QUALITY AND RATE CONTROL OF JPEG XR

A Thesis

by

YIDING LIU

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Chair of Committee,	Zixiang Xiong
Committee Members,	Erchin Serpedin
	Ulisses Braga-Neto
	Jyh-Charn Steve Liu
Head of Department,	Miroslav M. Begovic

August 2016

Major Subject: Electrical Engineering

ABSTRACT

Driven by the need for seismic data compression with high dynamic range and 32-bit resolution, we propose two algorithms to efficiently and precisely control the signal-to-noise ratio (SNR) and bit rate in JPEG XR image compression to allow users to compress seismic data with a target SNR or a target bit rate. Based on the quantization properties of JPEG XR and the nature of blank macroblocks, we build a reliable model between the quantization parameter (QP) and SNR. This enables us to estimate the right QP with target quality for the JPEG XR encoder.

DEDICATION

To my father, Dr. Xiaoyang Liu, and my mother, Wei Han.

ACKNOWLEDGEMENTS

I would like to especially thank my committee chair, Dr. Zixiang Xiong, and my committee members, Dr. Erchin Serpedin, Dr. Ulisses Braga-Neto and Dr. Jyh-Charn Steve Liu, for their guidance and support throughout the course of this research.

Thanks also go to my friends and colleagues, Bo Wang, Ziwei Xuan and Yan Meng for making my time at Texas A&M University a great experience.

Finally, thanks to my father, mother and Nan Jiang for their encouragement, support and love.

NOMENCLATURE

AC	Alternating Current
CR	Compression Ratio
DC	Direct Current
DCT	Discrete Cosine Transform
EBCOT	Embedded Block Coding with Optimal Truncation
FLC	Fixed Length Coding
HP	High Pass
LBT	Lapped Biorthogonal Transform
LP	Low Pass
MSE	Mean Square Error
PCT	Photo Core Transform
POT	Photo Overlap Transform
QP	Quantization Parameter
SF	Scaling Factor
SNR	Signal-to-Noise Ratio
VLC	Variable Length Coding

TABLE OF CONTENTS

Page

LIST OF FIGURES

1. INTRODUCTION

JPEG XR is a new generation still image compression standard which supports a number of competitive features including lossless and lossy coding, adaptive quantization, high dynamic range, overlapped transform, region-of-interest decoding and multiple color models, etc [1]. JPEG XR has excellent performance in the efficiency of coding, visual quality and low required memory compared to JPEG and JPEG 2000 image compression standard [2][3]. Therefore, JPEG XR is friendly to heavy workloads like large scale images and hardware with low configuration.

JPEG XR offers a quality control by dividing coefficients by the quantization parameter, but the quantization parameter does not lead to any specified parameter of output images like bit rate or signal-to-noise ratio (SNR). However, it is significant for streaming transmission limited by bandwidth or tradeoff between disk space and image quality in data storage so that images are able to be compressed with an expected bit rate or SNR efficiently.

Particularly, huge seismic data are generated nowadays for gas and oil exploration. In general, seismic data contains 3D wavefields and 2D traces where some 2D traces have millions of pixels and most seismic traces and wavefields are 32-bit floating-point data. Therefore, to meet the demand of industry, a new compression method is obviously needed to control the compression ratio and SNR of both high bit-depth and large scale images efficiently and precisely.

1.1 Related work

Nowadays, JPEG represents one of the most commonly used lossy compression standards in the world. However, since its first standard (ITU-T T.81 | ISO/IEC 10918-1) was released in 1992, many extended standards were released with the development of new technologies and the demand for higher performance. JPEG 2000 standard (ITU-T T.800 | ISO/IEC 15444-1) was firstly introduced to supersede the discrete cosine transform in JPEG with the wavelet transform and a number of new features in 2000. JPEG XR is the latest image compression standard (ITU-T T.832 | ISO/IEC 29199-2) released by the JPEG standardization committee in 2009 and was formerly known as HD Photo proposed by Microsoft since 2007. JPEG XR presents superior performance in terms of compression based on the comparison between H.264/AVC Intra-Frame, JPEG2000, and JPEG XR according to Tran et al. [2]. De Simone et al. [4] evaluated the subjective performance of JPEG, JPEG 2000 and JPEG XR, and JPEG XR is also competitive in terms of subjective visual quality.

There are lots of proposed rate control methods for multiple codecs. JPEG does not offer a feature to estimate bit rate, so several iterative control methods are designed based on fast converging functions [5][6]. Although JPEG 2000 supports to control rate distortion by its Embedded Block Coding with Optimal Truncation (EBCOT) scheme, the process requires encoding all the image before the truncating coefficients. Methods were proposed to increase the efficiency of JPEG 2000 encoder [7]. However, these methods are limited by their target formats. He et al. [8] proposed a general method to analyze the relationship between bit rate and SNR called ρ-domain analysis and Chan et al. [9] applied the method on rate control of JPEG XR. On the other hand, JPEG XR support block-based quantization [10] which means that adaptive bit rate methods for video are possible to be used in block-based JPEG XR image. However, the work from Richter et al. [11] indicates that uniform quantization is always optimal or close to optimality among adaptive quantization methods.

For seismic data compression, a licensed commercial wavelet-based compression method was proposed by [12], and it is widely used by the industry and also supports rate control.

1.2 Proposed approach

This thesis focuses on precise and efficient quality control of the JPEG encoder with seismic data. Although a subjective quality control scheme is embedded in the JPEG XR standard, it does not lead to stable objective results. Thus, this thesis focuses on controlling the quality with two kinds of parameters: SNR and bit rate.

This thesis proposes two methods to respectively control SNR and bit rate. The proposed SNR-control method is considered to be a fast solution by evaluating the quantization parameter directly during preprocessing. The additional complexity is limited to much less than the complexity of transform and coding. The proposed rate-control method reduces the processing time by accelerating the convergence of iterative parameter searching.

The critical features regarding the performance of the proposed encoder are:

1. The time complexity of the new encoder should be low enough;

2. The SNR of results from the proposed encoder should be as high as the standard results at the same bit rate;

3. The bit rate of results from the proposed encoder should be as low as the standard results at the same SNR.

4. The performance of the proposed methods should be competitive with the conventional scheme which should open the possibility of extensive uses of JPEG XR.

1.3 Contribution

The main contributions of this thesis to the JPEG XR encoding process are the precise and efficient control of the objective quality which should support high dynamic range floating data as well. The standard JPEG XR encoder scheme previously mentioned offers features including subjective quality control [4] and high dynamic range coding [17], but no quality control is supported to obtain stable images with given objective quality parameters like SNR and bit rate. Another limitation of built-in floating coding is low efficiency in the lossy mode. Although a number of methods were proposed to control the image compression ratio [5][6][7], these algorithms are format-related. In JPEG XR, only one proposed rate control algorithm [9] mainly focused on the relationship between bit rate and distortion and is not precise enough.

Another contribution of the proposed algorithms is to satisfy the demands of seismic data compression with an open source format in the industry since currently the widely used compression method is commercial. Therefore, the proposed algorithms are a great alternative for seismic data compression due to multiple features and efficient control performance no lower than the previous one.

1.4 Thesis overview

This thesis introduces two methods to precisely and efficiently control the SNR and compression ratio of output images in JPEG XR and an extended encoder to implement the algorithms based on the standard mechanism of JPEG XR encoder developed by Microsoft.

The second chapter illustrates the JPEG XR technology applied in image encoder including data hierarchy, two-stage transform, quantization, coefficient prediction and entropy encoding.

In the third chapter, a method to control SNR of JPEG XR is introduced. It indicates the nature of JPEG XR quantizer and the relationship between SNR and quantization parameter. An extended JPEG XR encoder is developed to implement the proposed algorithm. Assessment tests are performed on the encoder in terms of the accuracy of SNR results, the performance of the compression ratio and the computational complexity. The corresponding simulations results are presented.

In the fourth chapter, a method to control the compression ratio of JPEG XR is introduced. It presents the property of scaling factor and compression ratio and explains the iterative method to approach the expected compression ratio. The extended JPEG XR encoder is also developed to support this proposed algorithm. Assessment tests are performed on the encoder on the accuracy of the compression ratio and the processing time. The corresponding evaluations are described as well.

The last chapter contains the conclusion and introduces possible future work to expand the proposed algorithms.

2. JPEG XR TECHNOLOGY

2.1 Overview

JPEG XR (formerly HD Photo) is a still-image compression standard originally developed by Microsoft [13]. The steps of the algorithm are similar to JPEG. The source image is partitioned into blocks. Transforms are applied on each block respectively and the compression is achieved by quantization and entropy coding. However, there are many differences and modifications of JPEG XR standard compared with JPEG.

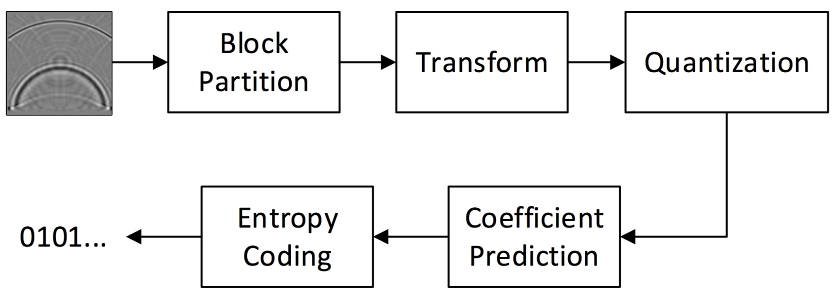


Figure 1: Block diagram of the JPEG XR encoder.

Compared with JPEG, JPEG XR also supports bit depth of 32 bits and floating-point data. Especially, for 32-bit images, JPEG XR only offer lossy compression. Whereas an 8×8 discrete cosine transform is applied in JPEG, JPEG XR introduces 4×4 integer transforms in 16×16 macroblocks and grouped DC coefficients in the second level to achieve multi-resolution hierarchy and lossless to lossy compression. Prediction of coefficients across blocks and macroblocks is applied. The entropy coding is more complex in JPEG XR and it includes adaptive coefficient reordering and adaptive Huffman coding.

2.2 Image hierarchy

As shown in Figure 2, images will be composed of one or more small tiles. Tiles might be padded to an integer number of macroblocks (16×16 pixel block) if they are on the right and bottom margin of images. Each macroblock contains 4×4 blocks and each block contains 4×4 pixels. Note that unlike JPEG with 8×8 pixel blocks and one-time transform in each block, JPEG XR applies a two-stage transform on each 4×4 pixel block and recombined lowpass block in the 16×16 pixel macroblock.

