IMAGE BIT-DEPTH ENHANCEMENT VIA MAXIMUM-A-POSTERIORI ESTIMATION OF GRAPH AC COMPONENT

Pengfei Wan[o], Gene Cheung[#], Dinei Florencio[$], Cha Zhang[$], Oscar C. Au[o]

[o] Hong Kong University of Science and Technology, [#] National Institute of Informatics, Japan
[$] Microsoft Research Redmond

ABSTRACT

While modern displays offer high dynamic range (HDR) with large bit-depth for each rendered pixel, the bulk of legacy image and video contents were captured using cameras with shallower bit-depth. In this paper, we study the bit-depth enhancement problem for images, so that a high bit-depth (HBD) image can be reconstructed from an input low bit-depth (LBD) image. The key idea is to apply appropriate smoothing given the constraints that reconstructed signal must lie within the per-pixel quantization bins. Specifically, we first define smoothness via a signal-dependent graph Laplacian, so that natural image gradients can nonetheless be interpreted as low frequencies. Given defined smoothness prior and observed LBD image, we then demonstrate that computing the most probable signal via maximum a posteriori (MAP) estimation can lead to large expected distortion. However, we argue that MAP can still be used to efficiently estimate the AC component of the desired HBD signal, which along with a distortion-minimizing DC component, can result in a good approximate solution that minimizes the expected distortion. Experimental results show that our proposed method outperforms existing bit-depth enhancement methods in terms of reconstruction error.

Index Terms— Bit-depth enhancement, graph signal processing

1. INTRODUCTION

It is undeniable that there exists an insatiable human desire to create bigger and more realistic displays. In terms of spatial resolution (number of pixels per image), television has evolved from VGA (640×480) to HD (1280×720), and soon to 4K and 8K ultra HD (3840×2160 and 7680×4320 respectively). In terms of bit-depth (number of bits per pixel), *high dynamic range* (HDR) technologies have promised 10 to 12 bits per pixel—as opposed to conventional 8 bits per pixel—for finer-grained quantization of real pixel values to discrete levels. However, though display technologies have continued to improve, the bulk of legacy image and video contents were captured using older capturing devices, often in lower spatial resolution and shallower bit-depth than what modern displays are capable. Thus there is a need to suitably increase the spatial resolution and/or bit-depth of legacy content. *Super-resolution* [1] addresses the first problem of increasing the spatial resolution. In contrast, in this paper we address the second problem of bit-depth enhancement.

Common in the literature of bit-depth enhancement [2–6] is the notion of *smoothness*. The key observation is that true image signals tend to be smooth, and thus given an observed low bit-depth (LBD) signal, applying a smoothing operator across consecutive quantization levels would likely result in a better quality signal. As an illustration, we see in Fig. 1 an 8-sample one-dimensional (1D) signal $\mathbf{y}$ quantized to integer values 1 and 2. If a smoothing operator is applied, it can result in $\mathbf{x}_1$ or $\mathbf{x}_2$, depending on the amount of smoothing applied. While the notion of smoothness is intuitive, defining it mathematically rigorously and applying it optimally for high bit-depth (HBD) signal reconstruction is not trivial. In this paper, leveraging on recent advances in *graph signal processing* (GSP) [7, 8], we first formally define "smoothness" via a signal-dependent graph Laplacian. Specifically, corresponding graph transform can decompose a graph-signal (pixel values in this work) into graph frequency components, and a graph-signal is deemed smooth if it contains mainly low graph frequencies. Further, unlike spectral decomposition based on fixed transforms like discrete cosine transform (DCT), signal structure can be embedded as edge weights in the graph, so that natural image gradients will nonetheless be interpreted as low graph frequencies, lowering the chance of over-smoothing.

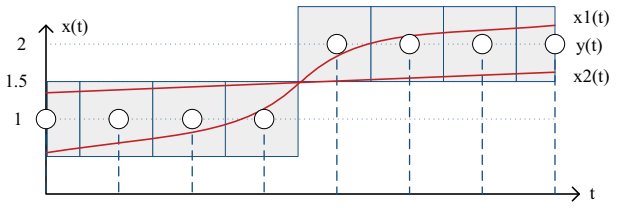


Fig. 1: Examples of quantized and smoothed 1D signals.

