo} \tag{5}$$

Where $\mathcal{T}_{data,pho,smo}$ is respectively called data, photometry and smoothness term. $\lambda_{s,p}$ are regularization parameters.

3.1.1. *Data Term*

Data term shows the primary anticipation for pixel intensities of the de-quantized image:

$$\mathcal{T}_{data} = \frac{1}{2}\|\mathbf{I_h} - \mathbf{I}^\star_\mathbf{h}\|^2_F \tag{6}$$

Where $\mathbf{I}^\star_\mathbf{h} = 2^{q-p}\mathbf{I_l} + (2^{q-p} - 1)\mathbf{SR}$, which is the output of linear interpolation using the VSM algorithm [5]. Frobenius norm is used to allow more flexibility for small deviations.

3.1.2. *Photometry Term*

Photometry term investigates the relation between pixel intensity and the second order derivative along the principle directions. By viewing de-quantized image surface as a smooth discrete 2D signal, we have the following propositions:

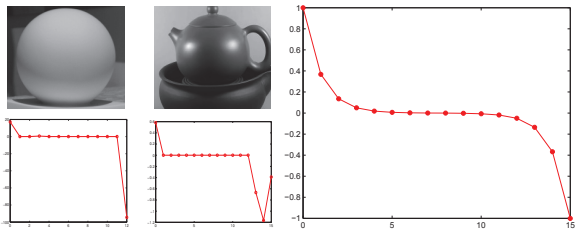


Fig. 4. Left figure shows the relation between image intensity (X axis) and the median value of second order derivative along the principle directions (Y axis) for sample images. Right is the proposed model (7) with $\lambda = 1$, $\delta = 15$.

1. In bright image regions, the second order derivative along the principle directions should be negative.
2. In dark image regions, the second order derivative along the principle directions should be positive.
3. In other regions, the second order derivative along the principle directions is approximately zero.

Inspired by above results, we propose a heuristic model f to approximate the relation between pixel intensity $\mathbf{I_l}$ and normalized second order derivative along the principle directions, see Fig. 4:

$$\mathbf{I_h}'' = f(\mathbf{I_l}) = \alpha\left(e^{-\lambda \mathbf{I_l}} - e^{\lambda(\mathbf{I_l}-\delta)}\right) \tag{7}$$

Where λ controls the decay, and $\delta = 2^p - 1$ is the largest pixel value of input image. Normalization parameter $\alpha = \frac{1}{1-e^{-\lambda\delta}}$ makes sure $f(0) = 1, f(\delta) = -1$.

Our Photometry term is then defined as:

$$\mathcal{T}_{pho} = \|\mathcal{D}_2 \otimes \mathbf{I_h} - f(\mathbf{I_l})\|_F^2 \tag{8}$$

Where operator $\mathcal{D}_2$ calculates the second order derivative of image intensity along the principle directions:

$$(\mathcal{D}_2 \otimes \mathbf{I_h})(k) = \mathbf{I_h}(\mathcal{N}_k^{up}) + \mathbf{I_h}(\mathcal{N}_k^{dn}) - 2\mathbf{I_h}(k) \tag{9}$$

$\mathcal{N}_k^{up}$ and $\mathcal{N}_k^{dn}$ are the pixel indices of immediate neighbors along k's principle direction. For the example in Fig. 3(a), they are neighbors indexed by 3 and 7.

Equivalently, photometry term encourages the sparsity of the second order derivative of image intensity along principle directions, while preserving the photometric characteristics of image intensity surface. Since the principle direction of each pixel k depends on image contents, $\mathcal{D}_2$ is spatially varying.

3.1.3. Smoothness Term

For smoothness of reconstructed 2D signal, we minimize the anisotropic total variation along the minor directions:

$$\mathcal{T}_{smo} = \sum_k |\mathbf{I_h}(k) - \mathbf{I_h}(\mathcal{N}_k^{lf})| + |\mathbf{I_h}(k) - \mathbf{I_h}(\mathcal{N}_k^{rt})| \tag{10}$$

Where $\mathcal{N}_k^{lf}$ and $\mathcal{N}_k^{rt}$ are the pixel indices of immediate neighbors along k's minor direction. For the example in Fig. 3(a), they are neighbors indexed by 1 and 5. Like the photometry term, calculations in (10) are also spatially varying.

