High-fidelity Image Compression for High-throughput and Energy-efficient Cameras

HIGH-FIDELITY IMAGE COMPRESSION FOR HIGH-THROUGHPUT AND ENERGY-EFFICIENT CAMERAS

BY

HENG WANG, B.Sc.

A THESIS

SUBMITTED TO THE DEPARTMENT OF ELECTRICAL & COMPUTER ENGINEERING

AND THE SCHOOL OF GRADUATE STUDIES

OF MCMASTER UNIVERSITY

IN PARTIAL FULFILMENT OF THE REQUIREMENTS

FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

Master of Applied Science (2010) McMaster University

(Electrical & Computer Engineering) Hamilton, Ontario, Canada

TITLE:	High-fidelity Image Compression for High-throughput and Energy-efficient Cameras
AUTHOR:	Heng Wang B.Sc., (Electrical Engineering) University of Science and Technology of China, Hefei, China
SUPERVISOR:	Dr. Xiaolin Wu
COSUPERVISOR:	Dr. Sorina Dumitrescu
NUMBER OF PAGES:	xiii, 51

To My Mom, For Everything

Abstract

In this thesis, we propose a new encoder-friendly image compression strategy for high-throughput cameras and other scenarios of resource-constrained encoders. The encoder performs L_∞-constrained predictive coding (DPCM coupled with uniform scalar quantizer), while the decoder solves an inverse problem of L_2 restoration of L_∞-coded images. Although designed for minimum encoder complexity (lower than distributed source coding and compressive sensing), the new codec outperforms state-of-the-art encoder-centric image codecs such as JPEG 2000 in PSNR for bit rates higher than 1.2 bpp, while maintaining much tighter L_∞ error bounds as well. This is achieved by exploiting the tight error bound on each pixel provided by the L_∞-constrained encoder and by locally adaptive image modeling.

Acknowledgements

I would like to first thank Dr. Xiaolin Wu for supervising and supporting my work. His enthusiasm, encouragement and patience were essential to the completion of this project. I am also grateful to my co-supervisor Dr. Dumitrescu Sorina for her patience and suggestions.

Special thanks goes to my readers, Dr. Nicola Nicolici and Dr. Shahram Shirani for their valuable input. Thanks to Cheryl, Helen and all other Electrical and Computer Engineering administrative staffs for their friendly assistance in the past few years. Sincere thanks goes to my colleagues Xiangjun, Mingkai, Ying, Huazhong, Xiao, Niu, Zhai, Zhou, Jie, Yong and Reza at the Multimedia Signal Processing Laboratory. Their help and friendship have made my master study a happy experience.

Last but not least, I would like to express my grateful thanks to my mother Jingfeng Liu and my father Kaiming Wang, for everything they did for me.

List of Abbreviations

2D	Two-dimension
CALIC	Context Adaptive Lossless Image Compression
CCD	Charge-coupled Device
CMOS	Complementary metal-oxide-semiconductor
CS	Compressive Sensing
DCT	Discrete Cosine Transform
DPCM	Differential Pulse Code Modulation
DSC	Distributed Source Coding
DWT	Discrete Wavelet Transform
EBCOT	Embedded Block Coding with Optimized Truncation
EZW	Embedded Zerotree Wavelet
GAP	Gradient-Adjusted Prediction
J2K	Joint Photographic Experts Group 2000
JPEG	Joint Photographic Experts Group
MRF	Markov Random Field
MSE	Mean Squared Error

PAR	Piecewise Autoregressive
PCM	Pulse Code Modulation
PSNR	Peak Signal-to-Noise Ratio
SPIHT	Set Partitioning In Hierarchical Trees
SQ	Scalar Quantizer
SSD	Solid-state Drive
VLC	Variable-Length Coding
WPAR	Weighted Piecewise Autoregressive

List of Notations

$(\bullet)^T$	the transpose of a vector or matrix
$\lvert \bullet \rvert$	cardinality of a set
$\lVert \bullet \rVert_\infty$	the infinity norm of a vector or a matrix
$\lVert \bullet \rVert_2$	the 2 norm of a vector or a matrix
$\mathbb{E}[\bullet]$	the statistical expectation operator
$\mathbb{Z}$	the field of integers
$\mathbb{R}$	the field of real numbers