Next, armed with our defined smoothness prior for graph-signals, we formulate an optimization problem for reconstructed signal $\mathbf{x}^*$ that minimizes the expected distortion given quantized signal $\mathbf{y}$. Observing that the optimization is difficult to solve directly, perhaps surprisingly, we next demonstrate that the most probable signal via *maximum a posteriori* (MAP) estimation given observed $\mathbf{y}$ and the smoothness prior can lead to arbitrarily large expected distortion, and thus applying MAP directly is a poor proxy for the original optimization objective. However, we argue that with a simple twist—computing the most probable AC component of the reconstructed signal first via MAP and then a distortion-minimizing DC component subsequently—MAP can still be a useful and efficient tool which results in good approximate solutions to the original problem. Experiments demonstrate that our proposed MAP estimation for AC component in graph transform domain has good objective bit-precision enhancement performances for both natural and synthesized images.

The outline of the paper is as follows. We first overview related work in Section 2. We then introduce the construction of a graph transform in Section 3. Proposed MAP formulation for image bit-depth enhancement and the corresponding optimization procedure are elaborated in Section 4 and Section 5. Finally, we show experiment results and draw conclusions in Section 6 and 7, respectively.

This work is supported in part by Hong Kong Research Grant Council (GRF), as well as the Microsoft Research CORE program.

2. RELATED WORK

Graph signal processing (GSP) is the study of signals that live on structured data kernels described by graphs [8]. GSP tools can also be applied to traditional signals such as images that live on regular kernels [9, 10] or point cloud structures [11], where the idea is to embed signal structure into the graph before signal processing. Similarly, to reconstruct a HBD signal we compute edge weights of a graph based on the signal structure (deduced from the observed input LBD image) and then define signal smoothness via the graph Laplacian, so that a signal with enhanced bit-precision can be reconstructed without over-smoothing natural image gradients.

Coarse quantization may cause false contours which degrade the visual quality of the image. In sense of objective quality, coarse quantization results in low bit-precision of the acquired image signal and thus large quantization error. Previous works on false contour removal and bit-precision enhancement [2–6] are typically smoothing schemes by filtering or spatial interpolation, which do not optimize an objective metric such as mean-square-error (MSE). Our work is a significant improvement over existing bit-depth enhancement works in that *we are the first in the literature to define signal smoothness formally using GSP tools for the bit-depth enhancement problem, and propose a computation-efficient MAP algorithm that produces good approximate solutions minimizing the expected distortion*. Our algorithm can also be used for broader applications that require bit-depth enhancement: *e.g.*, 3D surface refinement by enhancing the bit-precision of depth maps [12]; compression schemes that encode an image at shallower bit-depth than captured for bitrate saving, and then recover the least significant bits (LSB) at decoder [13].

We stress that the bit-precision enhancement problem is different from the inverse tone mapping (ITM) problem [14, 15] in HDR imaging. Specifically, the source of distortion is different: distortion in ITM is typically caused by non-linear tone mapping operator [14] or the over-saturation of camera sensors [15], while distortion in bit-precision enhancement problem is introduced by A/D conversion (quantization). Thus, the desired output of bit-precision enhancement does not hallucinate lost details for improved subjective quality, but estimates the original HBD signal by minimizing the expected distortion, as formulated in Section 4.

3. GRAPH TRANSFORM

We begin with a review of GSP concepts such as graph Laplacian and graph transform. A graph $\mathcal{G} = (\mathcal{V}, \mathcal{W})$ is defined by a set $\mathcal{V}$ of N vertices, and a set $\mathcal{W}$ of non-negative edge weights. Each vertex (pixel) i has associated signal intensity (pixel value) $x(i)$, so a graph-signal can be written as a vector $\mathbf{x} = [x(1), \cdots, x(N)]^T \in \mathbb{R}^N$. An undirected edge with weight $w(i,j)$ connects vertices i and j iff $w(i,j) > 0$.