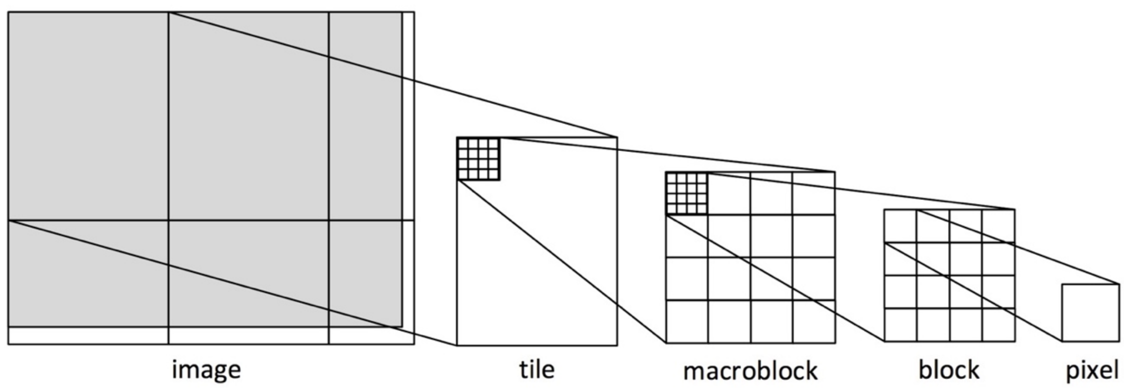


Figure 2: Image hierarchy.

2.3 Transforms

The transform model of JPEG XR is a series of independent transforms which can be presented as a two-stage transform hierarchy and a two-stage overlapped transform which enables fast and robust implementations and lifting-based lossless compression [14]. The encoder is highly flexible to switch the overlapped transform on or off in both stages of hierarchy.

2.3.1 Transform hierarchy

The transform of JPEG XR is an integer-based transform which also supports lossless compression. Each macroblock is transformed in 2 stages. The first stage transform is applied to each block which yields 16 lowpass (LP) coefficients and 240 highpass (HP) coefficients and the second stage transform is applied to the recombined block of the 16 LP coefficients and it provides 1 single direct current (DC) coefficient and 15 LP coefficients as shown in Figure 3. JPEG XR supports to encode the output bitstream in the order of DC, LP and HP coefficients for each tile. Note that this concept of recursive lifting technique is borrowed from the wavelet transform.

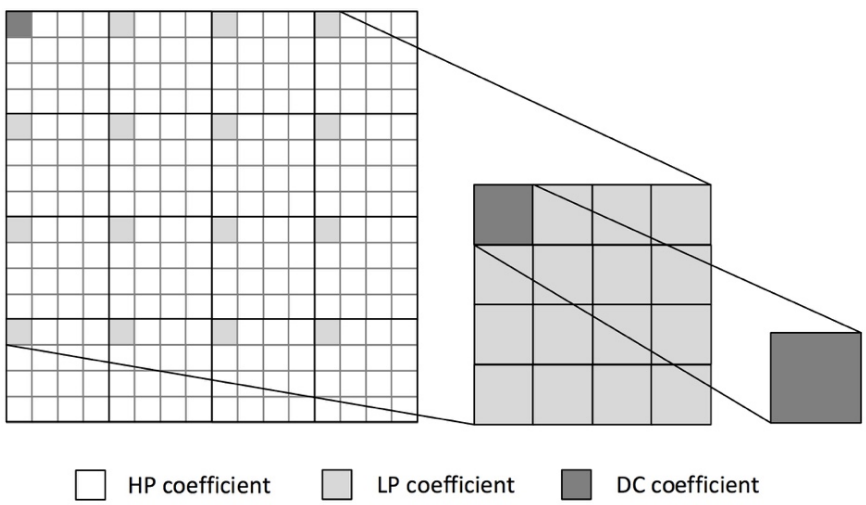


Figure 3: Two-stage transform hierarchy.

2.3.2 Lapped biorthogonal transform

The transform applied on JPEG XR is a two-stage lapped biorthogonal transform (LBT) which contains the photo core transform (PCT) and the photo overlap transform (POT) [15].

- The photo core transform

  The operator of the photo core transform is applied to 4×4 pixel blocks within macroblocks to compress the correlation between pixels of each block in the spatial domain and in the second stage transform it helps to compress the correlation of all lowpass coefficients of macroblocks in the frequency domain. The PCT is almost the same as the discrete cosine transform DCT since the inverse PCT offers a reliable reconstruction of images transformed by DCT.

- The photo overlap transform

  Sharp boundaries of blocks appear in lossy JPEG images because block DCT is unable to preserve correlation across block edges. Therefore, the photo overlap transform is designed to hold the correlation between blocks. As shown in Figure 4, the operator of POT is translated two pixels backward in both dimensions among blocks.

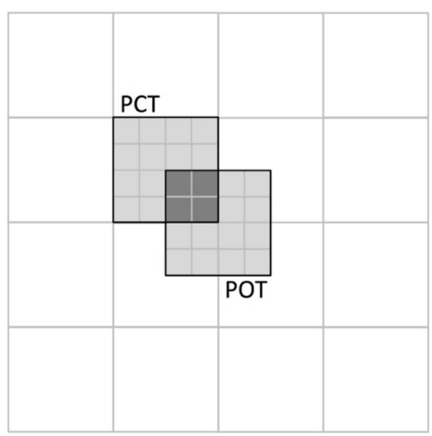


Figure 4: Two-stage lapped biorthogonal transform.

Since the PCT and POT are independent, the JPEG XR standard supports 3 modes of transform from low complexity to high complexity [16]:

- No POT is used in both two stages.
- POT is only used in the first stage (spatial domain).
- POT is used in both two stages (spatial and frequency domains).

2.4 Encoding

2.4.1 Quantization

Quantization in JPEG XR is highly flexible since quantization parameters can be different in tiles, macroblocks and frequency bands. The standard offers quantization parameters from 1 to 255 which are from lossless to extremely lossy. The mapping of quantization parameter and scaling factor is based on [16]:

$$SF = \begin{cases} QP & QP \in [1,16] \\ 2^{\lfloor QP/16 \rfloor - 1}(QP \; mod \; 16 + 16) & QP \in [17,255] \end{cases}$$

The scaled coefficient is obtained by dividing the original coefficient by the corresponding scaling factor of the selected quantization parameter and then rounded to an integer.

2.4.2 Prediction

The JPEG XR standard supports 3 different predictions with respect to DC, LP and HP coefficients [13].

The DC coefficients can be predicted from the left neighbor macroblock, the top neighbor macroblock, both left and top neighbor macroblocks inside the tile or not be predicted. The prediction is determined by computing the differences of DC coefficients of neighboring macroblocks and no prediction is applied if the difference is large.

If DC prediction is applied, the top or left LP edges of macroblocks can be predicted by the neighbor macroblocks which are involved in DC prediction. Therefore, LP prediction presents three options: the first column predicted from its left neighbor macroblock, the first row predicted from its top neighbor macroblock or no prediction applied.

The top or left HP edges of blocks can be predicted by the neighbor blocks inside macroblocks which are determined by the LP coefficients. HP prediction also presents three options: the first column predicted from its left neighbor block, the first row predicted from its top neighbor block or no prediction applied.

2.4.3 Entropy coding

In JPEG XR, the adaptive coefficient normalization is introduced as the entropy coding method [13]. The importance of bits in a coefficient is different. Variable length coding (VLC) is applied to compress the more significant bits with low entropy and fixed length coding (FLC) is applied to remaining bits with high entropy. The FLC data are arranged as Flex Bits band after HP band.

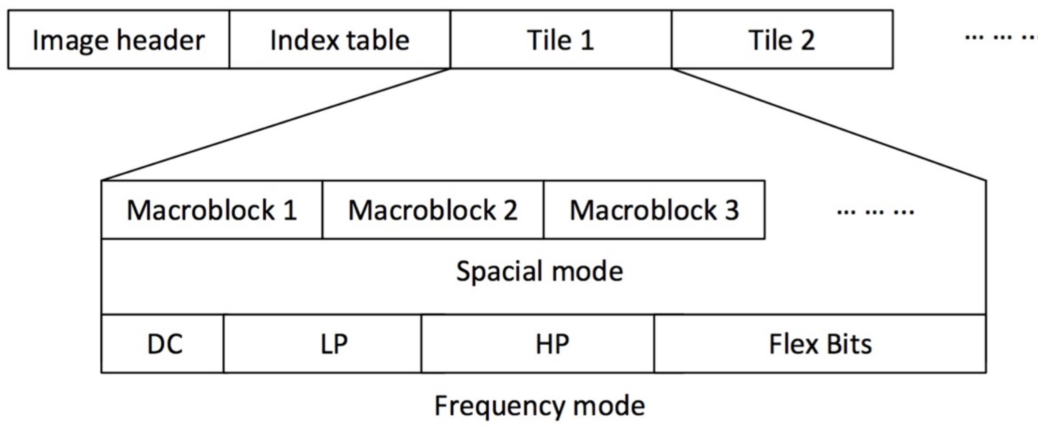


Figure 5: JPEG XR bitstream [16].

As shown in Figure 5, the bitstream can be merged in spatial mode or frequency mode. In the spatial mode, the data are arranged in the order of macroblocks. And in the frequency mode, the data is arranged in the order of DC, LP, HP and Flex Bits bands.

3. SNR CONTROL

In this chapter, we aim at efficiently encoding a 32-bit precision floating-point image into a JPEG XR file with a given SNR target. According to a general method to analyze the relationship between bit rate and SNR proposed by He et al. [8], we proposed an SNR control method based on the nature of the uniform quantizer and quantity of blank macroblocks.

Although the high dynamic range, including 32-bit float, is a built-in feature of JPEG XR standard [17], we select one scaling factor to achieve converting the input floating-point image into a 32-bit integer image and quantizing the integer image at one time for higher efficiency of JPEG XR compression, and the decoded integer image is scaled back to the final floating-point image (with the same scaling factor used at the encoder).

3.1 QP-SNR mapping

Quality control in JPEG XR is done by setting the quantization parameter (QP) between 1 and 255 (with QP=1 indicating lossless mode) [13]. The JPEG XR standard supports flexible quantization in different macroblocks and different bands. Each QP is mapped to a larger-scale scaling factor (SF) and the scaling factor is used as the quantization stepsize in its domain given by

$$m = \left\lfloor \frac{n + \Delta}{q} \right\rfloor$$

$$\Delta = \begin{cases} q/2 & DC\ band \\ (3q+1)/8 & LP\ and\ HP\ bands \end{cases}$$

where the n is the original coefficient and m is the quantized coefficient. However, according to the work from Richter et al. [11] that uniform quantization is always optimal or close to the optimality among adaptive quantization methods, only one single QP is chosen for the image in this work.