Essentially, smoothness term encourages the sparsity of first order derivative along the minor directions. Note that L_1 norm is friendly to true edges in the image.

3.2. Solving the Optimization Problem

Let's rewrite the full objective function in (5) below:

$$\begin{aligned}&\frac{1}{2}\|\mathbf{I_h} - \mathbf{I_h^\star}\|_F^2 + \lambda_p\|\mathcal{D}_2 \otimes \mathbf{I_h} - f(\mathbf{I_l})\|_F^2 + \\ &\lambda_s \sum_k |\mathbf{I_h}(k) - \mathbf{I_h}(\mathcal{N}_k^{lf})| + |\mathbf{I_h}(k) - \mathbf{I_h}(\mathcal{N}_k^{rt})|\end{aligned} \tag{11}$$

By re-arranging the images $\mathbf{I_h}, \mathbf{I_h^\star}, \mathbf{I_l}$ into column vectors, we get a simpler linear transformation form:

$$\frac{1}{2}\|\mathbf{I_h} - \mathbf{I_h^\star}\|_2^2 + \lambda_p\|\mathbf{D_2 I_h} - f(\mathbf{I_l})\|_2^2 + \lambda_s\|\mathbf{D_1 I_h}\|_1 \tag{12}$$

Where $\mathbf{D_2}$ is a $N \times N$ matrix. N is the number of pixels in image. Each row of $\mathbf{D_2}$ corresponds to a pixel k, with the column entries corresponding to $k, \mathcal{N}_k^{up}, \mathcal{N}_k^{dn}$ being $-2, 1, 1$ respectively. $\mathbf{D_1}$ is a $2N \times N$ matrix. Similarly, its column entries corresponding to $k, \mathcal{N}_k^{lf}$ (or $\mathcal{N}_k^{rt}$) are $-1, 1$.

In practice, not all pixels have principle and minor directions by definition, see Section 2. So rows in $\mathbf{D_1}$ and $\mathbf{D_2}$ that correspond to pixels whose upward and downward contour pixels are not on their opposite sides are set to zeros. That means our photometry term and smoothness term only operate on pixels with principle and minor directions. Irregular textures and image noise will not affect the effectiveness of our method.

Note the first two terms in (12) is a Tikhonov regularization term, so we get the final formulation as follows:

$$\hat{\mathbf{I}}_\mathbf{h} = \arg\min_{\mathbf{I_h}} \frac{1}{2}\|\mathbf{A I_h} - \mathbf{b}\|_2^2 + \lambda_s\|\mathbf{D_1 I_h}\|_1 \tag{13}$$

Where

$$\mathbf{A} = \begin{pmatrix} \mathbf{I}(N,N) \\ \sqrt{2\lambda_p}\,\mathbf{D_2} \end{pmatrix}_{2N\times N}, \mathbf{b} = \begin{pmatrix} \mathbf{I_h^\star} \\ \sqrt{2\lambda_p}\, f(\mathbf{I_l}) \end{pmatrix}_{2N\times 1}$$

Note $\mathbf{I}(N,N)$ is the $N \times N$ identity matrix. The problem now becomes a L_1-L_2 optimization problem, which can be efficiently solved by iterative shrinkage or L_1-regularized iterative least square methods [3]. For the memory cost, although N can be very large for general images, our method can handle the high dimensionality of $\mathbf{A}, \mathbf{b}, \mathbf{D_1}$ by sparse matrix representation.

The output of (13) is in nature a de-quantized discrete 2D signal with continuous intensity values. In our experiments, it is re-quantized to a HBD image with higher bit-depth q.

Table 1. Performance Comparisons in PSNR (dB) and SSIM [7]