Given graph $\mathcal{G}$, we can define the *adjacency matrix* $\mathbf{A} \in \mathbb{R}^{N\times N}$, where $A(i,j) = w(i,j) \geq 0$. The *degree matrix* $\mathbf{D} \in \mathbb{R}^{N\times N}$ is a diagonal matrix satisfying $D(i,i) = \sum_j A(i,j)$. The *combinatorial graph Laplacian* (graph Laplacian for short) for graph $\mathcal{G}$ is then:

$$\mathbf{L} = \mathbf{D} - \mathbf{A} \tag{1}$$

The graph Laplacian can be normalized to be the *normalized graph Laplacian* $\mathbf{L}^n = \mathbf{D}^{-\frac{1}{2}}\mathbf{L}\mathbf{D}^{-\frac{1}{2}}$. A *random walk Laplacian* is similarly defined as $\mathbf{L}^r = \mathbf{D}^{-1}\mathbf{L}$. For the purpose of this paper, we will only use the combinatorial graph Laplacian $\mathbf{L}$. See [16] for a more detailed discussion on different Laplacian variants.

Graph transform (GT), also known as *graph Fourier transform* [8], involves the eigen-decomposition of the graph Laplacian $\mathbf{L}$ into a matrix $\mathbf{T}$ composed of eigenvectors $\mathbf{v}_k$'s as rows. GT introduces the notion of frequencies into a graph-signal. Graph frequencies are $\mathbf{L}$'s eigenvalues λ_k's, with $\mathbf{v}_k$'s as associated basis vectors. In particular, λ_k measures the "smoothness" of $\mathbf{v}_k$:

$$\lambda_k = \mathbf{v}_k^T\mathbf{L}\mathbf{v}_k = \frac{1}{2}\sum_{i,j} w(i,j)(v_k(i) - v_k(j))^2 \tag{2}$$

So when the basis vectors are sorted in ascending order of their corresponding eigenvalues, the transform coefficient vector $\boldsymbol{\alpha} = \mathbf{T}\mathbf{x} = [\alpha_1, \cdots, \alpha_N]^T$ represents the amount of low-to-high graph frequency components in graph-signal $\mathbf{x}$.

4. PROBLEM SETUP

We begin with a set of definitions for later derivation. Without loss of generality, a length-N original signal $\mathbf{x}^o \in \mathbb{R}^N$ is quantized to $\mathbf{y} \in \mathbb{I}^N$ using quantization stepsize Q, *i.e.*, $\mathbf{y} = \text{round}(\mathbf{x}^o/Q)$. Q is known at both encoder and decoder. At the decoder, only $\mathbf{y}$ is observed and we are tasked to find the "best" estimate $\mathbf{x}^*$ to $\mathbf{x}^o$.

In this work, vertices (pixels) i and j are connected by an edge only when they are adjacent. Due to quantization, $|y(i) - y(j)| > 1$ means $x^o(i)$ and $x^o(j)$ cannot be equal, so we set edge weight $w(i,j) = 0$ in this case. For the other adjacent pixel pairs, we set $w(i,j) = \exp\{-(y(i) - y(j))^2/\sigma^2\}$. Proper assignment of edge weights embeds the image structure into the graph definition.

Graph-signal $\mathbf{x}$, whose entries are pixel values, takes on a prior probability $Pr(\mathbf{x})$. Here we define the following smoothness prior that favors low-frequency components in graph transform domain:

$$\begin{aligned} Pr(\mathbf{x}) = Pr(x_D)Pr(\mathbf{x}_A) &= C\frac{1}{K_1}\exp\left\{-\sigma_l^2\mathbf{x}^T\mathbf{L}^p\mathbf{x}\right\} \\ &= \frac{C}{K_1}\exp\left\{-\sigma_l^2\sum_{i=2}^{N}\lambda_i^p\alpha_i^2\right\} \end{aligned} \tag{3}$$

where λ_i and α_i are sorted graph frequency[1] and the corresponding graph transform coefficient respectively, and p is a positive integer. K_1 is the normalization factor for $Pr(\mathbf{x}_A)$, the prior on the AC component of $\mathbf{x}$. Constant $C = Pr(x_D)$ is the uniform prior on the DC component x_D. In words, (3) states that a signal $\mathbf{x}$ is more probable if it is smooth within a p-hop neighborhood. Different from DCT where a smooth signal prior tends to blur the natural image gradients, reconstructing a smooth signal in graph transform domain does not contradict with the natural image gradients because image structures are embedded in graph weights. Hence smoothness prior in the graph transform domain is more suitable for our problem.