3.1.1 Uniform quantization

JPEG XR quantization with a single QP can be regarded as a kind of uniform quantizer that maps continuous values into discrete values with a constant interval qp. Hence, the distortion error caused by a uniform quantizer in each pixel is given by

$$\epsilon(n) = f(n) - INT\left(\frac{f(n)}{qp}\right)qp$$

For instance, Figure 6(a) shows a general curve in the range $[0,14)$. With a uniform quantizer whose constant interval is 1, values of the curve are rounded to the nearest integers as shown in Figure 6(b) and the error caused by a 1-uniform quantizer is in the range $[-0.5,0.5)$ in Figure 6(c) which has an approximately uniform distribution as shown in Figure 6(d). Based on the nature of uniform distribution, the variance of error (mean square error) is approximately $qp^2/12$ [18] and it is given by

$$\epsilon_{MSE} = \lim_{\mathrm{N}\to\infty} \frac{1}{2N}\int_{-N}^{N} |\epsilon(n)|^2 dn = \frac{1}{qp}\int_{-qp/2}^{qp/2} n^2 dn = \frac{qp^2}{12}$$

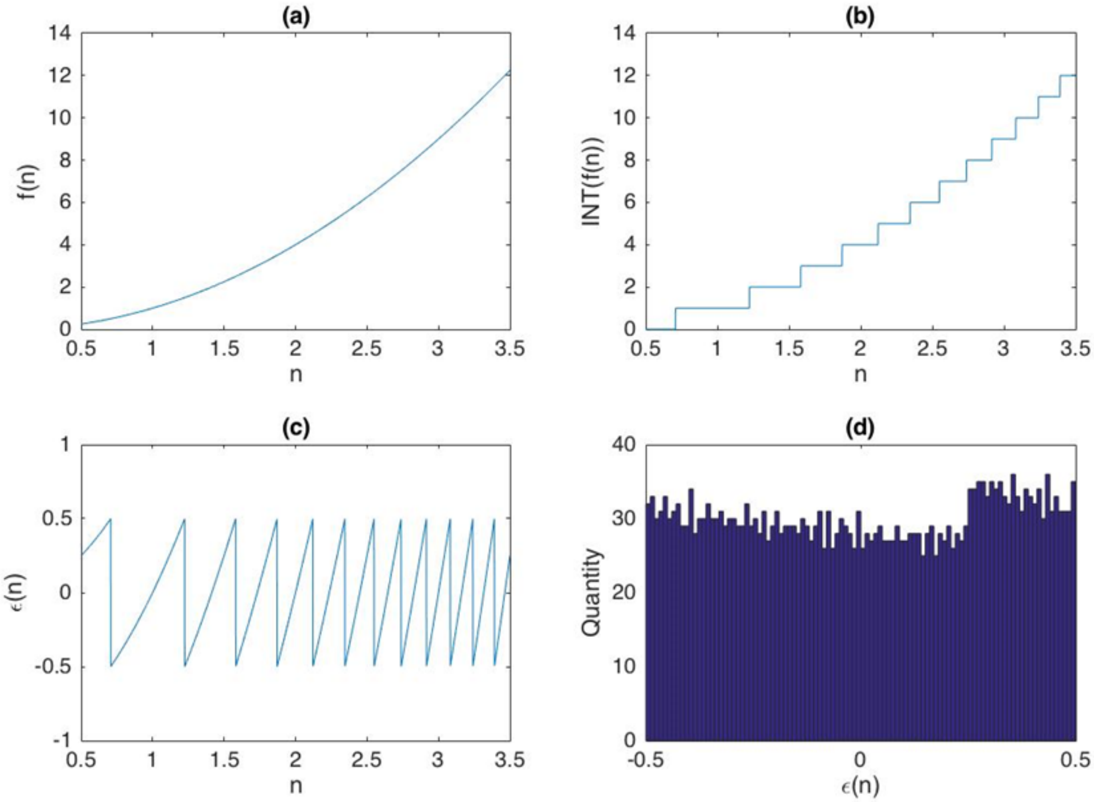


Figure 6: Error distribution of uniform quantization: (a) $f(n) = n^2, n \in [0.5,3.5]$; (b) rounding f(n) to the nearest integer; (c) quantization error; (d) distribution of quantization error.

The SNR is defined as the ratio of the power of signal and noise and it is usually expressed on a logarithmic scale. Because the error distribution has a zero mean, the average power of noise is the same as the variance of error. Hence, the relationship between QP and SNR in an optimal uniform quantizer is given by

$$SNR = 10 \log_{10} \frac{P_{signal}}{P_{noise}} = 10 \log_{10} \frac{\sigma^2/N}{qp^2/12}$$

However, since SNR is highly based on content, the equation is an ideal assumption for the average situation, but not a unifying QP-SNR model for the whole JPEG XR encoding scheme.

3.1.2 Blank macroblock ratio

It is noted that for image and video coding, a blank area among transformed coefficients after quantization has a significant effect on the compression ratio because of the encoding scheme including the prediction technique [8], and JPEG XR is also affected by the percentage of zero coefficients [9]. Although we have the relationship between QP and SNR in a uniform optimal quantizer, the quantization error for blank macroblocks as shown in Figure 7 is not uniformly distributed because of macroblock-based transforming and encoding, i.e., $qp^2/12$ is not a good estimate of the quantization error for these blocks.

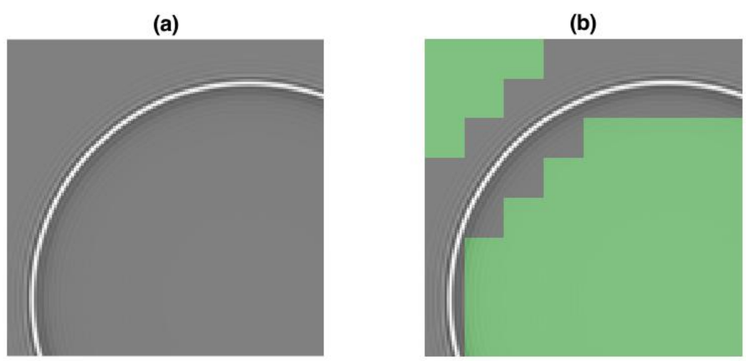


Figure 7: Blank macroblock: (a) raw image patch; (b) blank macroblocks (green) in the patch.

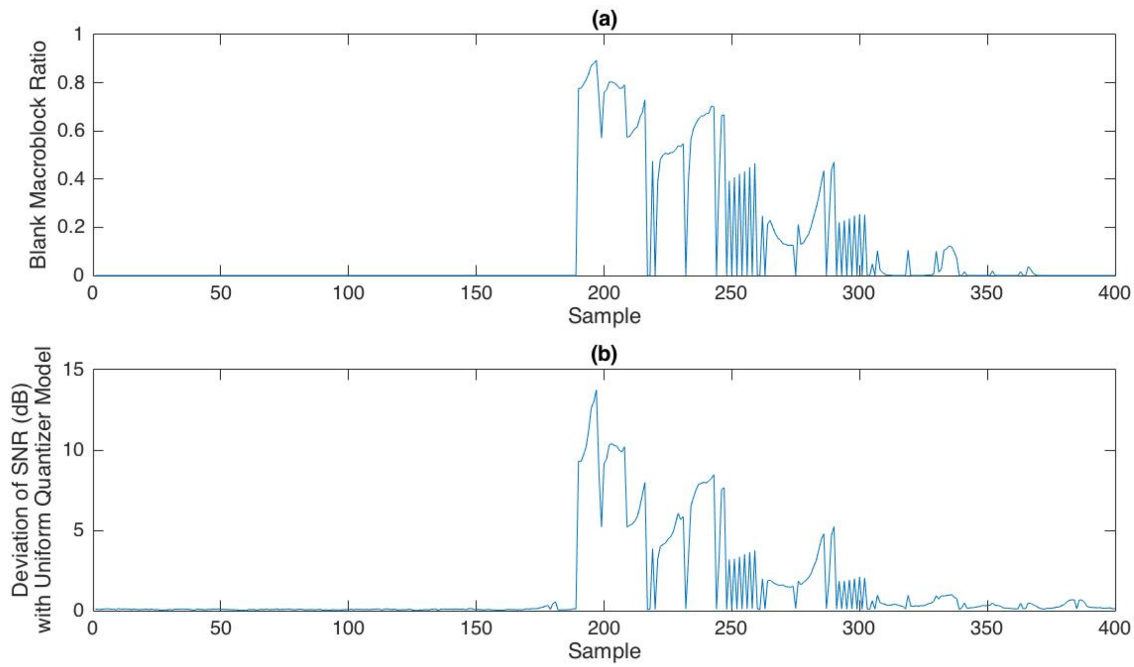


Figure 8: The correlation between blank macroblock ratio and the deviation of SNR: (a) blank macroblock ratio of original sample image; (b) deviation of SNR with target 80dB in ideal uniform quantizer model.

In this work, blank macroblocks are a special case of macroblocks that contain only zero coefficients after quantization as shown in Figure 7. We found the percentage of blank macroblocks after quantization has a high correlation with the deviation of the ideal uniform quantizer model, especially at low quality as shown in Figure 8. Thus, special care must be taken to remove these blank macroblocks to obtain an accurate estimate of the quantization error (mean square error). To correct the estimation, only pixels in non-blank macroblocks should be considered when calculating the average power of the input image. However, the quantity of blank macroblock could vary due to the quantization parameter as shown in Figure 9.

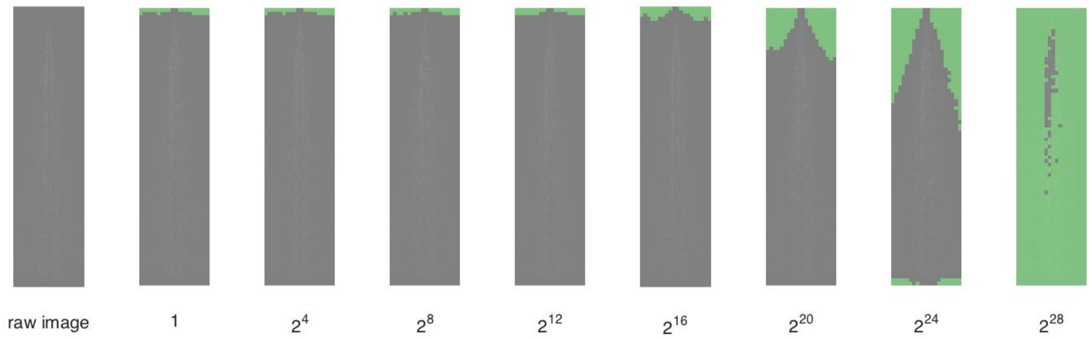


Figure 9: Blank macroblock areas (green) extended with increasing qp.

