

# Raw Data Compression in Computed Tomography: Noise Shaping

Yao Xie

(team member: Adam Wang)

Project Final Report for EE372  
Stanford University, Spring 2006-07

July 18, 2007

## Abstract

## 1 Introduction

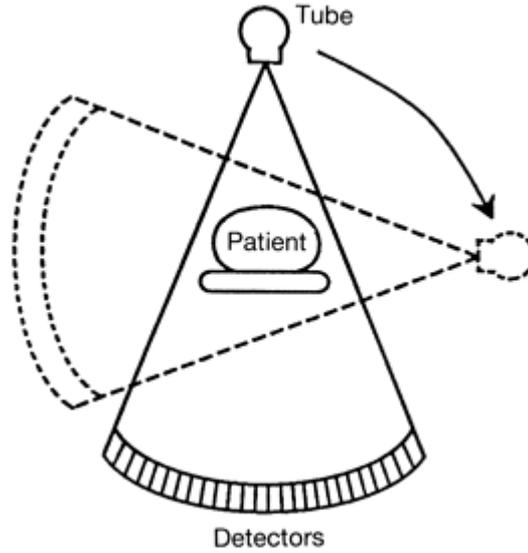
### 1.1 Background

X-ray computed tomography (CT) builds on the physical principles of radiography. It uses multiple views of an external x-ray source and images x-ray attenuation properties  $\mu$  as a function of location within the body.

In a CT scanner, large amounts of raw data are collected in the rotating gantry and must be transferred offline for processing and image reconstruction. Generally this transfer is made through a slip ring that has a limited data transfer rate. Compression can be used to reduce the data rate through the slip ring, but one needs to ensure that the compression does not degrade image quality.

This problem becomes more severe in the new inverse geometry CT (IGCT) system [GRGJ04][SFP05]. Because IGCT has multiple sources, it requires higher data rate than that of a typical cone beam CT system, which is already quite heavy. Consider a typical cone beam geometry CT system whose data transmission rate can be estimated as: 984 views/rotation  $\times$  888 detectors/row  $\times$  64 rows  $\times$  3.33 rotation/s  $\times$  16 bits/detector  $\approx$  400 MB/s.

However, our preliminary literature search has shown that no one has considered compressing data before transmitting it through the slip ring, and the only work that has been done on compressing raw data (sinogram) has used standard image compression techniques such as JPEG [BW01]. These techniques do not take into account the nature of the sinogram and how errors in the sinogram contribute to errors (noise) in the reconstructed image.



**Figure 1:** A fan beam CT system.

Compression can be applied not only to raw data transmission, but also to raw data storage. Currently, only reconstructed images are stored in PACS (picture archiving and communication system). This does not allow the flexibility to reconstruct different slices or do planar reformats since the original raw data is discarded. If we are able to compress the sinogram data with a tolerable amount of loss in the reconstructed image quality, we can store the raw data and then reconstruct desired images in the future.

## 1.2 Raw Data Model

To form images in CT systems, a set of x-ray beams are scanned through the entire field of view (where the object lies) in what is known as fan or, more generally, cone beam geometry. We will consider 2-D scan with fan beam geometry, as shown in Fig. 1<sup>1</sup>. Each x-ray beam (measured by the  $n$ th detector), is modelled as a line positioned at  $r_n$  with orientation  $\theta_i$ . The attenuation values  $\mu(x, y)$  along the path of the x-ray beam are superimposed, resulting in a line-integral of the attenuation. The attenuated x-ray beam intensities are measured using detectors. The expected value of the measured intensity for a particular beam is given by Beer's law:

$$I(r_n, \theta_i) = I_0 \exp \left\{ - \int_{l(r_n, \theta_i)} \mu(x, y) dl \right\},$$

---

<sup>1</sup>[www.medcyclopaedia.com](http://www.medcyclopaedia.com)

where  $I_0$  is the incident intensity of the x-ray beam passing through the object, and  $l(r_n, \theta_i)$  is the line through which the x-ray passing through the object.

This process is repeated for a large number of angles, yielding line attenuation samples of all angles  $\theta_i$ ,  $i = 0, \dots, N_\theta - 1$  and of all distances  $r_n$ ,  $n = 0, \dots, N_r - 1$  from the center of the detector array within the field of view (FOV). Then we have a complete data set of  $\{g_{\theta_i}(r_n)\}$ . Due to the discrete nature of the x-ray photon, each measurement  $g_{\theta_i}(r_n)$  follows a Poisson distribution with mean and variance  $I(r_n, \theta_i)$ .

The sinogram is a two-dimensional representation of the measured signal  $\{g_{\theta_i}(r_n)\}$ . We also refer to the sinogram as the raw data. From the sinogram, the actual attenuation at each voxel of the scanned slice can be reconstructed [DeM01]. Before image reconstruction, the raw data is often normalized by  $I_0$  and followed by a logarithmic operation:

$$-\log\left(\frac{g_{\theta_i}(r_n)}{I_0}\right) = \int_{l(r_n, \theta_i)} \mu(x, y) dl.$$

This logarithmic operation essentially corresponds to a logarithmic compander before quantization, which is widely used for A/D conversion. The logarithmic compander effectively reduces the data dynamic range. In addition to the logarithmic compander, we can also consider other companders such a square root compander (as studied in Adam’s project report). In this project we only use the logarithmic compander.

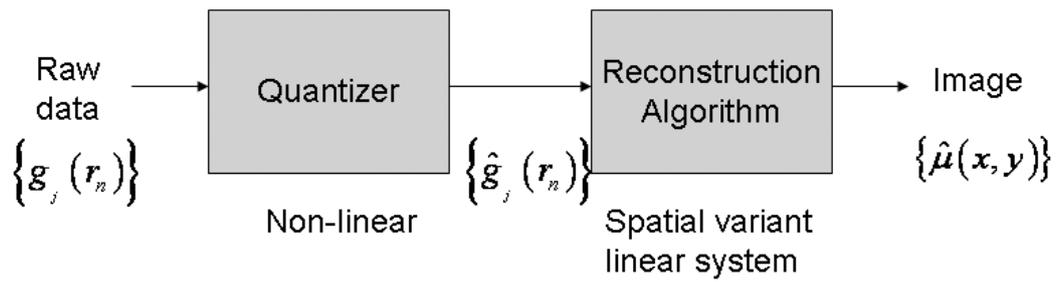
### 1.3 Project Goal

In our project, we will consider quantization and fixed rate lossy compression for CT imaging raw data (sinogram). We will design our quantizer considering the data properties of sinograms.

Our problem can be explained by the block diagram in Fig. 2. We will compress the raw data, and then the raw data is reconstructed to form the image  $\{\hat{\mu}(x, y)\}$ , and our ultimate goal is to keep the maximum difference between the image reconstructed from the compressed and from the uncompressed raw data to be “visibly indistinguishable.”

Typically, the reconstructed images are displayed in integer units of Hounsfield Units (HU) [DeM01], from -1000 to +3000, so errors that we introduce from compression of  $\pm 1$  HU are within the contrast resolution capability of CT systems. Therefore, when we subtract the reconstructed compressed image from the original reconstructed image, if the peak difference is within  $\pm 1$  HU, then “no one will complain.” Thus our distortion measure is defined in terms of this difference image.

However, if we use a distortion measure in the reconstructed image space, since the image reconstruction process is not spatial invariant, it is not a convenient distortion measure for Lloyd types of iterative algorithms (for each quantizer, we have to reconstruct the image, find out the distortion, and go back and forth.) To solve this problem, we study the relationship between the quantization errors in the sinogram space with the errors in the reconstructed image, and introduce a frequency weighted distortion measure in the sinogram space. The second distortion measure directly relates the quantized data with raw data.



**Figure 2:** Block diagram of CT raw data quantization/compression model.

The contributions of this project are: 1) study a distortion measure that relates the raw data space to the reconstructed image; 2) without changing rate, we consider two approaches to achieve a lower noise level due to quantization in the center region of the reconstructed image: first, noise shaping, and second, sub-band coding using DFT and bit-loading.