4.1. Distortion Minimization

Ideally, we would like to minimize the expected distortion of the reconstructed signal $\mathbf{x}^*$ with respect to the original signal $\mathbf{x}^o$,

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} \int_{\mathbf{x}^o\in\mathcal{F}(\mathbf{y})} \|\mathbf{x} - \mathbf{x}^o\|_2^2 \, Pr(\mathbf{x}^o \,|\, \mathbf{y}) \, d\mathbf{x}^o \tag{4}$$

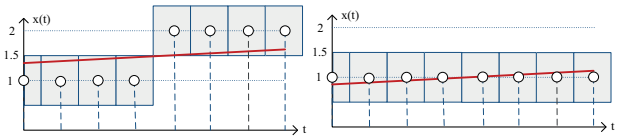
where $\mathcal{F}(\mathbf{y})$ is the feasible space of original signal $\mathbf{x}^o$ given observed quantized signal $\mathbf{y}$, *i.e.*:

$$\mathcal{F}(\mathbf{y}) = \{\mathbf{x}^o \mid \text{round}(\mathbf{x}^o/Q) = \mathbf{y}\} \tag{5}$$

(4) can be viewed as a "sum" of squared errors each weighted by $Pr(\mathbf{x}^0|\mathbf{y})$. It is difficult to solve directly because each signal $\mathbf{x}^o$ in the sum induces a different squared error with a different weight, and given the large space of signals $\mathbf{x}^o \in \mathcal{F}(\mathbf{y})$ to consider, it is hard to keep track of them all in one optimization procedure.

[1](3) uses the fact that $\lambda_1 \equiv 0$ in graph transform (zero frequency).

4.2. MAP Formulation for Reconstructed Signal



(a) quantization bin set 1 (b) quantization bin set 2

Fig. 2: Example showing the same AC signal $\mathbf{x}_A$ being quantized to different sets of quantization bins.

(a) Input image $\mathbf{y}$ (b) Output by (7) (c) Our output

Fig. 3: Example showing MAP estimation of the signal leads to a DC output (b), while MAP of AC component reconstructs a smooth signal (c).

Suppose instead of minimizing the expected distortion using (4), we simply find the most probable $\mathbf{x}$ given observed $\mathbf{y}$ instead:

$$\mathbf{x}^* = \arg \max_{\mathbf{x} \in \mathcal{F}(\mathbf{y})} Pr(\mathbf{x} \,|\, \mathbf{y}) \tag{6}$$

In relation to original objective (4), it means that we first find $\mathbf{x}^o$ in the space $\mathcal{F}(\mathbf{y})$ with the largest weight $Pr(\mathbf{x}^o|\mathbf{y})$, and then set the variable $\mathbf{x}^*$ to $\mathbf{x}^o$ so their corresponding squared error is zero in the sum in (4). (6) can be rewritten as follows via MAP:

$$\mathbf{x}^* = \arg \max_{\mathbf{x} \in \mathcal{F}(\mathbf{y})} Pr(\mathbf{y}|\mathbf{x})\, Pr(\mathbf{x}) \tag{7}$$

where likelihood $Pr(\mathbf{y}|\mathbf{x})$ evaluates to 1 if $\mathbf{x} \in \mathcal{F}(\mathbf{y})$ and 0 otherwise. (7) is thus equivalent to finding the signal $\mathbf{x}$ with the largest probability $Pr(\mathbf{x}) = Pr(\mathbf{x}_A)Pr(x_D)$ in $\mathcal{F}(\mathbf{y})$.

We show that *MAP solution (7) can lead to arbitrarily large expected distortion given quantization step size Q*, and thus (6) is not a good proxy for the original objective (4). See Fig. 2(a) for an example set of quantization bins corresponding to observed $\mathbf{y}$. According to prior probability (3), the most probable signal is one that is essentially DC, *e.g.* $\mathbf{x} = [1.5 - 4\epsilon, \ldots, 1.5 + 4\epsilon]$ for arbitrarily small ϵ, regardless of Q. It is clear that for large Q, this solution will result in large expected distortion.[2] The reason for this result is because the MAP formulation for reconstructed signal $\mathbf{x}$ only accounts for the most probable signal $\mathbf{x}^o$ with probability $Pr(\mathbf{x}_A)Pr(x_D)$ in the sum in (4) while ignoring all other signals in space $\mathcal{F}(\mathbf{y})$.