		Zero padding		Daly et al. [6]		Cheng et al. [4]		VSM [5]		Proposed	
Setup	Image	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
5-8 bit	bell	35.59	0.9405	38.76	0.9130	45.37	0.9935	52.12	0.9958	**52.85**	**0.9964**
	vally	35.76	0.9463	37.85	0.9037	45.61	0.9874	46.12	0.9897	**47.50**	**0.9941**
	sinc	35.94	0.9206	37.56	0.9082	41.16	0.9456	42.28	0.9711	**43.21**	**0.9770**
4-8 bit	bell	29.04	0.8584	30.99	0.7069	36.47	0.9888	48.24	0.9949	**48.89**	**0.9953**
	vally	29.36	0.8365	31.63	0.7425	38.56	0.9761	39.65	0.9848	**40.11**	**0.9872**
	sinc	29.89	0.8110	31.75	0.7568	33.04	0.8758	38.42	0.9612	**39.42**	**0.9689**
3-8 bit	bell	22.64	0.8045	24.88	0.6906	27.13	0.9735	41.14	0.9924	**41.40**	**0.9926**
	vally	23.11	0.6894	25.35	0.6046	29.80	0.9454	31.61	**0.9759**	**31.77**	0.9729
	sinc	24.04	0.5879	25.71	0.5725	26.75	0.7890	29.47	0.8582	**30.01**	**0.8718**

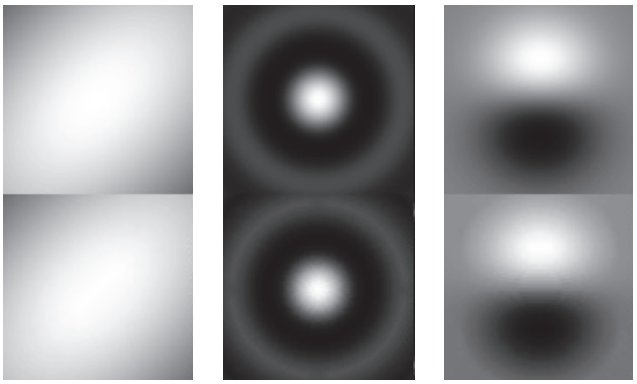


Fig. 5. Results in 4-bit to 8-bit expansion. The first row are the ground-truth 8-bit test images and the second row are the output of proposed method. From left to right, they are respectively bell, sinc and vally.

4. EXPERIMENTS

In this section, we show the effectiveness of proposed method, by comparing with a variety of bit-depth expansion methods [4–6]. Note the input LBD images are obtained by LSBs truncation of 8-bit test images (128×128).

Three sets of image bit-depth expansion experiments (5-8 bit, 4-8 bit and 3-8 bit) are conducted, using peak signal to noise ratio (PSNR) and structural similarity (SSIM) [7] as similarity metrics.

The numerical results show that proposed method is able to achieve best PSNR and SSIM performance in most cases. It is noteworthy that proposed method has very pleasing SSIM index levels. That's because in proposed method the discontinuities caused by piecewise linear interpolation are effectively removed, and more realistic intensity patterns are reconstructed in bright and dark image regions, see Fig. 5.

For detailed performance comparisons, please refer to Table. 1.

5. CONCLUSIONS

In this paper we address the image de-quantization especially bit-depth expansion problem and propose a clean and effective image de-quantization method using L_1-L_2 optimization. The effectiveness of proposed method is verified in our experiments. Future work could be more complicated modelling of image intensity surface and the study of proposed method in computer graphics applications.

6. REFERENCES

[1] R. Mantiuk, A. Efremov, K. Myszkowski, and H. Seidel, "Backward compatible high dynamic range mpeg video compression," in *ACM SIGGRAPH*, 2006, pp. 713–723.

[2] "Meeting report of the fourth meeting of the joint collaborative team on video coding (jct-vc), daegu, kr," 20-28 January, 2011.

[3] M. Zibulevsky and M. Elad, "L1-l2 optimization in signal and image processing," *Signal Processing Magazine, IEEE*, vol. 27, no. 3, pp. 76–88, 2010.

[4] C. Cheng, O. C. Au, C. Liu, and K. Yip, "Bit-depth expansion by contour region reconstruction," in *Proc. of IEEE Int. Sym. on Circuits and Systems*, 2009.

[5] P. Wan, O. C. Au, K. Tang, Y. Guo, and L. Fang, "From 2d extrapolation to 1d interpolation: Content adaptive image bit-depth expansion," in *Proc. IEEE Int. Conf. Multimedia and Expo*, 2012.

[6] S. Daly and X. Feng, "Decontouring: Prevention and removal of false contour artifacts," in *Proc. SPIE Human Vision and Electronic Imaging IX*, 2004, vol. 5292, pp. 130–149.

[7] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *Image Processing, IEEE Transactions on*, vol. 13, no. 4, pp. 600–612, 2004.