Note that the number of blank macroblocks is related to qp. Hence, the relationship between QP and SNR in a JPEG XR codec is given by

$$SNR = 10\log_{10}\frac{P_{signal}}{P_{noise}} = 10\log_{10}\frac{\sigma^2/\big(N - N_{bm}(qp)\big)}{qp^2/12}$$

where N is the number of all pixels and N_{bm} is the number of pixels in blank macroblocks which is determined by qp. However, since the number of macroblocks would not change sharply, a precise qp is not required before the blank macroblocks are counted. Therefore, to save the complexity of the algorithm, a polynomial fitting function is applied to estimate qp'.

$$\log_2 qp' = \begin{cases} 0.001174\,SNR^2 - 0.2731\,SNR + 28.07 & SNR < 100 \\ 12.5 & SNR \geq 100 \end{cases}$$

$$SNR = 10\log_{10}\frac{\sigma^2/\big(N - N_{bm}(qp')\big)}{qp^2/12}$$

3.1.3 Factoring qp

After an estimated qp' is obtained from the given SNR target, we factor it into three parts:

$$qp = scaling \cdot q \cdot 2^{shift}$$

a prescaling factor $scaling$ that is used to convert the floating-point input image into a 32-bit integer image, an integer quantization stepsize q as the embedded JPEG XR scaling factor, and a bit shift $shift$ used in JPEG XR to make sure that the transform coefficients can still be fit into 32-bit resolution. In our current implementation, $shift$ is set as 3. We choose the largest internal q in JPEG XR's quantization table that is less than $qp/2^{shift}$. And $scaling$ is automatically set as the remaining factor left in qp. The quantization scheme is shown in Figure 10.

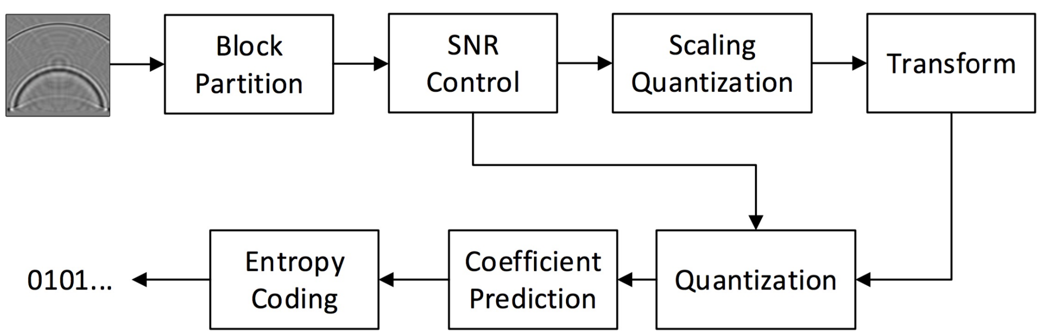


Figure 10: Block diagram of SNR control.

One can potentially set $qp = 1$, i.e., with lossless compression in JPEG XR, while letting $scaling = qp/2^{shift}$. However, our experiments indicate that it is better to utilize JPEG XR standard for quantization since it achieves a better rate-distortion tradeoff than the prescaling step which does not involve entropy coding.

3.2 SNR control algorithm

The detailed algorithm for SNR control is given below.

Input: floating image I with N pixels, SNR target snr

Output: 32-bit integer image M, integer q

1. *Find the maximum value ma and minimum value mi of I.*
2. *Calculate the lower bound of quantization parameters quantStep*

$$quantStep = \frac{max(ma, mi)}{2^{31} - 1}$$

3. *Calculate the power* σ^2 *of image I.*
4. *Estimate a rough quantization parameter qp' according to SNR target snr.*

$$qp' = \begin{cases} 2^{0.001174snr^2 - 0.2731snr + 28.07} & snr < 100 \\ 2^{12.5} & snr \geq 100 \end{cases}$$

5. *Count the number of all pixels* N_{bm} *in blank macroblocks in the image I quantized by qp'.*
6. *Calculate expected quantization parameter qp.*

$$qp = \sqrt{\frac{12\sigma^2}{10^{\frac{snr}{10}} 2^{shift}(N - N_{bm})}}$$

7. *If target SNR snr is higher than 100 and qp is less than quantStep, go to a); If target SNR snr is higher than 100 and qp is larger than quantStep, go to b); If target SNR snr is lower than 100, go to c).*
   a) *Show a message of the highest SNR* snr_{max}*. Then go to 8.*

$$snr_{max} = 10\, log \frac{\sigma^2/(N - N_{bm})}{2^{shift} quantStep^2/12}$$

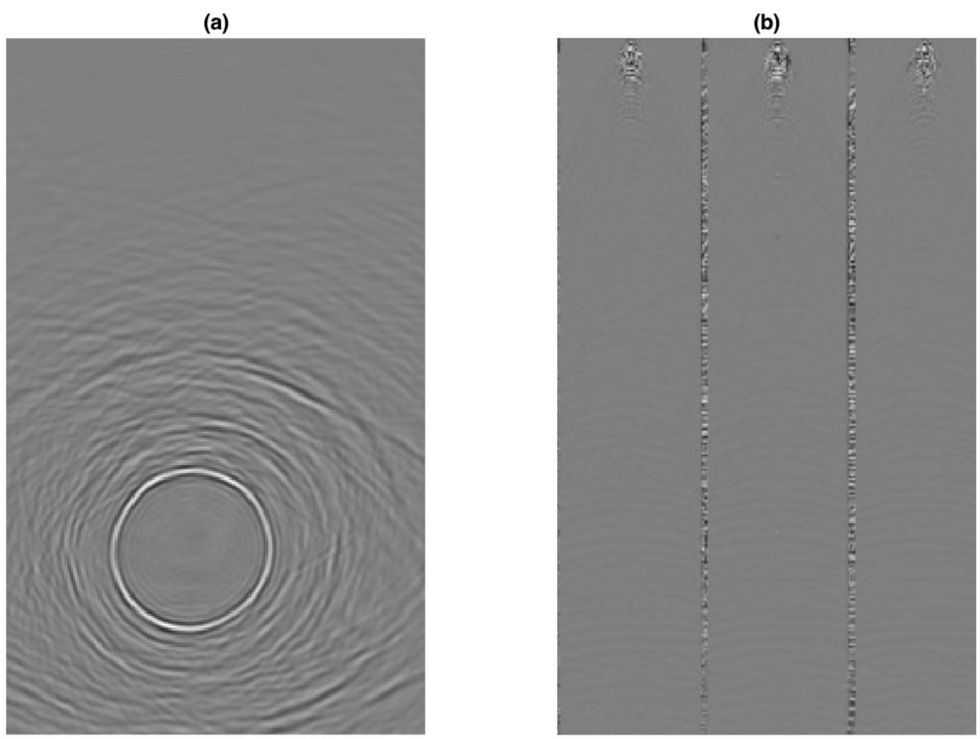
   b) *Assign integer q as 1. Then go to 8.*

c) Assign integer q as the largest integer in JPEG XR's quantization table that is less than quantStep, and then divide quantStep by q. Then go to 8.

8. *Scale image I with quantStep from floating-point numbers to 32-bit integers.*

3.3 SNR control results

The codec is tested on a large seismic dataset including over 20 large scale floating-point trace images and over 1000 slices of 3D wavefield image as shown in Figure 11.



.

Figure 11: Sample images: (a) a slice image of a 3D wavefield data; (b) a trace image.

3.3.1 SNR results

Figure 12 shows the SNR results with tens and general SNR targets in wavefield and trace test sets. Our method gives the high precision of SNR control with high SNR targets and the variance increases a little with low SNR targets. This method also supports precise SNR control with images of any scale including small, large and even 1-pixel-width stripe images.

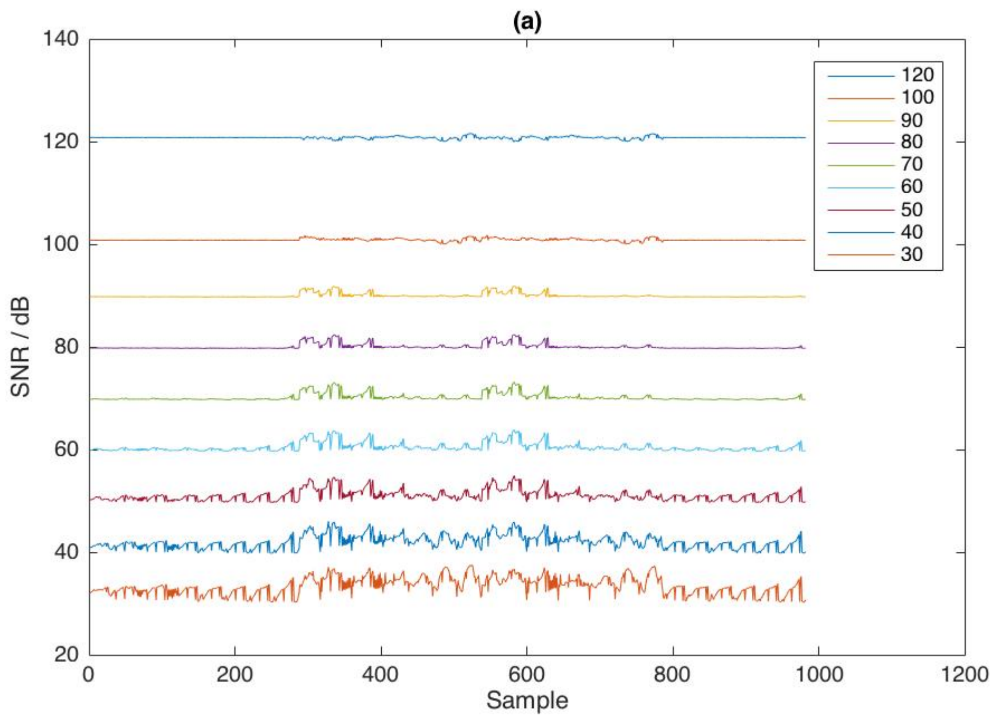
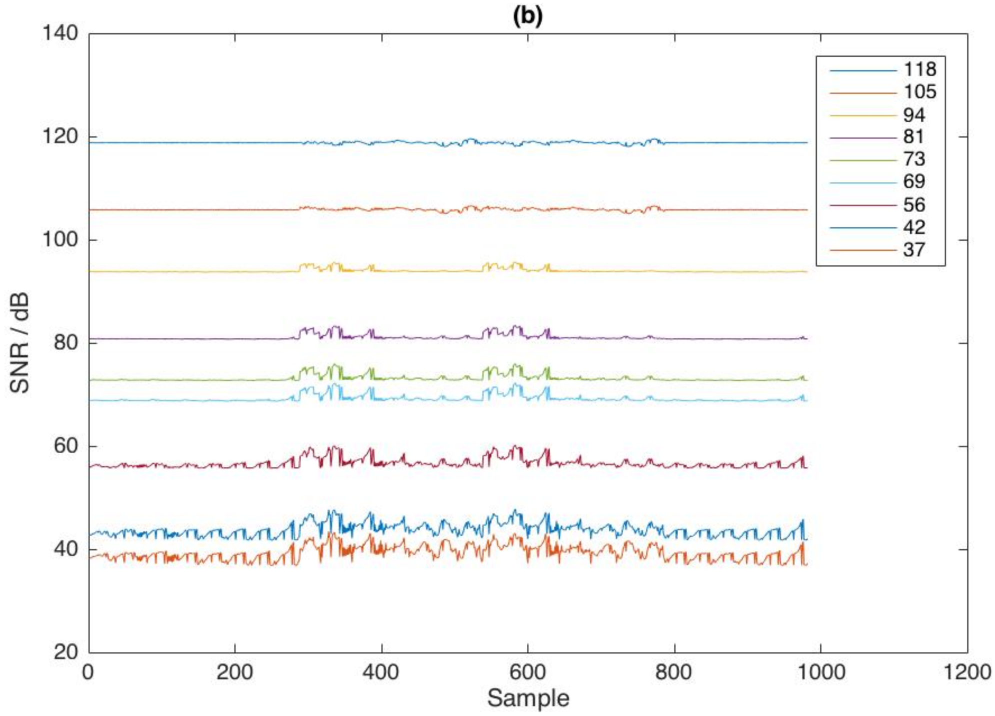


Figure 12: SNR results: (a) Wavefield samples results with tens SNR target; (b) Wavefield samples results with general SNR target; (c) Trace samples results with tens SNR target; (d) Trace samples results with general SNR target.