## 2 Quantizer Model

We will consider a scalar uniform quantizer with bin width  $\Delta$  and  $N$  levels. In our project we will use 8 to 10 bits for CT data compression, which corresponds to a minimum number of quantizer levels of  $N = 2^8$ . Under these high rate and small distortion assumptions, we can use the high rate Bennett’s approximation [Gra07], and model the quantizer error as additive white process uncorrelated with each other and uncorrelated with the quantizer output. The quantizer model is given by:

$$g_{\theta_i}(r_n) = \hat{g}_{\theta_i}(r_n) + e_{\theta_i}(r_n), \quad n = 1, \dots, N_r, \quad i = 1, \dots, N_{\theta}.$$

where  $\hat{g}_{\theta_i}(r_n)$  is the quantizer output and  $e_{\theta_i}(r_n)$  is the quantizer error.

Under the high rate assumptions, the quantizer error  $e_{\theta_i}(r_n)$  follows the uniform distribution in  $[-\Delta/2, \Delta/2]$ , with mean 0 and variance  $\Delta^2/12$ .

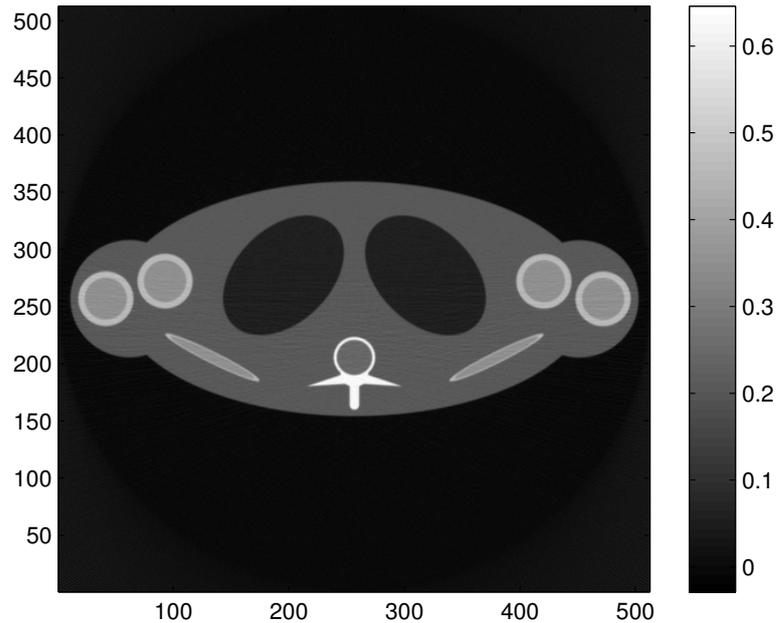
## 3 Distortion Measure

### 3.1 Quantization Error Frequency Property Study

We found that the low frequency band of the quantization errors contributes to the center of the reconstructed image, and the high frequency band error contributes more to the peripheral region of the image. (In fact, noise in different detectors will also contribute to different parts in the image, as studied in Adam’s project report. In this report, we will explore the frequency property of quantization error.)

To study the quantizer noise distribution property, we created a thorax phantom as shown in Fig. 3, and simulated the CT data using CatSim [DBC<sup>+</sup>07], a GE proprietary package developed for the express purpose of simulating various CT configurations. In fact, without such simulation software, raw data is exceedingly difficult to acquire since all of the processing to produce the final reconstructed image is “hidden away” within a commercial system. The image reconstruction block is also implemented using CatSim. The so-obtained CT raw data are represented with 32 bit floating point accuracy.

To study the effects of different frequency components of the quantization errors in the reconstructed image, we performed the following experiment. We added three frequency banded (with equal total bandwidth) noise signals to the 1D DFT of the uncompressed sinogram (32 bits floating point) in the view direction (as opposed to 2D DFT). The noise is uniformly distributed in  $[-\Delta/2, \Delta/2]$ , and the total power is about the same in these three frequency bands. We find the dynamic range of the raw data to be  $[U_{\min}, U_{\max}]$ , and then use an 8-bit uniform scalar quantizer, with  $\Delta = \frac{U_{\max} - U_{\min}}{2^8}$ .



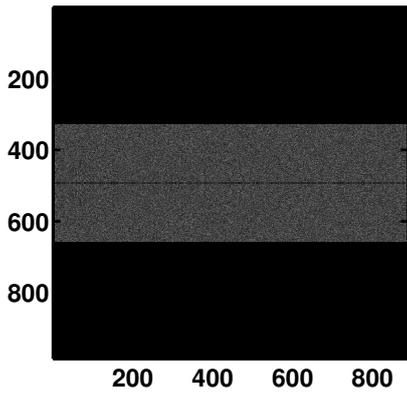
**Figure 3:** Reconstructed image for the thorax phantom, using raw data (32 bit floating point).

The frequency banded noises are shown in Fig. 4(a), Fig. 4(b) and Fig. 4(c). The resulted noisy sinograms are then reconstructed, and the difference images using the quantized and unquantized data are then found, as shown in in Fig. 4(b), Fig. 4(d), and Fig. 4(e). We can clearly see that the high frequency noise has less error in the center of the reconstructed image, whereas the low frequency noise has almost uniform noise throughout the image. Fig. 5(a) shows the absolute peak value (in the unit  $\text{cm}^{-1}$ ) in the difference image, as a function of distance  $d_i$  from the center of the image.

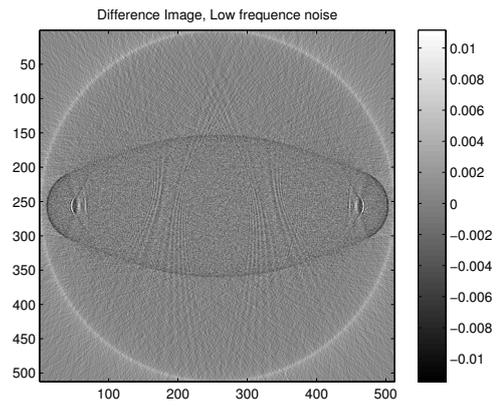
### 3.2 Distortion Measure

In many clinical applications, the center of the image contains the object of interest. So we would like to keep the error due to quantization/compression small in the center of the image, and allow a large error in the peripheral region of the image. For these reasons, we define the following distortion measure in the image space:

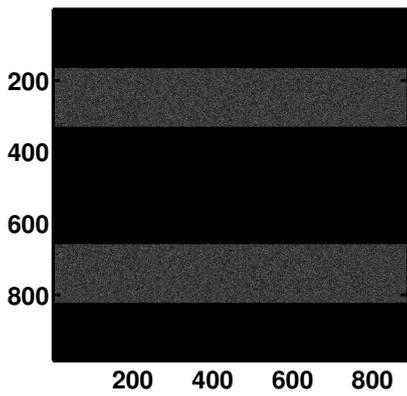
$$D_I = \frac{1}{N_p} \sum_{i=0}^{N_p-1} \alpha^2(d_i) \epsilon_i^2,$$



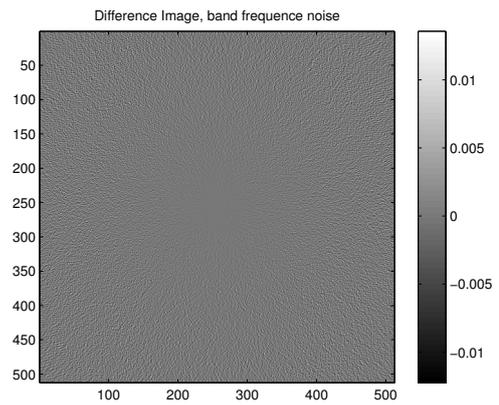
(a): Low band frequency band noise,



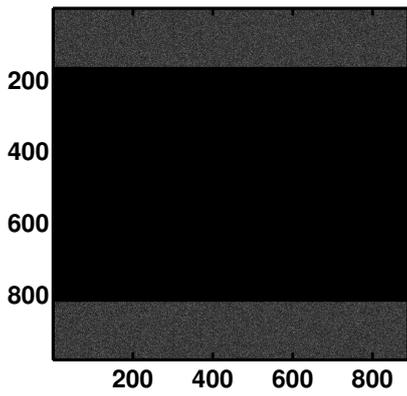
(b): Difference image, adding (a),



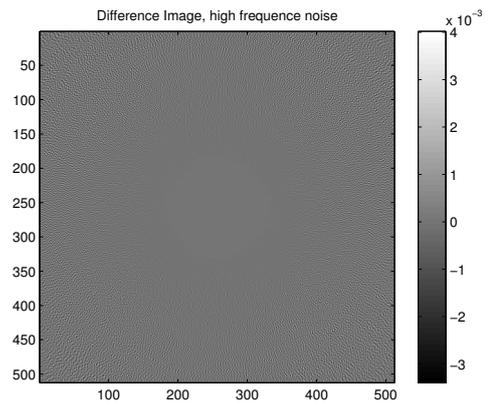
(c): Mid band frequency band,



(d): Difference image, adding (c)