4.3. MAP Formulation for Reconstructed AC Signal

Instead of solving for the most probable $\mathbf{x}$ via a MAP formulation, *we propose to estimate the most probable AC component $\mathbf{x}_A$ via MAP first*. After acquiring $\mathbf{x}_A^*$, if we can assume that $\mathbf{x}_A^*$ is reasonably accurate and $\mathbf{x}_A^* \approx \mathbf{x}_A^o$ (to be argued below), the distortion-minimizing DC component x_D^* can then be solved easily:

$$\arg \min_{x_D} \int_{\mathbf{x}^o \in \mathcal{F}(\mathbf{y})} \|x_D + \mathbf{x}_A^* - (x_D^o + \mathbf{x}_A^*)\|_2^2 \; Pr(x_D^o + \mathbf{x}_A^*|\mathbf{y}) \; d\mathbf{x}^o \tag{8}$$

[2]This is not an isolated case in which MAP solution performs poorly. A similar case is that quantized $\mathbf{y}$ alternates between 1 and 2.

By (3), x_D is uniformly distributed, so $Pr(x_D^o + \mathbf{x}_A^*|\mathbf{y})$ is constant so long as x_D^o leads to quantized $\mathbf{y}$ according to (5), *i.e.*

$$y(i) - \frac{Q}{2} \le x_A^*(i) + x_D^o \le y(i) + \frac{Q}{2}, \quad \forall i \tag{9}$$

Thus (9) establishes the integration bounds in (8), resulting in:

$$x_D^* = \arg \min_{x_D} \int_{z_{\max} + y_D - \frac{Q}{2}}^{z_{\min} + y_D + \frac{Q}{2}} \|x_D - x_D^o\|_2^2 \; dx_D^o \tag{10}$$

where vector $\mathbf{z} = \mathbf{y}_A - \mathbf{x}_A^*$, $z_{\min}$ and $z_{\max}$ are the smallest and largest scalars in vector $\mathbf{z}$. The solution to (10)—distortion-minimizing DC component—is simply $x_D^* = y_D + (z_{\min} + z_{\max})/2$.

We now argue that the MAP solution $\mathbf{x}_A^*$ does give a good estimate to true AC signal $\mathbf{x}_A^o$, *i.e.* the expected distortion is small:

$$\int_{\mathbf{x}_A^o} \|\mathbf{x}_A^* - \mathbf{x}_A^o\|_2^2 \; Pr(\mathbf{x}_A^o|\mathbf{y}) \; d\mathbf{x}_A^o \tag{11}$$

Seeking the MAP estimate $\mathbf{x}_A^*$ means solving:

$$\arg \max_{\mathbf{x}_A} Pr(\mathbf{x}_A|\mathbf{y}) = \arg \max_{\mathbf{x}_A} Pr(\mathbf{y}|\mathbf{x}_A)\, Pr(\mathbf{x}_A) \tag{12}$$

$$= \arg \min_{\mathbf{x}_A} - \log Pr(\mathbf{y}|\mathbf{x}_A) - \log Pr(\mathbf{x}_A) \tag{13}$$

By total probability theorem, the likelihood can be written as:

$$Pr(\mathbf{y}|\mathbf{x}_A) = \int Pr(\mathbf{y}|\mathbf{x}_A, x_D) Pr(x_D) \; dx_D \tag{14}$$

where again $Pr(\mathbf{y}|\mathbf{x}_A, x_D)$ evaluates to 1 if $x_D + \mathbf{x}_A \in \mathcal{F}(\mathbf{y})$ and 0 otherwise, and $Pr(x_D) = C$. So we can rewrite (12) as:

$$\mathbf{x}_A^* = \arg \max_{\mathbf{x}_A} \int_{x_D | x_D + \mathbf{x}_A \in \mathcal{F}(\mathbf{y})} C \; dx_D \; Pr(\mathbf{x}_A) \tag{15}$$

We can interpret (15) as follows. MAP solution $\mathbf{x}_A^*$ and other AC signals $\mathbf{x}_A^o$ in the "sum" in (11) with large weight $Pr(\mathbf{x}_A^o|\mathbf{y})$ must have large feasible ranges for x_D to integrate over, as shown in Fig. 2(b) (as opposed to Fig. 2(a)). By (10), x_D feasible range given AC signal $\mathbf{x}_A^*$ is $Q + z_{\min} - z_{\max}$. That means AC signals with large feasible range of x_D have small $z_{\max} - z_{\min}$, *i.e.* they are all close to $\mathbf{y}_A$ and thus are similar to each other. Hence solving for $\mathbf{x}_A^*$ via MAP means it reduces distortion for all signals in the sum in (11) with the largest weights. Thus, MAP solution $\mathbf{x}_A^*$ has small expected distortion and provides a good estimate to the true AC signal $\mathbf{x}_A^o$.