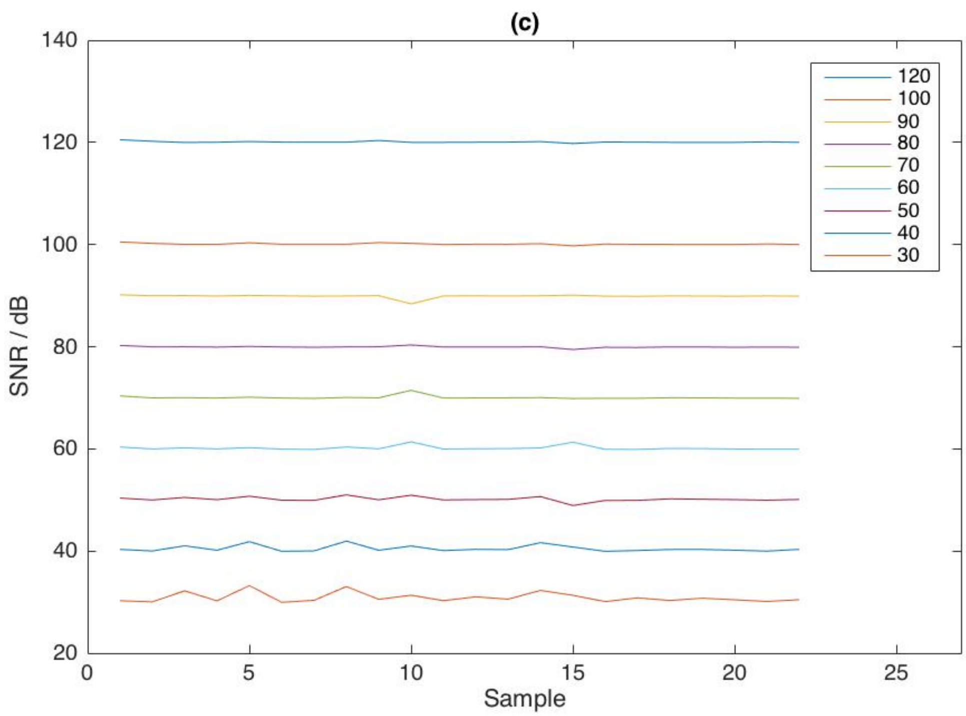


Figure 12: Continued.

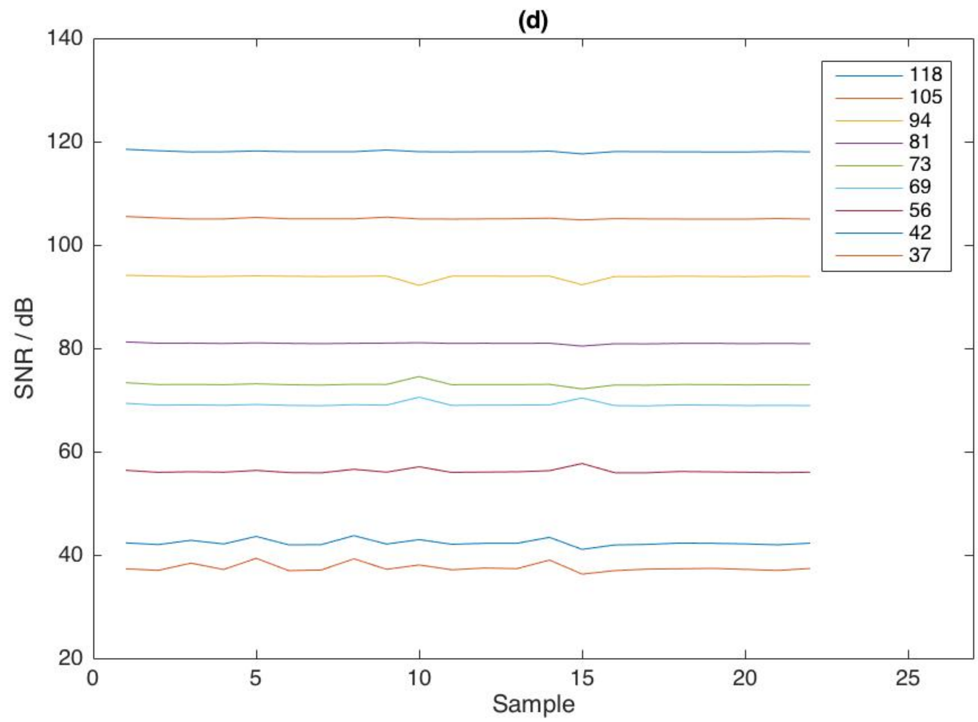


Figure 12: Continued.

3.3.2 Compression results comparison

Figure 13 shows the comparison in terms of SNR vs. compression ratio between original qp-control method and our SNR-control method. Our method gives almost the same performance as the original qp-control method and it greatly extends the scope of quantization.

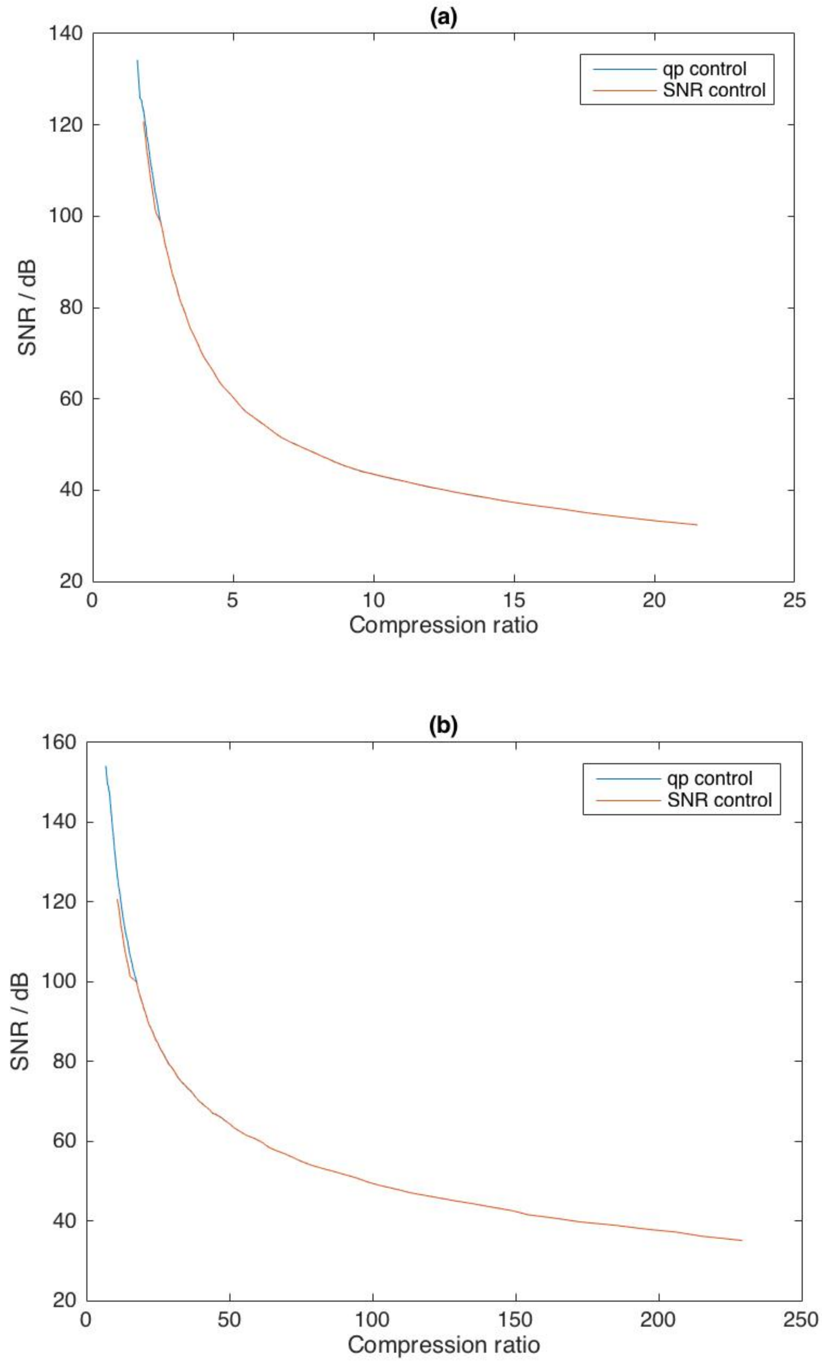
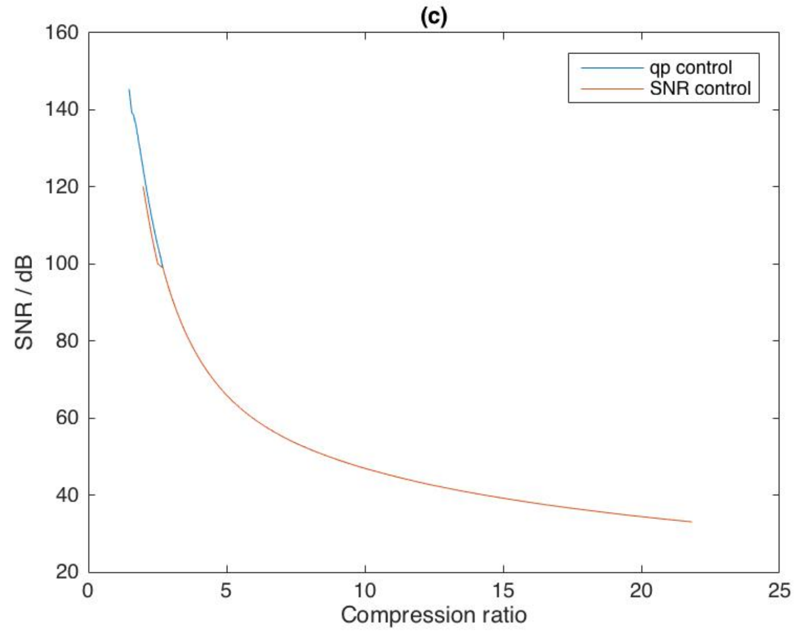


Figure 13: Compression ratio results: (a) a wavefield sample image with low blank macroblock ratio; (b) a wavefield sample image with high blank macroblock ratio; (c) a trace sample image with low blank macroblock ratio; (d) a trace sample image with high blank macroblock ratio.