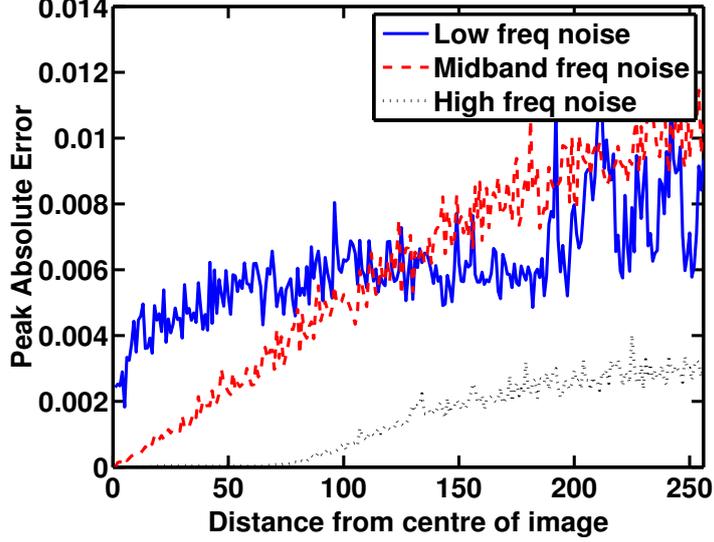


(e): High band frequency,



(f): Difference image, adding (e).

**Figure 4:** Different band frequency noise (used to modelling the quantizer error), added to DFT of the sinogram, and their corresponding difference images.



**Figure 5:** (a): peak error versus distance from the center of the image (pixels).

where  $N_p$  is the number of pixels,

$$\epsilon_i = \max_{x^2+y^2=d_i^2} |\mu(x, y) - \hat{\mu}(x, y)|$$

is the peak error in the difference image (with distance  $d_i$  from the center of the image), and weights  $\{\alpha_i^2\}$  are inversely proportional to distances with

$$\alpha_i^2 \propto 1/d_i^2, \quad \sum_i \alpha_i^2 = 1.$$

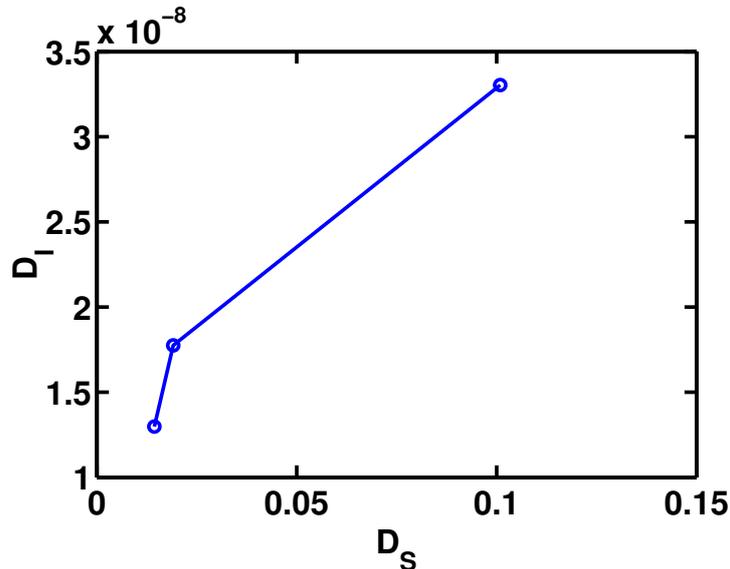
However, this distortion measure  $D_I$  is not easy to handle because it is not directly related to the raw sinogram data. From our study, we found that high frequency noise contributes to the peripheral region, whereas the low frequency noise contributes to the center of the reconstructed image. For this consideration, we can define a frequency weighted distortion measure. If we denote by capital letters the components of the discrete Fourier transform (DFT) of a signal, then in the frequency domain, the weighted MSE is given by:

$$D_S = \frac{1}{N_\theta N_d} \sum_{j=0}^{N_\theta-1} \beta_j^2 \sum_{n=0}^{N_d-1} |\hat{G}_n(f_j) - G_n(f_j)|^2 = \frac{1}{N_\theta N_d} \sum_{j=0}^{N_\theta-1} \beta_j^2 \sum_{n=0}^{N_d-1} |E_n(f_j)|^2.$$

The weights  $\{\beta_j\}$  are inversely proportional to frequency of the noise:

$$\beta_j^2 \propto 1/f_j^2, \quad \sum \beta_j^2 = 1.$$

Here we let  $\beta_0 = 1$ . This distortion directly measures the distortion in the sinogram space due to quantization and compression. Note that it is similar to the (relate to the Itakura



**Figure 6:**  $D_i$  versus  $D_f$ , in Figs. 4(b), 4(d), and 4(f).

distortion measure [Ita75]) It can be used as a distortion measure for quantization, such as in the Lloyd algorithm [Gra07].

Then we measure the two distortions,  $D_S$  and  $D_I$ , for these three cases as shown in Fig. 6(b). Although we only have three points, we can see these two distortion measures are positively correlated. In the following, as a heuristic, we can minimize  $D_S$  to achieve the goal of minimizing  $D_I$  (because  $D_S$  is under the control of the quantizer, whereas  $D_I$  is affected by image reconstruction algorithm).

## 4 Error Diffusion Coding

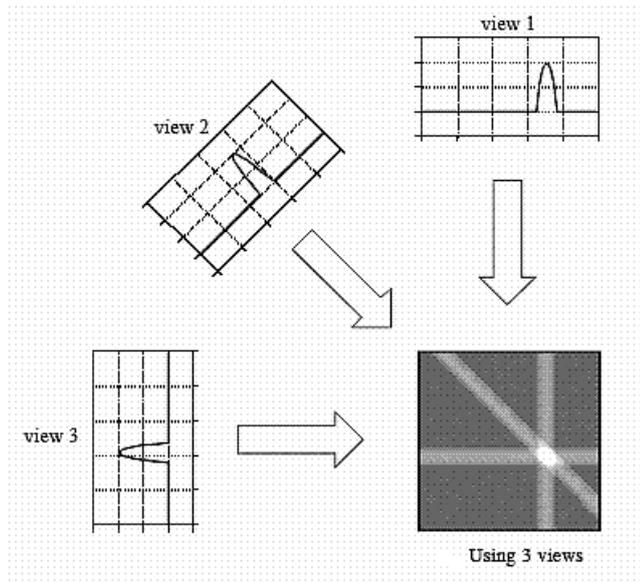
### 4.1 Error Feedback Quantization

Our motivation for error feedback quantization is to ensure that the quantization error does not accumulate in the reconstructed image. We find that it is related to the early work of error diffusion code [Ana89]. However, we come to our algorithm from a different perspective.

To simplify our analysis, we consider the parallel beam CT image reconstruction algorithm (which is a good approximation to the fan-beam reconstruction algorithm in most cases). The parallel beam reconstruction algorithm forms the reconstructed image by back-projecting all the filtered rays, as shown in Fig. 7<sup>2</sup>:

$$\hat{\mu}(x, y) = \sum_{i=0}^{N_r-1} \sum_{j=0}^{N_\theta-1} (\hat{g}_{\theta_j}(r_i) * c(r_i)) \delta(x \cos \theta_j + y \sin \theta_j - r_i),$$

<sup>2</sup>Courtesy: [www.dspguide.com](http://www.dspguide.com)



**Figure 7:** Back projection for parallel beam geometry.

the backprojection filter  $c(r_i)$  can be well approximated by a delta function  $\delta(\cdot)$ . Then the reconstructed pixel value is simply determined by summing all the projection ray passing through that pixel.