We discuss how (13) can be efficiently solved next.

5. MAP SIGNAL ESTIMATION

5.1. Computing Likelihood

Likelihood (14) is still hard to compute when $\mathbf{x}_A$ becomes the optimization variable. We approximate $Pr(\mathbf{y}|\mathbf{x}_A)$ as follows. We have argued earlier that an AC signal $\mathbf{x}_A$ close to the AC component of $\mathbf{y}$ (the centers of quantization bins) allows x_D to vary to a large extent while still keeping $x_D + \mathbf{x}_A \in \mathcal{F}(\mathbf{y})$, thus leads to a large $Pr(\mathbf{y}|\mathbf{x}_A)$. Generalizing the above observation, we approximate the likelihood $Pr(\mathbf{y}|\mathbf{x}_A)$ as follows for efficient solving:

$$\approx \begin{cases} \frac{1}{K_2} \exp\left\{-\sigma_q^2 \|\mathbf{y}_A - \mathbf{x}_A\|_2^2\right\} & \text{if } |x_A(i) - y_A(i)| < \frac{Q}{2}, \;\; \forall i \\ 0 & \text{o.w.} \end{cases} \tag{16}$$

where K_2 is the normalization factor. In words, (16) states if $\mathbf{y}_A$ and AC signal $\mathbf{x}_A$ are a good match, then the likelihood of $\mathbf{x}_A$ is high. Further, we impose a sufficient condition where each $x_A(i)$ cannot deviate from $y_A(i)$ by more than $Q/2$ to ensure feasibility in (5).

(a) 4-bit input image, 8-bit ground-truth

Image	Anchor	De-cont	Interp	Proposed
lampshade2	3.554e-4	7.250e-4	3.275e-4	**2.529e-4**
plastic	2.233e-4	5.510e-4	3.189e-4	**1.499e-4**
midd2	2.709e-4	5.993e-4	1.854e-4	**1.687e-4**
monopoly	3.404e-4	6.129e-4	3.061e-4	**2.395e-4**

(b) 8-bit input image, 12-bit ground-truth

Image	Anchor	De-cont	Interp	Proposed
dude	1.217e-6	2.443e-6	6.020e-7	**3.227e-7**
sphere	1.279e-6	2.577e-6	4.752e-7	**1.561e-7**

Table 1: MSE Performance Comparisons

5.2. Linearly Constrained Quadratic Programming

By combining (3), (5) and (16), we formulate the problem (13) as:

$$\min_{\mathbf{x}_A} \quad \sigma_l^2 \mathbf{x}_A^T \mathbf{L}^p \mathbf{x}_A + \sigma_q^2 \|\mathbf{x}_A - \mathbf{f}_A\|_2^2 \tag{17}$$
$$\text{s.t.} \quad |x_A(i) - y_A(i)| < Q/2, \forall i, \qquad \sum_i x_A(i) = 0$$

where fidelity $\mathbf{f}_A$ is initialized to be $\mathbf{y}_A$ by (16). After solving the quadratic problem (17) for the first time, we obtain the output image denoted by $\mathbf{x}_A^{*(1)}$. We then iteratively solve (17), where the solution $\mathbf{x}_A^{*(k+1)}$ for $(k+1)$-th iteration is obtained by solving (17) with updated fidelity $\mathbf{f}_A = \mathbf{x}_A^{*(k)}$, as $\mathbf{x}_A^{*(k)}$ is the best estimate of $\mathbf{x}_A^o$ to-date. At each iteration we update the edge weights based on the latest computed pixel values. Solving (17) yields $\mathbf{x}_A^*$, which is subsequently used to obtain the distortion-minimizing DC component x_D^* via (10). Our reconstructed image signal is thus $\mathbf{x}^* = \mathbf{x}_A^* + x_D^*$.