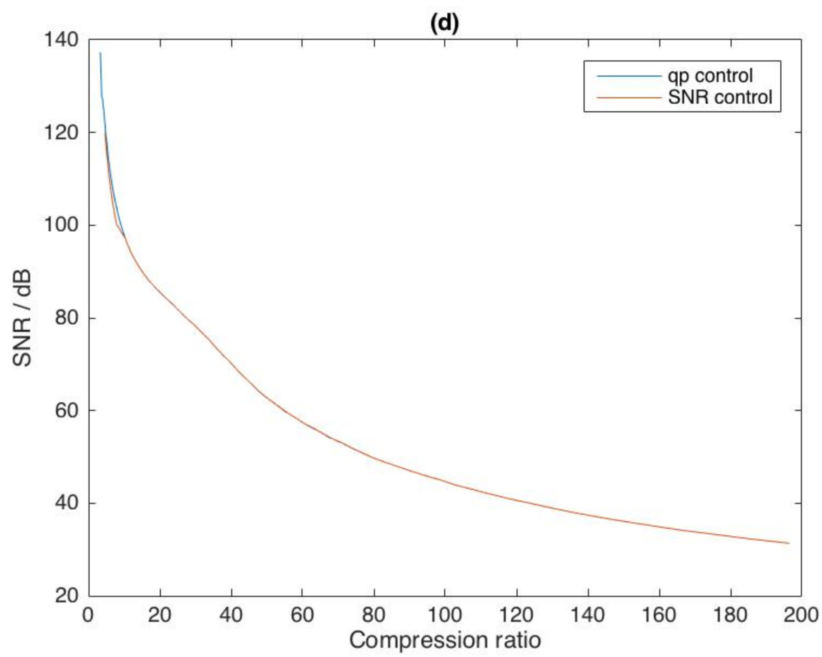


Figure 13: Continued.

3.3.3 Complexity comparison

Figure 14 displays the additional runtime ratio between our SNR-control method and original qp-control method with same SNR targets. It shows that no more than 10% complexity is added to achieve SNR-control.

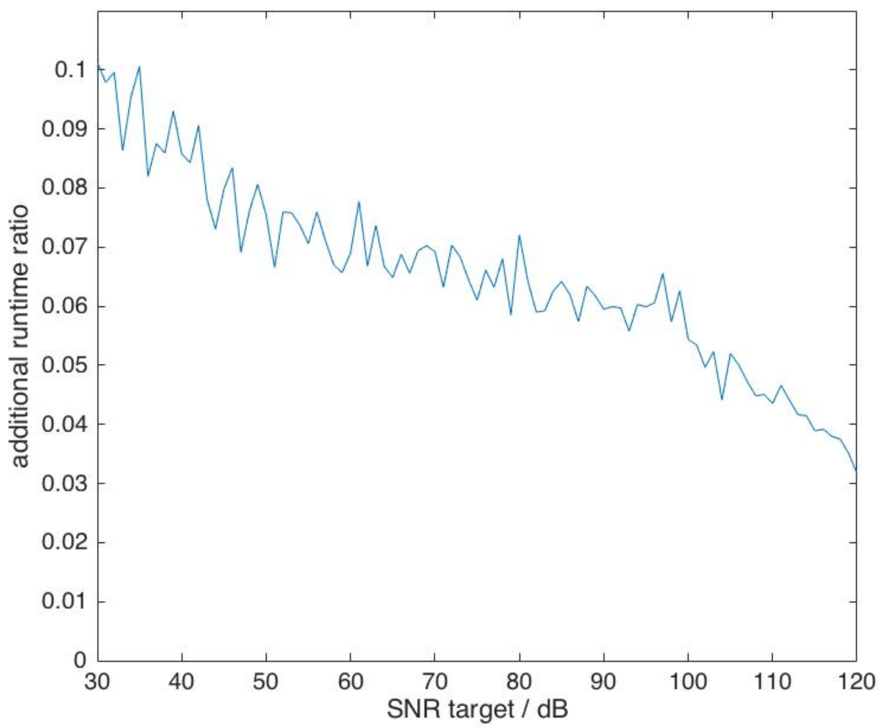


Figure 14: Additional runtime ratio: (new method runtime - original runtime) / original runtime.

4. RATE CONTROL

Quality control in JPEG XR is achieved by setting the quantization parameter (QP) to between 1 and 255 with QP=1 indicating lossless mode. Each QP is mapped to a scaling factor (SF) with the uniform quantization stepsize used by JPEG XR. However, quantization parameter does not lead to stable bit rate which is significant for streaming transmission limited by bandwidth or tradeoff between disk space and image quality in data storage.

We aim at modeling the Compression Ratio-Scale Factor (CR-SF) curve of images during encoding process and finding out the right scaling factor corresponding to our specified target compression ratio. Inspired by a JPEG rate control method [5] based on fast converging functions, an iterative rate control method is proposed to approximate the linear nature of JPEG XR image based on a converging function.

4.1 QP-CR mapping

The quantization parameters which are between 1 and 255 are mapped in scaling factors which are in a much larger range [1,1016808] as shown in Figure 15. It significantly helps to expand the limit of quantization in an exponential scale and store a very large range scaling factor into 8 bits in the parsed syntax element.

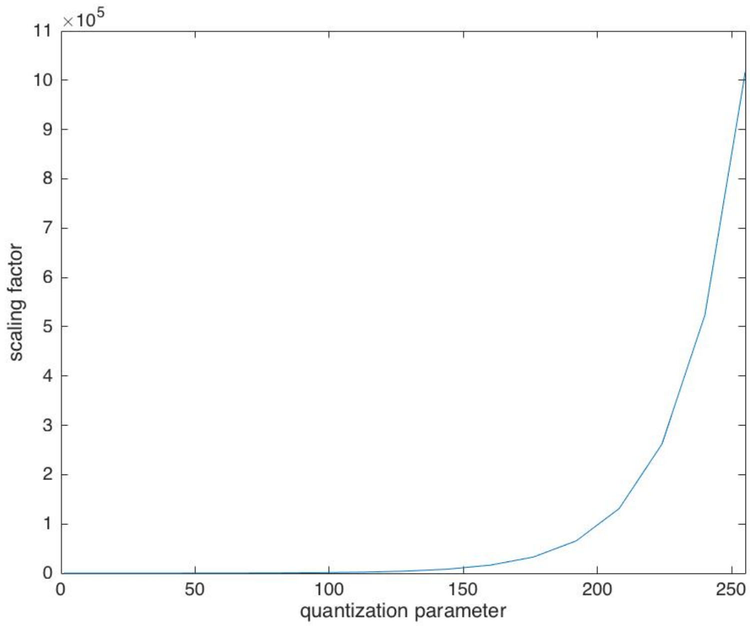


Figure 15: SF-QP curve.

Note that the JPEG XR quantizer has a harmonic scale because SF increases linearly with respect to QP initially but grows exponentially later to allow high compression ratios for high bit-depth images. This leads to fast implementation by bit shifts (together with a few multiplications due to the limited digits of the mantissa).

$$SF = \begin{cases} QP & QP < 16 \\ 2^{\lfloor QP/16 \rfloor}(QP \bmod 16 + 16) & QP \geq 16 \end{cases}$$

4.2 CR-SF modeling

For different JPEG XR images, we have unique CR-SF curves. However, In Figure 16, four images are selected with either trace or wavefield image and either high or low blank macroblock ratio. Like JPEG [5], CR-SF curves from the whole sample set share a strong linear characteristic.

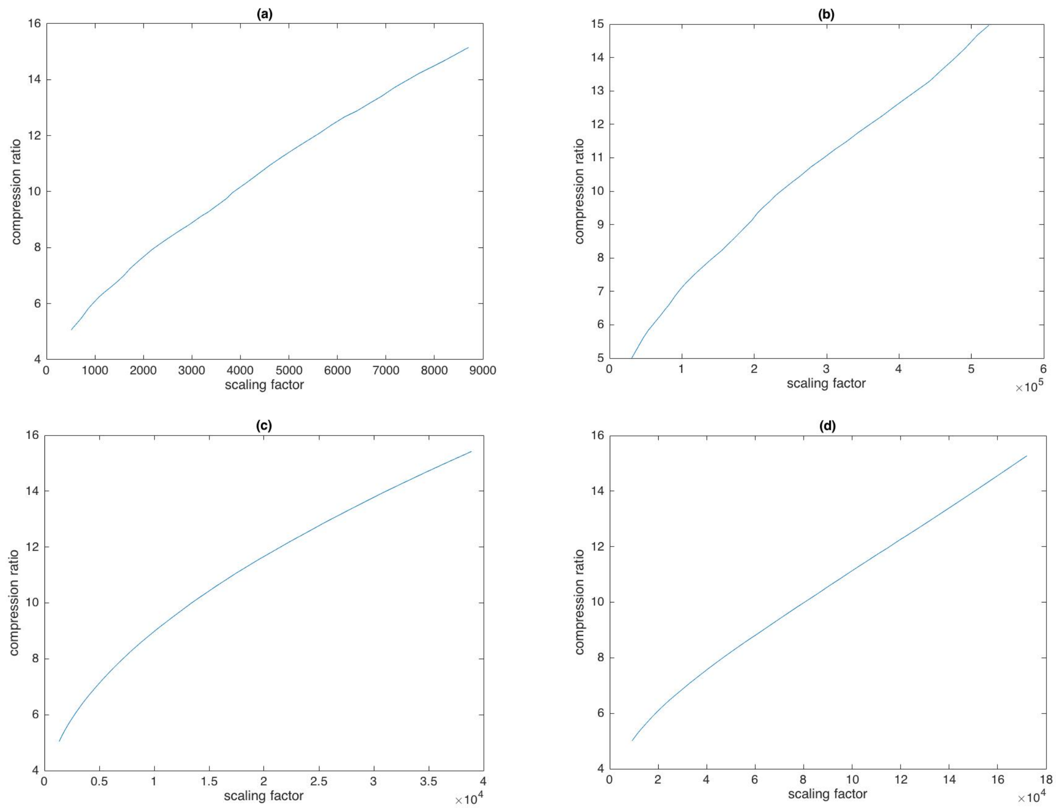


Figure 16: CR-SF curves for four different images.

Therefore, the linear approximation $CR = k \cdot SF + b$ is selected to model the CR-SF curves. The approach to controlling the compression ratio is to find the right slope k and intercept b of the approximate line close to the target CR and to calculate the target SF with $SF = (CR - b)/k$ and then to quantize the input image with the corresponding QP to target SF as shown in Figure 17.