Our error feedback quantization method is given by (for notational simplicity, we drop the dependence of variables on  $r_n$ ):

- Set  $e_{\theta_1} = 0$ ;
- For each reading  $g_{\theta_i}$

$$\begin{aligned}
 g'_{\theta_i} &= g_{\theta_i} + e_{\theta_{i-1}}; \\
 \hat{g}_{\theta_i} &= \mathcal{E}(g'_{\theta_i}) \quad (\text{the codeword to be transmitted}); \\
 e_{\theta_i} &= g'_{\theta_i} - \mathcal{D}(\hat{g}_{\theta_i})
 \end{aligned}$$

where  $\mathcal{E}$  is the encoder and  $\mathcal{D}$  is the decoder. We note that this scheme is related to error diffusion coding with a particular noise shaping filter [Ana89].

The quantization error will not accumulate in the reconstructed image in this way. As an example, we consider the pixel at the center of the image, whose value is determined by the sum of all the measurements from the center detector. Consider all  $N_\theta$  measurements made by the center detector, and recall that compression introduces a zero mean error with

an RMS value of  $\Delta^2/12$ . If the error is independent for each of the  $N_\theta$  readings, the sum of the  $N_\theta$  readings will have an RMS value (DC error) of  $\sqrt{N_\theta}\Delta^2/12$ . However, using our proposed scheme, the RMS value of the sum will reduce to  $\Delta^2/12$  (see Appendix A for more detailed explanation).

This can also be interpreted from the frequency domain perspective. Indeed, the process of feeding the error due to compression into another reading will reduce the low temporal frequency components of the noise. If we consider the  $z$ -transform of all the quantities in Fig. 8(a), we have the  $z$  transform of the quantization error

$$E_n(z) = \hat{G}_n(z) - U_n(z) = \hat{G}_n(z) - G_n(z) + z^{-1}H(z)E(z) = \hat{E}_n(z).$$

And in our error feedback system, as shown in Fig. 8(a),  $H(z) = 1$ . Hence, the quantization error  $\hat{e}_n(i) = \hat{g}_n(i) - g_n(i)$  is related to the quantizer error  $e_n(i) = \hat{g}_n(i) - u_n(i)$  by

$$\hat{H}(z) = \frac{\hat{E}_n(z)}{E_n(z)} = 1 + z^{-1}H(z).$$

The filter  $\hat{H}(z)$  has one zero at  $z = -1$ , hence creates a high-pass filter, whose frequency response is given in Fig. 10(a). Then the quantization error is a filtered version of the quantizer error. If the quantizer error is assumed to be an independent white noise sequence in the linearized model then the noise shaping is achieved by properly choosing  $H(z)$ .

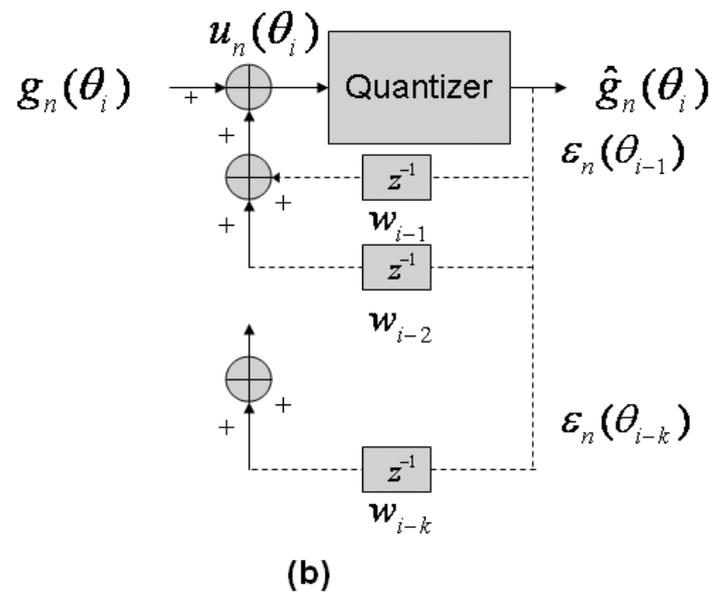
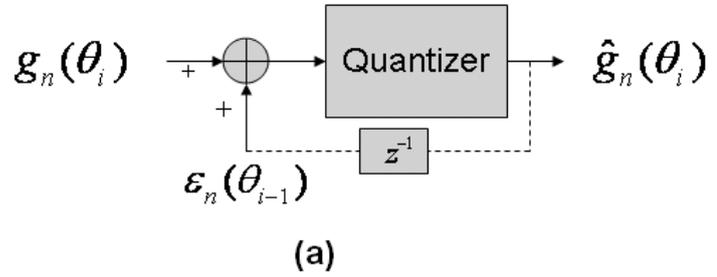
We apply this simple error feedback quantizer to the thorax phantom raw data in Fig. 3, reconstruct the image and find the difference image. Fig. 9(b) gives the spectrum of the noise (pixel value of the difference image). Note that when compared with the noise spectrum of the linear scalar quantizer without feedback in Fig. 8(a), the error feedback system passes the error through a high pass filter and achieves a desired noise spectrum shape.

Note that herein we only perform error feedback for each detector independently. We can also consider feeding errors across different detectors, however, it is not yet clear to us how this will benefit the final image reconstruction. Also, implementing error feedback online for each detector independently is easier in the CT gantry hardware.

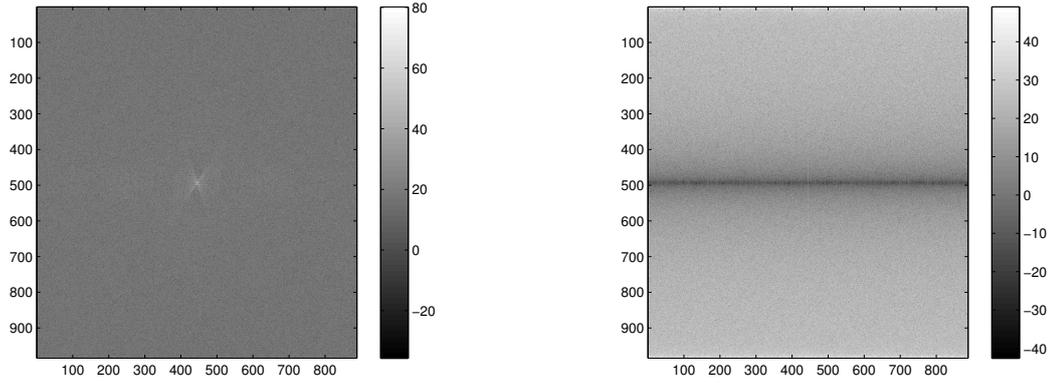
We may also consider feeding errors for a single view (independent of other views) across all the detectors. However, our preliminary simulation results have shown that this scheme will create more noise in the center region of the reconstructed image than the peripheral region. One possible explanation for this is that, in the image reconstruction algorithm, we have applied a high pass ramp filter across detector measurements, for each view. The image reconstruction ramp filter will counteract the efforts of our noise shaping filter, so we will not consider this type of error feedback.

## 4.2 Error feedback filter comparison

A more general case to the first order error feedback system in Fig. 8(a), is to design the filter  $H(z)$  to achieve our desired noise spectrum. Particularly, we would like to push the quantizer noise power to the high frequency region as much as possible.



**Figure 8:** Block diagram for the error feedback schemes: (a): our error feedback scheme is a first order system; (b): general error feedback scheme.



**Figure 9:** (a): quantizer error spectrum, without error feedback; (b): quantizer error spectrum with feedback.

In case of an FIR filter, the transfer function of the  $k$ th-order filter is given by

$$H(z) = \sum_{i=0}^{k-1} w_i z^{-i},$$

and the corresponding system is shown in Fig. 8(b).

However, stability is a major problem in nonlinear recursive systems [TA88]. In our case, we should have

$$w_i < 0, \quad i = 0, \dots, k-1.$$

for the system to be stable. If not, say, when we have some noise oscillating from  $+1$  to  $-1$ , then the feedback error can easily accumulate (which also happens in our simulations). So we would rather not use this two pole system.

Note that these weights should add to  $-1$ ,

$$\sum_{i=0}^{k-1} w_i = -1;$$

i.e.,  $H(z)$  should be unity-gain, so that the exact value of the quantizer error is diffused. Because of the unity-gain property,  $H(1) = 1$ ,  $\hat{H}(z)$  is guaranteed to have a zero at  $z = 1$  (dc value).