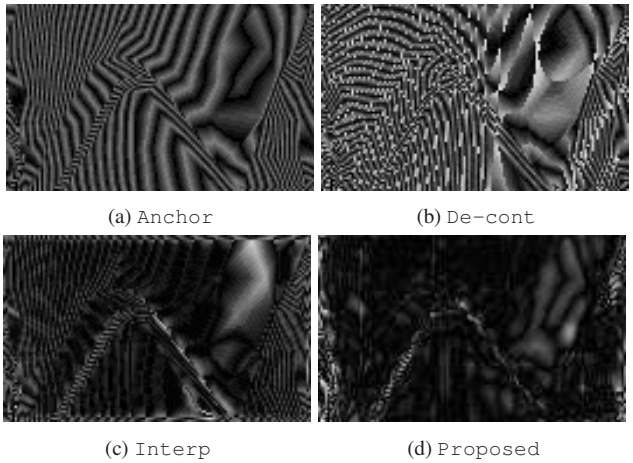
6. EXPERIMENTATION

We conduct the following bit-depth enhancement experiments using both natural color images and synthesized images. A HBD image serves as the ground-truth signal $\mathbf{x}^o$. Our input is the LBD image $\mathbf{y}$ with quantization step size Q. Input image is divided into overlapped blocks of size 64×64 for robust and efficient computation of our posed quadratic problem (17) with $p = 5$. The quality of the reconstructed image $\mathbf{x}^*$ is measured in MSE with respect to $\mathbf{x}^o$. For color images, average MSE of R, G and B channels is used. Smaller MSE indicates higher bit-precision of $\mathbf{x}^*$.

The competing methods in our experiments are: 1) `Anchor` which picks quantization bin centers as the reconstructed image; 2) `De-cont` [6] which is a filtering-based method for removing false contours; 3) `Interp` [4] which is a linear interpolation method in spatial domain; and 4) our proposed MAP estimation for AC component in graph transform domain.

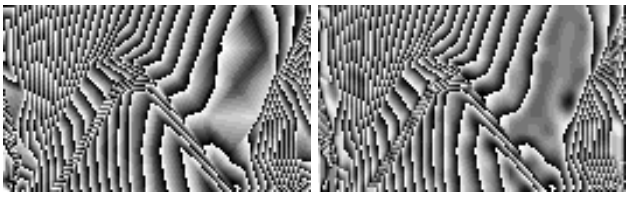
Two sets of bit-depth enhancement experiments are conducted. In 4-bit enhancement experiments, 8-bit natural color images (`lampshade2`, `plastic`, `midd2`, `monopoly`) serve as ground-truth $\mathbf{x}^o$ and input $\mathbf{y}$ is coarsely quantized to 4-bit. These 4 color images are all from the Middlebury 2006 stereo dataset [17]. In 8-bit enhancement experiments, grayscale images `dude` and `sphere` with 12-bit precision are used as $\mathbf{x}^o$ and the input $\mathbf{y}$ is coarsely quantized to 8-bit. These 2 images are computer-generated depth maps of a human body model and a 3D sphere, respectively.

Fig. 4 shows the scaled absolute reconstruction error $|\mathbf{x}^* - \mathbf{x}^o|$ for the four methods. We see that our method achieves the best performance with the smallest error energy. That is because our method



(a) `Anchor` (b) `De-cont` (c) `Interp` (d) `Proposed`

Fig. 4: Absolute error maps of four comparing methods in 8-bit enhancement experiment for `dude` (scaled for visibility).



(a) LSBs of ground-truth $\mathbf{x}^o$ (b) LSBs of $\mathbf{x}^*$ by proposed method

Fig. 5: LSB maps corresponding to the same experiment as Fig. 4.

embeds image patch structures into the graph weights, enabling the reconstruction of complicated signals without over-smoothing.

Alternatively, the bit-precision enhancement problem can be viewed as the LSBs reconstruction problem if the target bit-depth of $\mathbf{x}^o$ is known. Fig. 5 shows the LSB maps for $\mathbf{x}^o$ and our output $\mathbf{x}^*$. We see that proposed method well restores the ground-truth LSBs.