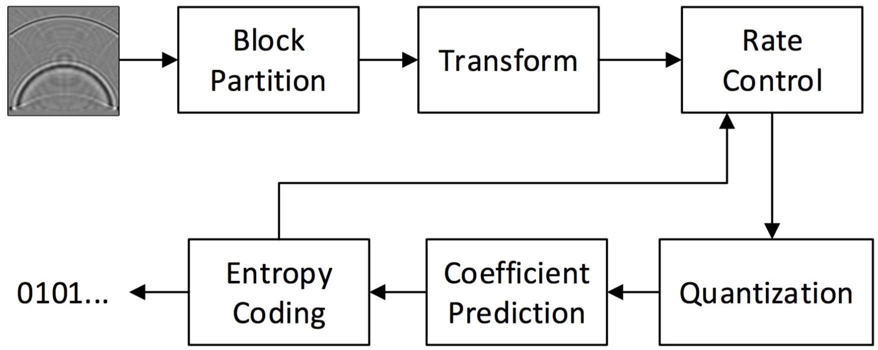


Figure 17: Block diagram of rate control.

In this method, at least two points on the curve are needed to estimate the slope and intercept which means the complexity of repeated encoding would be at least 3 times higher than the original encoding. As shown in Figure 18, the slopes and intercepts of the fitting line are highly concentrated in the sample set.

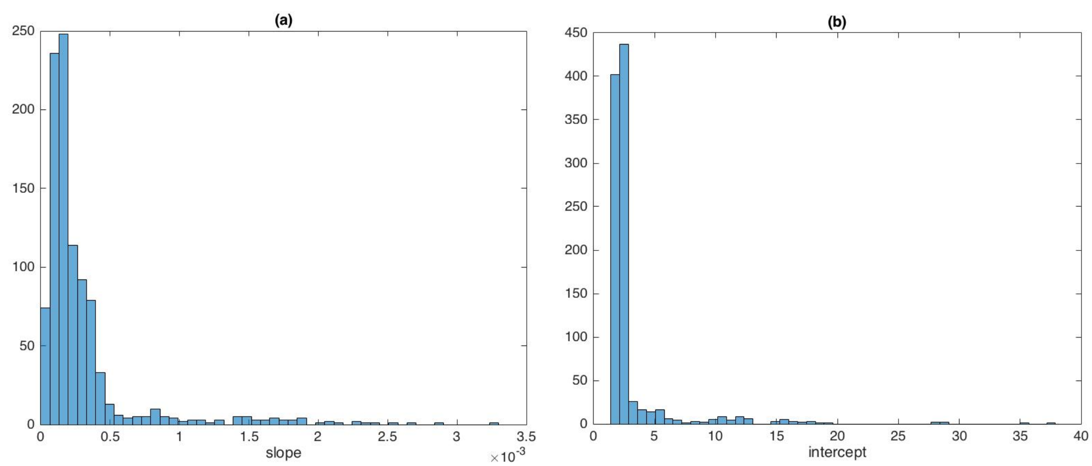


Figure 18: Histograms of slopes and intercepts of fitting lines.

Therefore, to speed up the iterative process, we use the average slope and intercept obtained from the large sample set to initially estimate the target point as shown in Figure 19. The linear model is given by

$$CR = k \cdot SF + b$$
$$k = 2.9783\times10^{-4}$$
$$b = 3.1378$$

where k is the average slope and b is the average intercept. Hence, given a target compression ratio CRT, we estimate the corresponding SF1 with the straight line and we can obtain the first estimated scaling factor $SF1 = (CRT - b)/k$. Compressing the image with SF1, we should obtain the compression ratio CR1 which is on the real CR-SF curve and could be different from CRT.

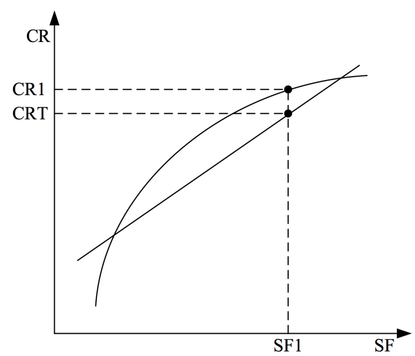


Figure 19: Linear approximation (iteration 1).

If the target is not achieved, the known CR-SF point (SF1, CR1) could help enhance the approximate line. We can give up the average intercept and shift the line across the known point with the average slope as shown in Figure 20. Then, we estimate the scale factor SF2 corresponding to the target compression ratio CRT with the updated straight line.

$$CRT = k\,(SF - SF1) + CR1$$

Hence, the second estimated scaling factor is $SF2 = (CRT - CR1)/k + SF1$ and we should obtain the compression ratio CR2 which is on the real CR-SF curve and might be different from CRT.

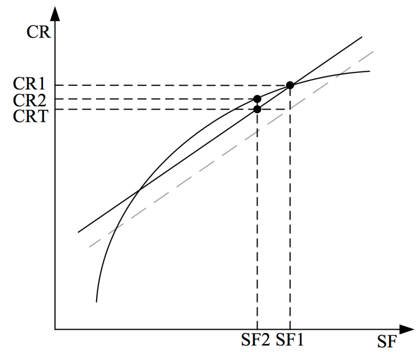


Figure 20: Linear approximation (iteration 2).

If the target is still not achieved, we can modify the estimated line more precisely with the two known CR-SF points (SF1, CR1) and (SF2, CR2) by connecting these two points as shown in Figure 21.

$$CRT = \frac{(CR2 - CR1)(SF - SF1)}{SF2 - SF1} + CR1$$

Hence, given the target compression ratio CRT, we estimate the corresponding SFT with the latest straight line and we can obtain the estimated scaling factor $SFT = (CRT - CR1)(SF2 - SF1)/(CR2 - CR1) + SF1$ and we should obtain the compression ratio CRT since most scaling factor corresponding to the target compression ratio (with an error margin of 5%) is obtained in three iterations. However, for some instances away from average instances, continuous iterations for linear approximation is needed to use the straight line of the latest updated CR-SF points.

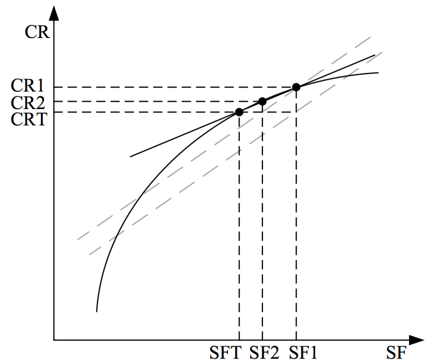


Figure 21: Linear approximation (iteration 3).

4.3 Rate control algorithm

The detailed algorithm for rate control is given below.

Input: transformed image I, CR target CRT, tolerance t, default slope k and intercept b

Output: integer QP q

1. *Calculate SF1 corresponding to CRT in the default fitting model.*

$$SF1 = (CRT - b)/k$$

2. *Search out the QP1 mapping to the SF closest to SF1. Encode image I with QP1, count the encoded image size and calculate the ratio CR1. If CR1 is greater than CRT and QP1=1, stop here with q=1. If the error between CRT and CR1 is less than t, stop here with q=QP1. If CR1 is less than CRT and QP1=255, stop here with q=255.*

3. *Estimate a CR-SF curve with point (QP1, CR1) and the average slope k in the default fitting model and calculate the corresponding SF2.*

$$SF2 = \frac{CRT - CR1}{k} + QP1$$

4. *Search out the QP2 mapping to the SF closest to SF2, encode image I with QP2, count the encoded image size and calculate the ratio CR2. If CR2 is greater than CRT and QP2=1, stop here with q=1. If the error between CRT and CR2 is less than t, stop here with q=QP1. If CR2 is less than CRT and QP2=255, stop here with q=255.*
5. *Estimate a CR-SF curve with 2 known SF-CR points (SFa, CRa) and (SFb, CRb) closest to CRT from all know SF-CR points. calculate the corresponding SFnext.*

$$SFnext = SFa - \frac{(SFa - SFb)(CRa - CRT)}{CRa - CRb}$$

6. *Search out the QPnext mapping to the SF closest to SFnext, encode image I with QPnext, count the encoded image size and calculate the ratio CRnext. If CRnext is greater than CRT and QPnext=1, stop here with q=1. If the error between CRT and CRnext is less than t, stop here with q=QPnext. If CRnext is less than CRT and QPnext=255, stop here with q=255 Otherwise, go to 5.*

4.4 Rate control results

4.4.1 Compression ratio results

Figure 22 shows the rate control results with compression ratios in the integer compression ratio target in the range [5,15] with over 800 wavefield test sets. The tolerance of error is set to be less than 5% of the target compression ratio and only a compression ratio higher than the target is allowed. Not like SNR control, this method does not design an outer quantizer to extend the upper and lower limits of built-in quantization. Therefore, only those images whose range of compression ratio wider than [5,15] are included in the test set.

The results show high precision for low compression ratio targets. However, since there is no additional quantizer, the intervals between scaling factors are larger when the compression ratio target raises and corresponding quantization parameter increases. So for high compression ratio targets, the method actually gives closest results which is higher than the compression ratio target.

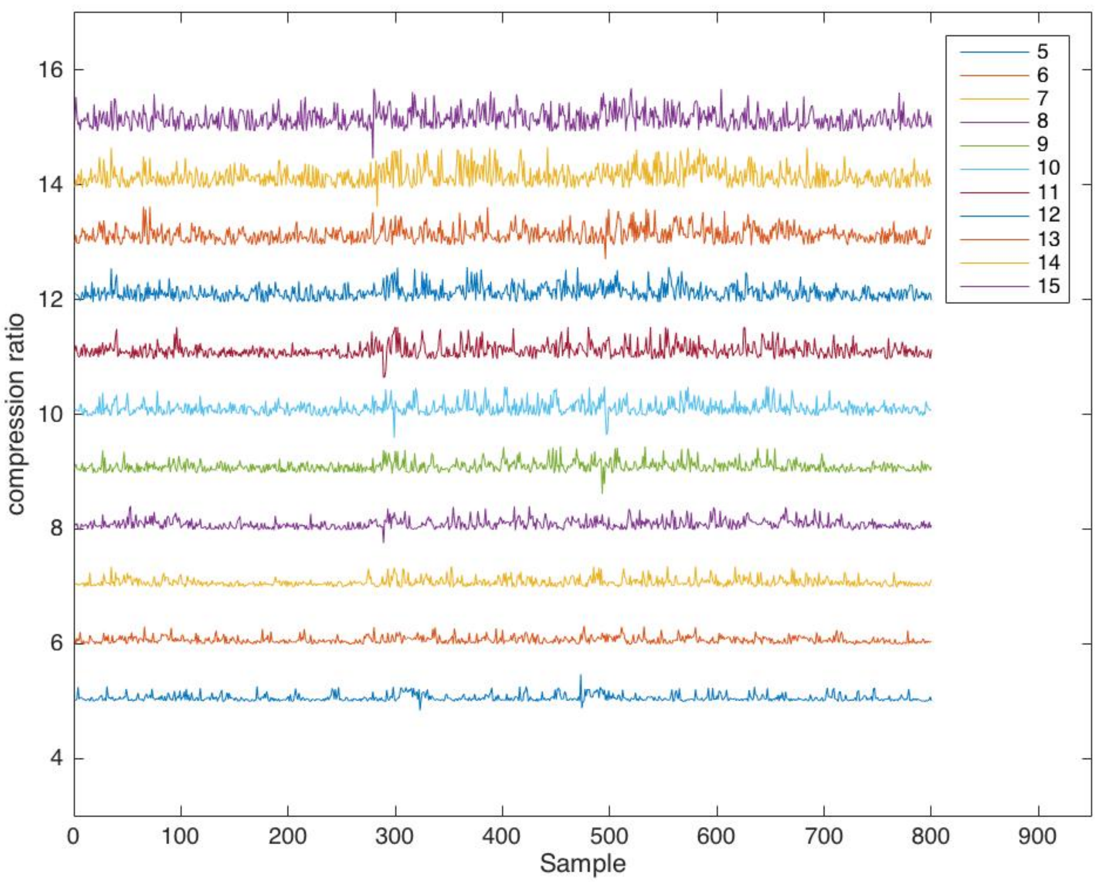


Figure 22: Compression ratio results.