Also note that it is possible to arrange the  $N$  weights  $w_i$  so that they all sum to one and there are  $N$  zeros of  $H(z)$  at  $z = 1$  (so that we can have a sharper cutoff frequency). For example, for a second order system  $N = 2$ , the choice is  $w_1 = 2$ ,  $w_2 = -1$ , which creates two poles at  $z = 0$ , and whose frequency response is given in Fig. 10(b).

We examine several different types of FIR filters: another second order FIR filter,  $h = [1, -0.5, -0.5]$ , and a third order FIR filter,  $h = [1, -0.4, -0.3, -0.3]$ , all of which guaranteed the stability of the system. The frequency responses for them are given in Fig. 10.

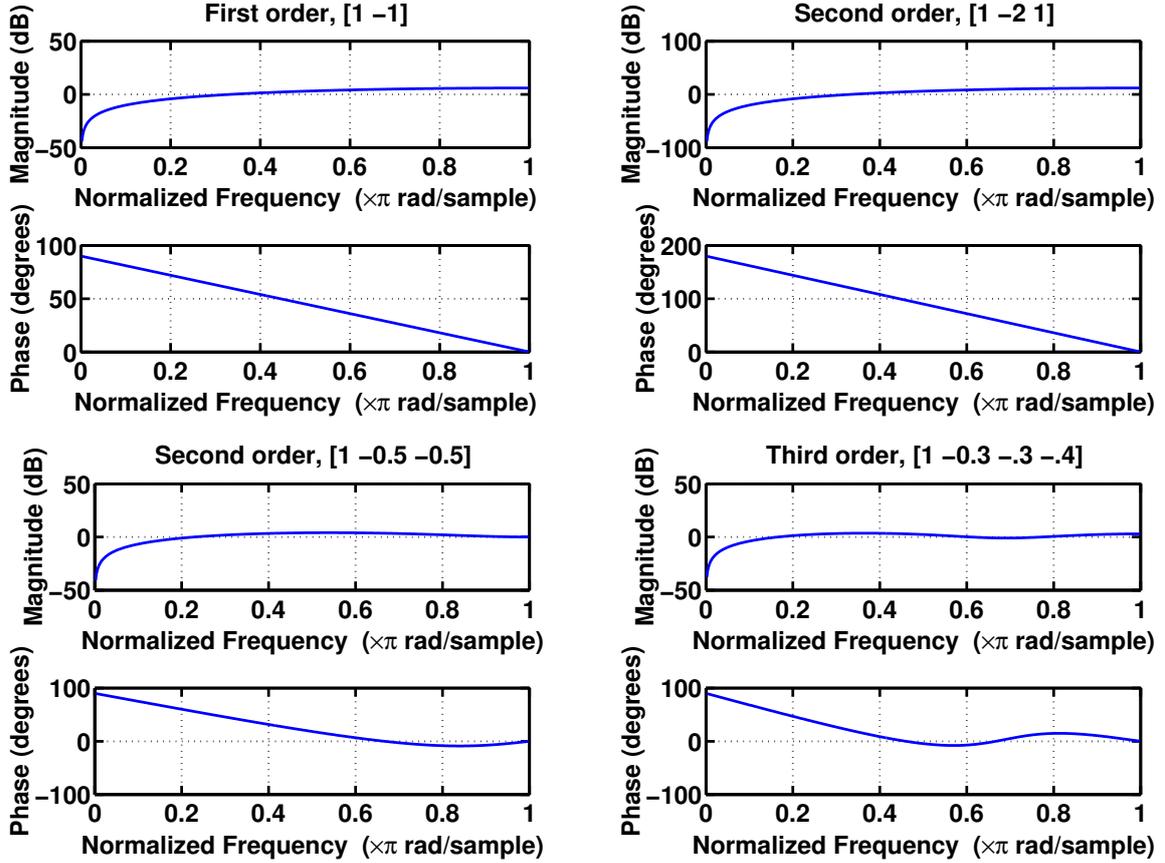


Figure 10: Comparison frequency response of different feedback filter.

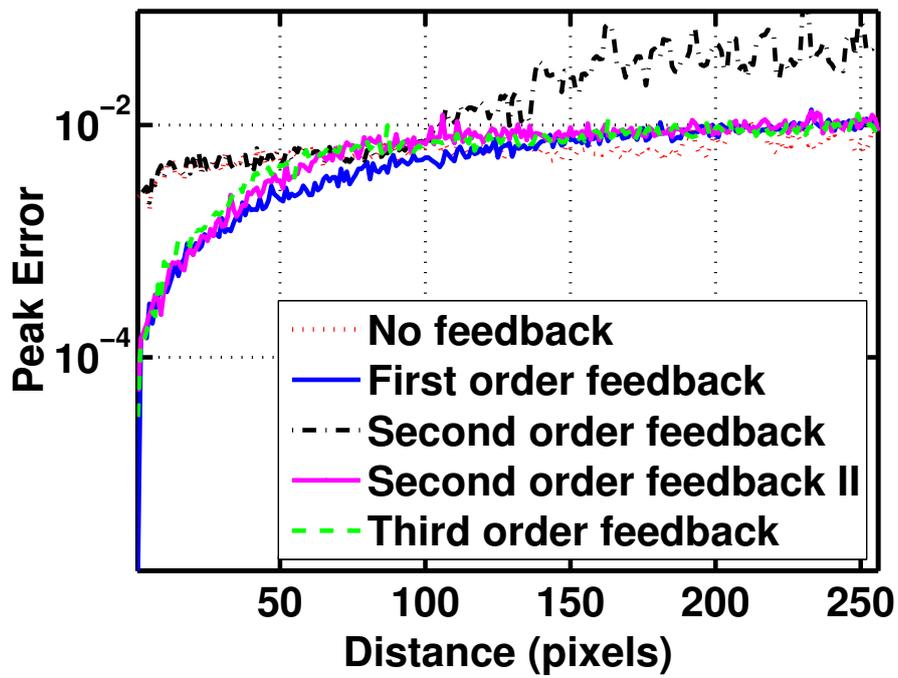
Fig. 11 gives the peak error in the image as a function of distance  $d_i$  from the center of the image. The first, second, and third order stable filters have similar performances. So we would rather use the simplest first order error feedback.

Fig. 12 shows the rate distortion curves for different error feedback quantization schemes. With more bits for quantization, we have less distortion  $D_i$ , but at the cost of higher rate. Note that we have an error floor, i.e.,  $D_s$  stops decreasing even when the number of bits continues to increase.

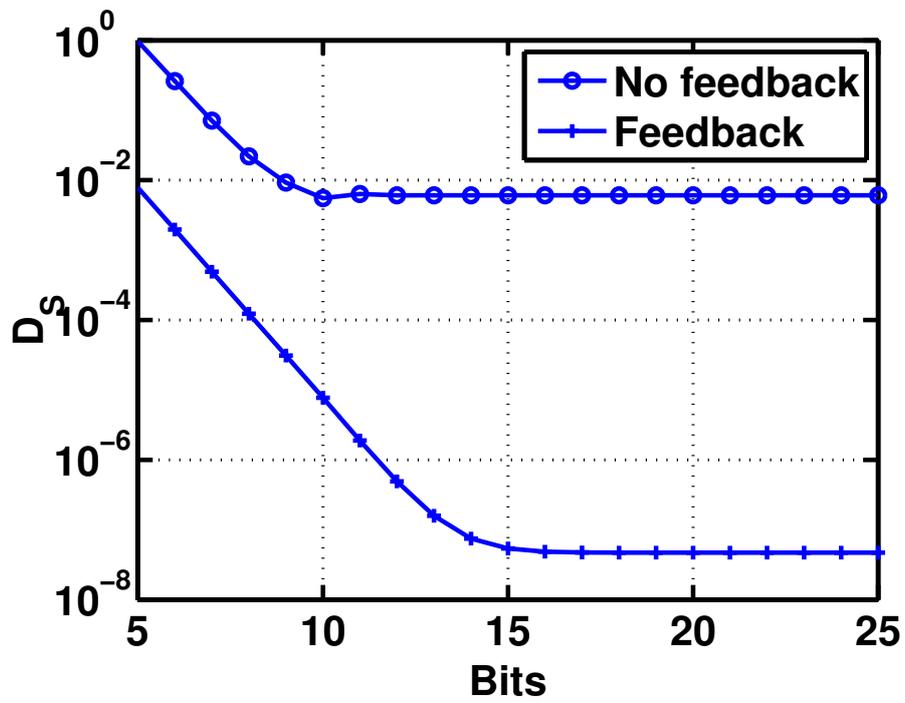
Interestingly, we found out that our goal for CT data compression is similar to the case of oversampled  $\Sigma\Delta$  modulation [Ana89], where it is desired to minimize the error at the very low frequencies (around  $z = 1$ ).