The numerical MSE results are summarized in Table. 1, which show that our method achieves the best objective performance in both experiments. `Anchor` does not take inter-pixel correlations into consideration, so the reconstruction error is large. `De-cont` always performs the worst in terms of MSE because of its inability to capture the image structures (various shapes and sizes of quantized regions). `Interp` works quite well for synthesized images (`dude`, `sphere`); but when dealing with natural images, its simple linear interpolation cannot well approximate the original signal $\mathbf{x}^o$ which can be highly irregular. In contrast, our method performs consistently well in 4-bit/8-bit enhancement experiments using signal-dependent graph transform and proposed MAP formulation.

7. CONCLUSION

We proposed an image bit-depth enhancement algorithm that uses MAP to estimate the AC component of an image patch in graph frequency domain; smoothness defined in the graph domain means image structures can be embedded into the signal prior, avoiding the problem of over-smoothing. Our MAP formulation can be efficiently solved via quadratic programming with linear constraints. Experiments show that our proposed method outperforms competing methods in terms of objective reconstruction error. Our investigation also serves as a cautionary tale: though MAP remains an effective tool, if improperly used, it can lead to arbitrarily large expected errors, as we demonstrated in our problem setting.

8. REFERENCES

[1] P. Milanfar, *Super-Resolution Imaging (Digital Imaging and Computer Vision)*, CRC Press, September 2010.

[2] Min-Ho M.-H. Park, J. W. Lee, Rae-Hong R.-H. Park, and J.-S. Kim, "False contour reduction using neural networks and adaptive bi-directional smoothing," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 2, pp. 870–878, 2010.

[3] X. Jin, S. Goto, and K. N. Ngan, "Composite model-based dc dithering for suppressing contour artifacts in decompressed video," *IEEE Transactions on Image Processing*, vol. 20, no. 8, pp. 2110–2121, 2011.

[4] 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 International Conference on Multimedia & Expo (ICME)*, Melbourne, Australia, 2012.

[5] P. Wan, O. C. Au, K. Tang, and Y. Guo, "Image de-quantization via spatially varying sparsity prior," in *IEEE International Conference on Image Processing*. IEEE, 2012, pp. 953–956.

[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] P. Milanfar, "A tour of modern image filtering: new insights and methods, both practical and theoretical," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 106–128, 2013.

[8] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," in *IEEE Signal Processing Magazine*, May 2013, vol. 30, no.3, pp. 83–98.

[9] W. Hu, G. Cheung, X. Li, and O. Au, "Depth map compression using multi-resolution graph-based transform for depth-image-based rendering," in *IEEE International Conference on Image Processing*, Orlando, FL, September 2012.

[10] Y. Mao, G. Cheung, and Y. Ji, "Graph-based interpolation for dibr-synthesized images with nonlocal means," in *Symposium on Graph Signal Processing in IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, Austin, TX, December 2013.

[11] C. Zhang, , D. Florencio, and C. Loop, "Point cloud attribute compression with graph transfrom," in *IEEE International Conference on Image Processing*. IEEE, October 2014.

[12] P. Wan, G. Cheung, P. A. Chou, D. Florencio, C. Zhang, and O. C. Au, "Precision enhancement of 3D surfaces from multiple quantized depth maps," in *11th IEEE IVMSP Workshop: 3D Image/Video Technologies and Applications*, Seoul, Korea, 2013.

[13] V. Nguyen, D. Min, and M. Do, "Efficient techniques for depth video compression using weighted mode filtering," in *IEEE Transactions on Circuits and Systems for Video Technology*, February 2013, vol. 23, no.2.

[14] F. Banterle, P. Ledda, K. Debattista, and A. Chalmers, "Inverse tone mapping," in *Proceedings of the 4th international conference on Computer graphics and interactive techniques in Australasia and Southeast Asia*. ACM, 2006, pp. 349–356.

[15] A. Rempel et. al, "Ldr2hdr: on-the-fly reverse tone mapping of legacy video and photographs," *ACM Transactions on Graphics (TOG)*, vol. 26, no. 3, pp. 39, 2007.

[16] T. Biyikoglu, J. Leydold, and P. F. Stadler, "Nodal domain theorems and bipartite subgraphs," in *Electronic Journal of Linear Algebra*, November 2005, vol. 13, pp. 344–351.

[17] H. Hirschmuller and D. Scharstein, "Evaluation of cost functions for stereo matching," in *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, June 2007, pp. 1–8.