4.4.2 Complexity comparison

For the range of QP in [1,255], our method requires on average three times the complexity of the baseline JPEG XR codec. This is much faster than the benchmark binary search method, which on average needs eight iterations. Figure 23 displays the additional run-time ratio between our rate-control method and the original JPEG XR algorithm (assuming the right QP is known). It is seen that due to the iterative process of our rate control algorithm, the complexity increase is on average 160% over that of the original JPEG XR algorithm (without rate control).

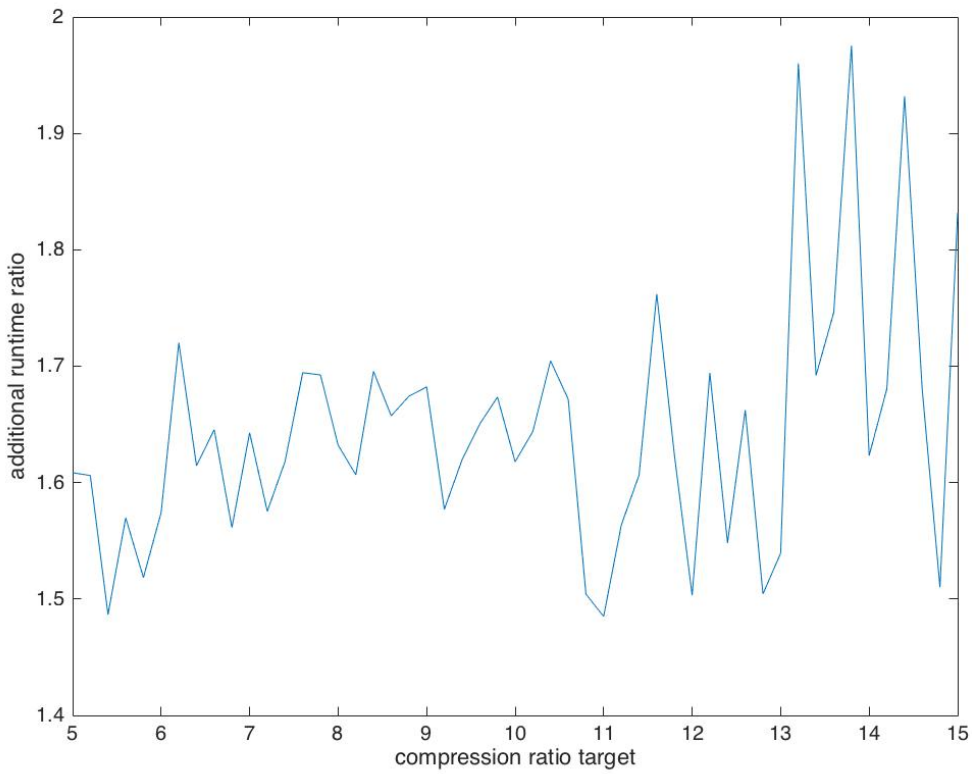


Figure 23: Additional runtime ratio: (new method runtime - original runtime) / original runtime.

5. CONCLUSION AND FUTURE WORK

5.1 Conclusion

In this thesis, two quality control algorithms in JPEG XR are proposed. The first algorithm controls the SNR of encoded data quickly and accurately by evaluating appropriate quantization parameters during preprocessing. In the second algorithm, the suitable quantization parameter with respect to given bit rate is searched by fast iterative convergence.

The first part of this thesis summarizes our work on SNR control based on the fast Microsoft implementation of JPEG XR. We added SNR control with less than 10% extra complexity. Extensive tests on over 1000 trace images and 22 wavefield images indicate that our SNR control algorithm's output SNR (after decoding) matches the input target very well at the relatively high SNR (e.g., < 40 dB). As the target SNR gets lower, our estimate of the MSE becomes less accurate, leading to a slight mismatch between the target and output SNRs, which is on average within 3dB.

The second part describes a fast implementation of rate control in JPEG XR, using a linear approximation of the CR-SF curve. For a given compression ratio, we limit our average number of iterations to three to find the approximate scaling factor with an error margin of 5%. The method is limited when the target is out of the range of quantization parameters [1,255].

5.2 Future work

The proposed algorithms are considered to be a starting point to control the quality of seismic data. As an encoder for seismic data, it is expected to support high dynamic range floating coding, large scale picture and efficient rate control for streaming/3D data.

In terms of SNR control, the upper and lower limitations of SNR range can improve. Firstly, the upper limitation of SNR is causes by the default bit shift before the transform. As we mentioned briefly in SNR control, although the standard claims that 32-bit floating is supported in the raw data, a default bit shift is necessarily applied before transform in case the DC coefficient and even other coefficients after the transform are longer than 32 bits. Therefore, the lossless mode is not accessible for the 32-bit data type and the upper bound is around 180dB which is determined by the default bit shift. In order to support 32-bit lossless encoding, an additional preprocessing or novel encoding is needed. On the other hand, a lower bound of SNR control exists since the SNR variance of results becomes larger when the SNR target is lower than 40dB. Hence, a new method with a deeper understanding of JPEG XR property in low SNR is required to make the control more accurate.

With respect to rate control, the proposed algorithm is not as fast as the standard encoder because of the iterative approach involved. As we mentioned in the related work, Chan et al. [9] proposed a reliable algorithm to estimate rate-SNR curves for JPEG XR images. Moreover, in this thesis, a fast and precise algorithm to control SNR for JPEG XR is proposed. Therefore, there is a new possible approach to controlling bit rate which is to estimate the corresponding SNR according to the given rate target and then applying the SNR control method to obtain all appropriate parameters for the quantization process.

REFERENCES

[1] Dufaux, Frédéric, Gary Sullivan, and Touradj Ebrahimi. "The JPEG XR image coding standard." IEEE Signal Processing Magazine 26, no. MMSPL-ARTICLE-2009-004 (2009): 195-199.

[2] Tran, Trac D., Lijie Liu, and Pankaj Topiwala. "Performance comparison of leading image codecs: H. 264/AVC Intra, JPEG2000, and Microsoft HD Photo." In Optical Engineering+ Applications, pp. 66960B-66960B. International Society for Optics and Photonics, 2007.

[3] Richter, Thomas. "Visual quality improvement techniques of HDPhoto/JPEG-XR." In Image Processing, 2008. ICIP 2008. 15th IEEE International Conference on, pp. 2888-2891. IEEE, 2008.

[4] De Simone, Francesca, Lutz Goldmann, Vittorio Baroncini, and Touradj Ebrahimi. "Subjective evaluation of JPEG XR image compression." In SPIE Optical Engineering+ Applications, pp. 74430L-74430L. International Society for Optics and Photonics, 2009.

[5] Bracamonte, Javier, Michael Ansorge, and Fausto Pellandini. "Bit-Rate Control for the JPEG Algorithm." In Intelligent Integrated Media Communication Techniques, pp. 263-301. Springer US, 2003.

[6] Bruna, Arcangelo, Stewart Smith, Filippo Vella, and Filippo Naccari. "JPEG rate control algorithm for multimedia." In Consumer Electronics, 2004 IEEE International Symposium on, pp. 114-117. IEEE, 2004.

[7] Yeung, Yick Ming, and Oscar C. Au. "Efficient rate control for JPEG2000 image coding." Circuits and Systems for Video Technology, IEEE Transactions on 15, no. 3 (2005): 335-344.

[8] He, Zhihai, and Sanjit K. Mitra. "A unified rate-distortion analysis framework for transform coding." Circuits and systems for video technology, IEEE Transactions on 11, no. 12 (2001): 1221-1236.

[9] Chan, Duncan, Jie Liang, and Chengjie Tu. "ρ-domain rate control for JPEG XR." In 2010 Conference Record of the Forty Fourth Asilomar Conference on Signals, Systems and Computers. 2010.

[10] Testoni, Vanessa, Max HM Costa, Darko Kirovski, and Henrique S. Malvar. "On the adaptive coefficient scanning of JPEG XR/HD photo." In Data Compression Conference (DCC), 2010, pp. 69-78. IEEE, 2010.

[11] Richter, Thomas. "Spatial constant quantization in JPEG XR is nearly optimal." In Data Compression Conference (DCC), 2010, pp. 79-88. IEEE, 2010.

[12] Averbuch, Amir Z., F. Meyer, Jan-Olov Strömberg, R. Coifman, and Anthony Vassiliou. "Low bit-rate efficient compression for seismic data." Image Processing, IEEE Transactions on 10, no. 12 (2001): 1801-1814.

[13] ITU-T Rec. T.832 | ISO/IEC 29199-2, Information technology – JPEG-XR image coding system – Image coding specification, http://www.itu.int/rec/T-REC-T.832.

[14] Liang, Jie, Chengjie Tu, and Trac D. Tran. "Fast lapped transforms via time-domain pre-and post-processing." Proc. IEEE ICICS'2001 (2001).

[15] Tu, Chengjie, Sridhar Srinivasan, Gary J. Sullivan, Shankar Regunathan, and Henrique S. Malvar. "Low-complexity hierarchical lapped transform for lossy-to-lossless image coding in JPEG XR/HD Photo." In Optical Engineering+ Applications, pp. 70730C-70730C. International Society for Optics and Photonics, 2008.

[16] Srinivasan, Sridhar, Chengjie Tu, Shankar L. Regunathan, and Gary J. Sullivan. "HD Photo: a new image coding technology for digital photography." In Optical Engineering+ Applications, pp. 66960A-66960A. International Society for Optics and Photonics, 2007.

[17] Srinivasan, Sridhar, Zhi Zhou, Gary J. Sullivan, Robert Rossi, Shankar Regunathan, Chengjie Tu, and Anirban Roy. "Coding of high dynamic range images in JPEG XR/HD Photo." In Optical Engineering+ Applications, pp. 707315-707315. International Society for Optics and Photonics, 2008.

[18] Jain, Anil K. Fundamentals of digital image processing. Prentice-Hall, Inc., 1989.