## 5 Sub-Band Coding

We can also achieve our quantizer noise shaping goal by frequency subband coding. We can transform the sinogram into the frequency domain, and then use a different number of bits for different frequency bands, while still maintaining the same rate. In our CT data



**Figure 11:** Comparison of various error feedback filters (8 bits uniform quantizer), in terms of peak errors (1/cm) in pixel values versus the distance from the center of the image.



**Figure 12:** Rate distortion curves of the first order feedback system and that without feedback (8 bits uniform quantizer).

compression case, to have an equivalent high pass filtering effect on the quantizer error, we can use more bits for low frequency bands and less bits for high frequency bands.

## 5.1 Coefficients Truncation

By examining the DFT coefficients of the sinogram, we find that most of the coefficients are quite small. So we can truncate the small coefficients while keeping the distortion of sinogram due to truncation under a certain level.

We will quantize/compress the real and imaginary parts of the Fourier coefficients of the sinogram separately (we could also quantize the magnitude and phase of the Fourier coefficients; but the real and imaginary parts of the Fourier coefficients should have similar dynamic range and be easier to design quantizers for).

Since we have a real object, the Fourier coefficients are Hermitian symmetric, so we could only quantize half of the Fourier coefficients. At the same time, we have to quantize the real and imaginary parts of the coefficients. So when comparing the rates of transform coding methods with the quantizer in image space, we should take this into consideration.

Since lower frequency components are more important to the error level in the center of the reconstructed image, we will define the frequency weighted truncation level. The DFT of the sinogram is frequency weighted,

$$\mathbf{Re}\hat{G}_n(\theta_i) = \begin{cases} \mathbf{Re}\hat{G}_n(f_i), & |\beta_i^2 \mathbf{Re}\hat{G}_n(f_i)| > \max\{\beta_i^2 \mathbf{Re}\hat{G}_n(f_i)\}/c; \\ 0, & \text{Otherwise.} \end{cases}$$

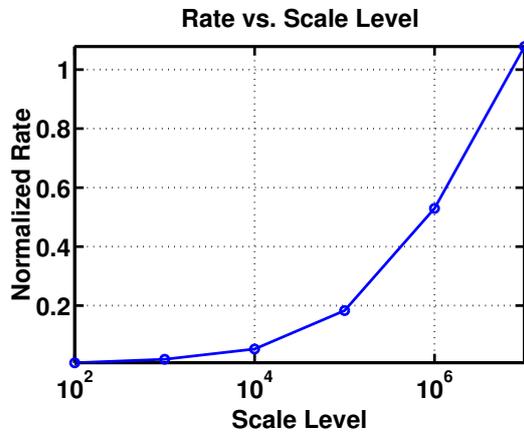
where  $c$  is the truncation level. A similar truncation algorithm was applied to imaginary part of the DFT sinogram.

Determining how many coefficients need to be kept (or deciding the truncation level) involves a trade off between the distortion and rates. The more non-zero DFT/DCT coefficients we keep, the less the distortion of the sinogram to the original sinogram, however, at the same time the rates are higher. This trade-off is shown in Fig. 13. We can choose the truncation level according to our desired MSE level and rates. Herein we choose the truncation level to be  $2e^{-6}$ .

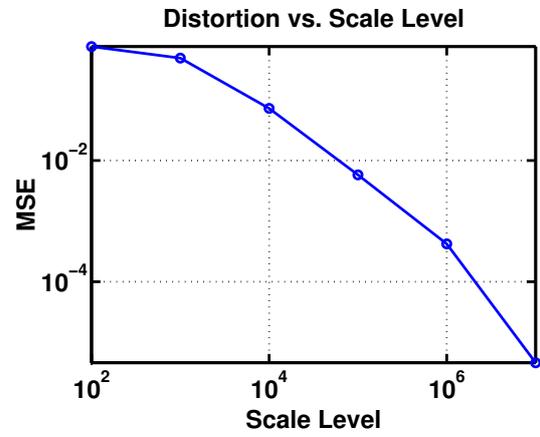
## 5.2 Bit allocation

Keep the same overall rates, we can allocate more bits to quantize low frequency bands of the sinogram, and less bits for high frequency bands of the sinogram. In this way we can manually form the quantized error spectrum. For example, while keep the same rate as that of an  $N_b$  bit quantizer, we can use  $N_b + v$  bits,  $N_b$  bits, and  $N_b - v$  bits, for the high, mid, and low frequency bands data respectively, where  $v$  is bits variation with respect to the center band.

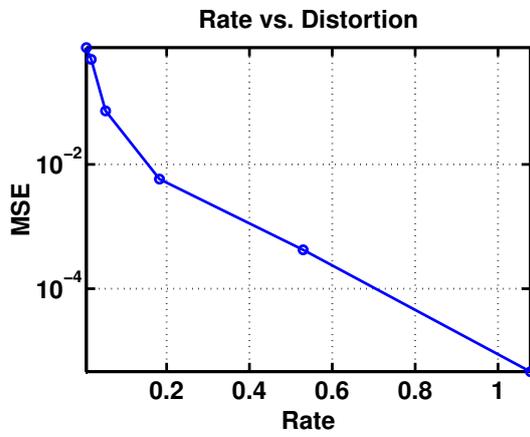
We will study the noise shaping effects for  $v = 1, \dots, 5$ , for  $N_b = 8$  bits. The difference image of the reconstructed image using different bit allocation schemes are shown in Fig. 14. We can achieve the same goal of pushing the error to the peripheral region of the



(a)

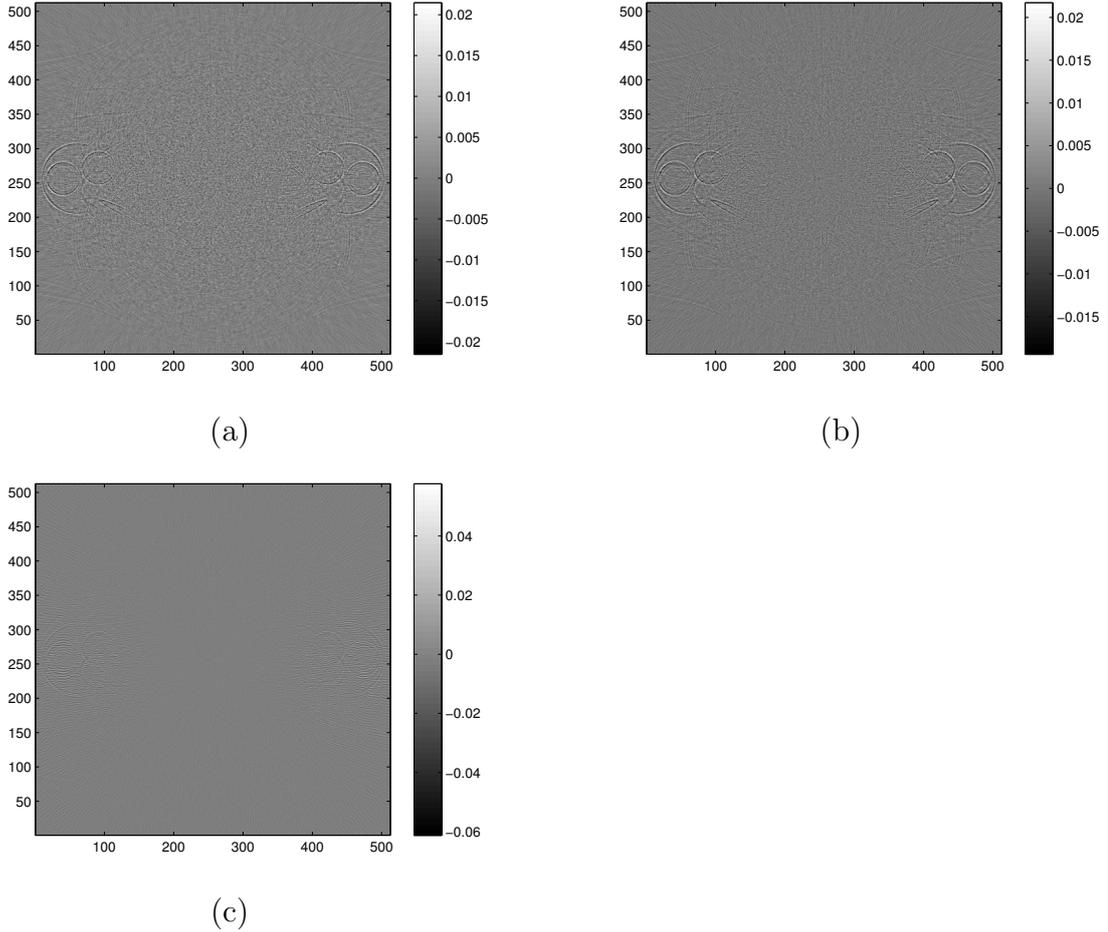


(b)



(c)

**Figure 13:** (a): normalized rate as a function of scale level; (b): distortion (MSE) as a function of scale level; (c): rate-distortion curve.



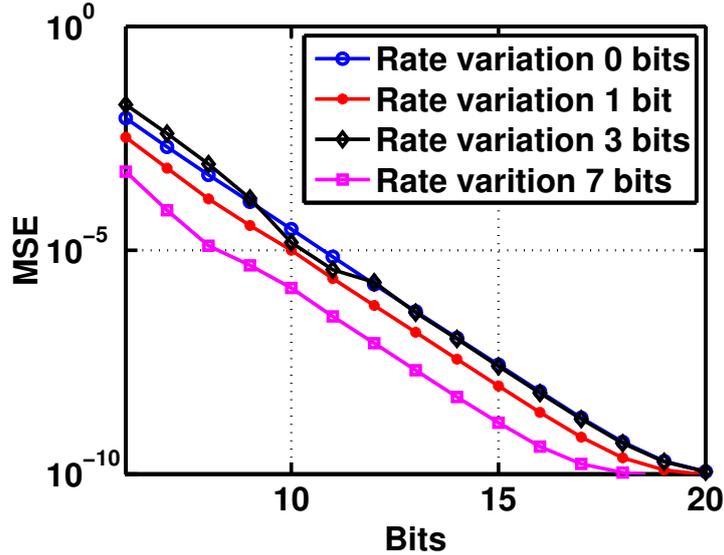
**Figure 14:** (a): sub-band bit allocation, 8-8-8 bits (low-mid-high frequency band); (b): bit allocation, 9-8-7 bits (low-mid-high frequency band); (c): bit allocation, 11-8-5 bits (low-mid-high frequency band). The images are displayed in the 1/cm unit.

reconstructed image by using sub-band coding and bit allocation as that achieved using our error feedback system.

The peak error in the image as a function of distance  $d_i$  from the center of the image is shown in Fig. 18 (compared with a Lloyd quantizer as described below).

The rate distortion curves for different bit allocation schemes are shown in Fig. 15. We found that bit allocation can achieve a lower distortion than that can be achieved by error feedback quantization.

The main advantage of using sub-band coding with bit allocation is that it can achieve better distortion performance compared with the error feedback quantization. However, the DFT approach can be implemented only after we have acquire data from all views, so unlike the error feedback quantization, it cannot be implemented online. The DFT sub-band coding approach may be more useful for sinogram storage compression.



**Figure 15:** Rate distortion curve of different bit loading schemes (The vertical axis should be  $D_s$ . Sorry about the confusion).

### 5.3 Lloyd algorithm

We can find a better quantizer by using Lloyd algorithm. Here we use Lloyd algorithm to design quantizer for our sub-band coding. We apply Lloyd algorithm on the real and the imaginary parts of the coefficients, for different sub-bands data, respectively.

We use half of the thorax data as training data to the Lloyd quantizer, and apply that on the entire thorax phantom raw data. We test two sets of initial conditions for the Lloyd algorithm, the uniform quantizer and the quantizer level that follows the quantizer point density as predicted by the Gersho's approximation [Gra07]. They converge to different quantizers, which shows that our problem is a nonconvex problem and has many local minima.

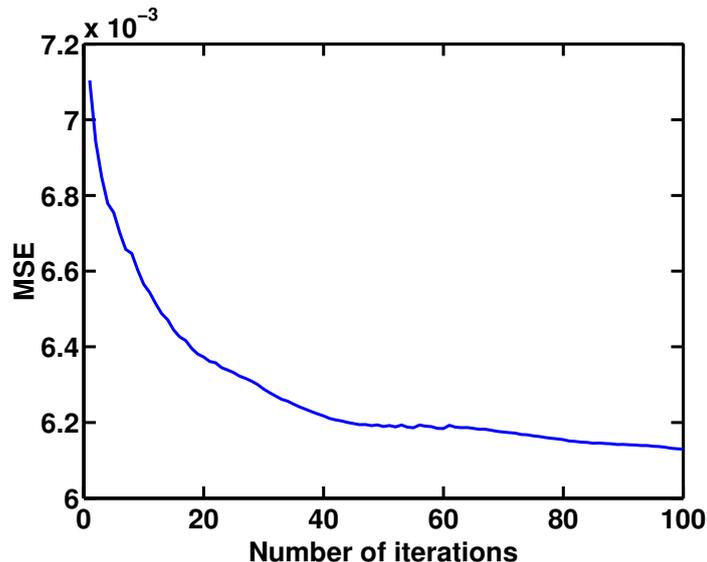
The results shown below are using a uniform quantizer as the initial condition. The MSE as a function of iteration number for the Lloyd algorithm is shown in Fig. 16. The so obtained Lloyd quantizer for different frequency bands data are shown in Fig. 17.

Fig. 18 shows the peak error in the image for all bit allocation schemes. Clearly we can have a better performance by using Lloyd optimal quantizer.

However, we found that in our problem, the quantizer is quite object dependent. The quantizer trained from the thorax raw data may perform quite poorly on the shoulder raw data (this issue is discussed in more detail in Adam's report.)

## 6 Discussion and future work

In our project, we considered the raw data compression problem for CT imaging. In designing our quantizer, we take into consideration the properties of sinogram. We found that the low



**Figure 16:** MSE convergence for a particular Lloyd algorithm realization.

frequency noise is responsible for the error in the center of the image, whereas the high frequency noise contributes to the peripheral region of the image. Thus we considered two types of noise shaping approaches: error feedback scheme, and sub-band coding with bit allocation.

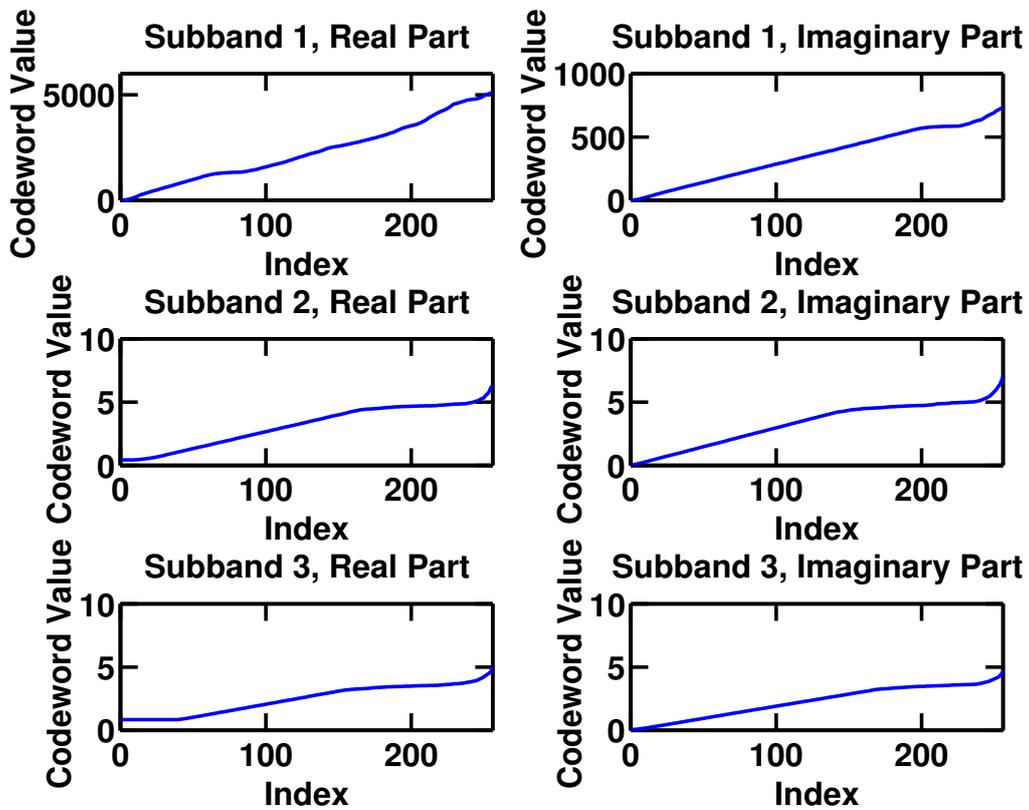
We found that, error feedback quantization (with first order filter) is the simplest one and has similar performance with other higher order FIR feedback filter. The second approach, sub-band coding with bit allocation, has better MSE performance when the number of bits is large; it may be more useful for sinogram data storage compression since it cannot be implemented in real time.

Given more time, we would like to explore the following aspects of this problem.

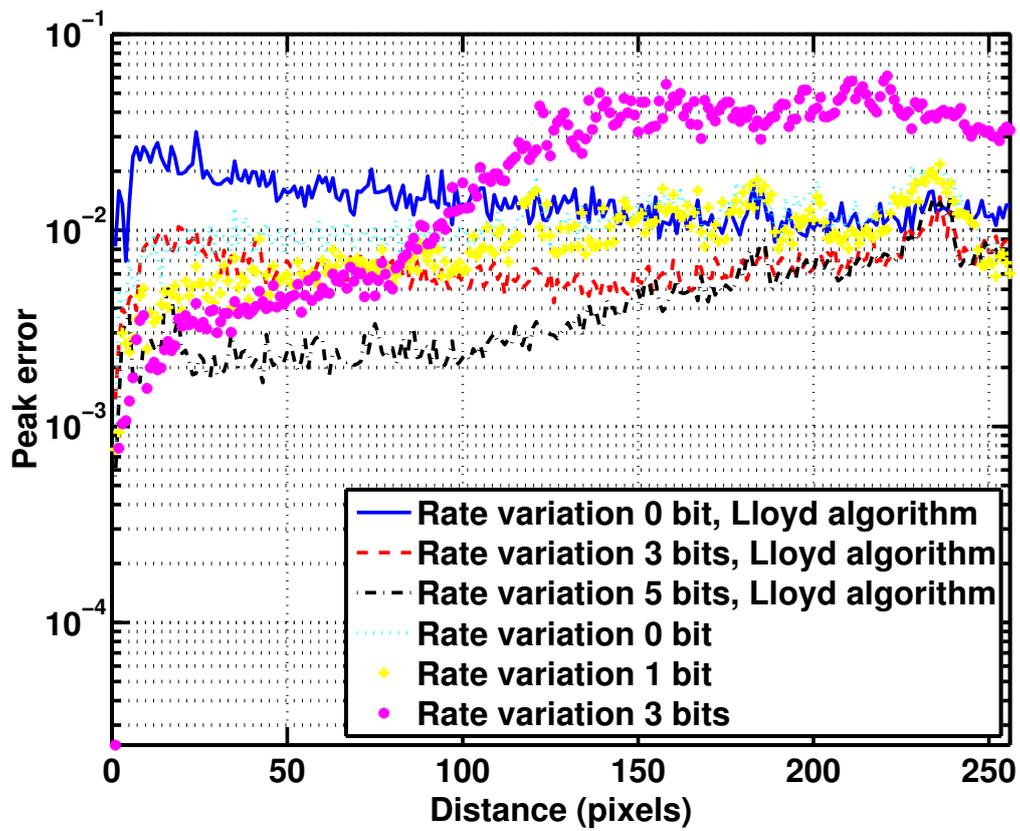
1. Vector quantizer. Herein in our project, we only consider scalar quantizers. We may achieve better compression by using vector quantization. For example, we can consider quantizing the data from all detectors, for one view, as a vector in the raw data space. We can use the LGB algorithm [GG92] on training data to obtain a vector codebook.

2. Bit allocation across all the detectors. Besides the frequency property of the sinogram, we can also consider the bit allocation schemes for different detector bands (the dependence as explained in Adam's report). We can use more bits to quantize the data from center detectors, and less bits for side detectors.

3. For now we only consider the fixed rate quantization. We can also consider variable rate coding (such as Huffman coding), to achieve an even better data compression.



**Figure 17:** Codewords found using the Lloyd algorithm (8 bits) for three subbands and their real and imaginary parts of the DFT coefficients.



**Figure 18:** Comparison of different subband bit allocation schemes, in terms of peak errors in pixel values versus the distance from the center of the image.

## A Noise Variance with Error Feedback

The pixel values near the center of the image are approximately determined by summing all the projections from a same detector. For each reading  $g_{\theta_i}$ , we have (using the additive noise model)

$$\begin{aligned}g_{\theta_1} - e_{\theta_1} &= \hat{g}_{\theta_1}; \\(g_{\theta_2} + e_{\theta_1}) - e_{\theta_2} &= \hat{g}_{\theta_2};\end{aligned}$$

So

$$\hat{g}_{\theta_1} + \hat{g}_{\theta_2} = g_{\theta_1} - e_{\theta_1} + (g_{\theta_2} + e_{\theta_1}) - e_{\theta_2} = g_{\theta_1} + g_{\theta_2} - e_{\theta_2};$$

In a similar way, we have

$$\sum_{i=1}^{N_\theta} \hat{g}_{\theta_i} = \left( \sum_{i=1}^{N_\theta} g_{\theta_i} \right) - e_{\theta_{N-1}};$$

so the noise variance in the sum with error feedback is  $\frac{\Delta^2}{12}$ .

Without using error feedback, we have

$$\sum_{i=1}^{N_\theta} \hat{g}_{\theta_i} = \sum_{i=1}^{N_\theta} g_{\theta_i} - \sum_{i=1}^{N_\theta} e_{\theta_i};$$

the noise variance accumulates to  $N_\theta \frac{\Delta^2}{12}$ .

Similar argument exists for neighboring pixels to the center of the image. However, for the peripheral pixels in the image, the pixel values are determined by summing the projection from different detectors. So the above analysis does not hold. So the error energy has actually be pushed from the center of the image to the peripheral region of the image.

## References

- [Ana89] D. Anastassiou. Error diffusion coding for A/D conversion. *IEEE Trans on Circuits and Systems*, 36(9):1175–1186, Sept. 1989.
- [BW01] K. T. Bae and B. R. Whiting. CT data storage reduction by means of compressing projection data instead of images: Feasibility study. *Radiology*, 219:850 – 855, 2001.
- [DBC<sup>+</sup>07] B. DeMan, S. Basu, N. Chandra, B. Dunham, P. Edic, M. Iatrou, S. McOlash, P. Sainath, C. Shaughnessy, B. Tower, and E. Williams. *CATSIM: a new computer assisted tomography simulation environment*, volume 6510 of *Medical Imaging 2007: physics of medical imaging*. Proc. of SPIE, 2007.

- [DeM01] Bruno DeMan. *Iterative reconstruction for reduction of metal artifacts in computed tomography*. PhD Thesis, 2001.
- [GG92] A. Gersho and R. M. Gray. *Vector Quantization and Signal Compression*. Kluwer Academic Press/Springer, 1992.
- [Gra07] R. Gray. *EE 372 Notes*. Stanford University, Spring 2007.
- [GRGJ04] Schmidt T. G., Fahrig R., Solomon E. G., and Pelc N. J. An inverse-geometry volumetric CT system with a large area scanned source: A feasibility study. *Med. Phys.*, 31(9):2623 – 2627, 2004.
- [Ita75] F. Itakura. Line spectrum representation of linear predictive coefficients of speech signals. *J. Acoust. Soc. Amer.*, 57, 1975.
- [SFP05] T. G. Schmidt, R. Fahrig, and N. J. Pelc. A three-dimensional reconstruction algorithm for an inverse-geometry volumetric ct system. *Med. Phys.*, 32:3234–3245, 2005.
- [TA88] T. C. Tsai and D. Anastassiou. Stable nonlinear recursive filters for edge enhancing noise smoothing of images. *Proc. IEEE Int. Symp. on Circuits and Systems*, 1988.