Contents

Abstract iv

Acknowledgements v

List of Abbreviations vi

List of Notations viii

1 Introduction 1

1.1 Image Compression 1

1.2 Problem and Motivation 5

1.3 Organization 9

2 L_∞-Constrained Near-Lossless Compression 11

2.1 Review of Near-Lossless Image Compression 11

2.2 Near-Lossless CALIC 13

2.3 Compression Noise 16

3 Piecewise Autoregressive Image Model 20

3.1 Piecewise Autoregressive Image Model 20

3.2 Overfitting Problem 24
3.3 Weighted PAR model (WPAR) 26

4 Soft-Decision Decoding Algorithm with PAR modeling **28**
4.1 Formulation of the Optimization Problem 28
4.2 Mean Squared Error (MSE) Analysis 32
4.3 Experimental Results and Remarks 36

5 Conclusions **44**

List of Tables

4.1 Performance of Different Methods For $\tau = 1$. 38

4.2 Performance of Different Methods For $\tau = 3$ 38

4.3 Performance of Different Methods For $\tau = 5$ 38

List of Figures

1.1 Image Compression Process 2

1.2 Schematic description of DPCM 3

1.3 Schematic description of J2K 5

2.1 Schematic description of L_∞-constrained CALIC 14

2.2 CALIC prediction table 15

2.3 Compression artifacts in decompressed image with different L_∞ bounds 17

2.4 Median Mask 18

2.5 Median filter effect 19

3.1 8-order AR parameters 22

3.2 PAR models in two different directions 24

3.3 Example of non-stationary window 26

4.1 CALIC Error Prediction Distribution 33

4.2 Five Test Images in the test set 37

4.3 (a) Original image; (b) J2K decoded image; (c) Soft-decoded image; (d) Enlarged version of the red box of (a); (e) Enlarged version of the red box of (b);(f) Enlarged version of the red box of (c). 40

4.4 (a) Original image; (b) J2K decoded image; (c) Soft-decoded image; (d) Enlarged version of the red box of (a); (e) Enlarged version of the red box of (b);(f) Enlarged version of the red box of (c). 41
4.5 (a) Original image; (b) J2K decoded image; (c) Soft-decoded image; (d) Enlarged version of the red box of (a); (e) Enlarged version of the red box of (b);(f) Enlarged version of the red box of (c). 42
4.6 (a) Original image; (b) J2K decoded image; (c) Soft-decoded image; (d) Enlarged version of the red box of (a); (e) Enlarged version of the red box of (b);(f) Enlarged version of the red box of (c). 43

Chapter 1

Introduction

1.1 Image Compression

Digital images are common and convenient means of storing and transmitting information. As the saying goes, "One picture is worth more than ten thousand words", the spatial information about objects, such as positions, sizes, and shapes, is conveyed in the digital images. Since we have innate visual and mental abilities, human beings are good at deriving information from such digital images.

In an ideal situation, we should have enough space to store, process, and transmit digital images. However, given the fact that we are facing the massive size of image data and limited channel capacity, it would not be possible to store, process, and transmit digital images without image compression. Image compression, as the terminology implies, is to reduce the number of bits required to represent a digital image without degrading the quality to an unacceptable level. As shown in Fig. 1.1, the encoder compresses the original image into bitstream before any operation on the image, such as storage and transmission. After that, the decoder decompresses the

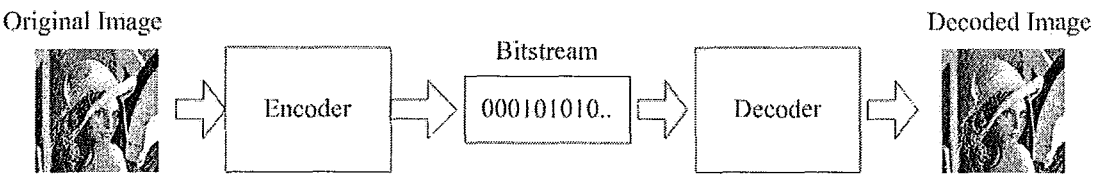


Figure 1.1: Image Compression Process

bitstream to reconstruct the original image or an approximation of it.

Based on the requirements of reconstruction, image compression schemes can be divided into two broad categories: lossless and lossy compression. Lossless image compression process does not incur any loss of information, which means the decompressed image is bit for bit identical of the original image. In contrast, certain loss of information is allowed in lossy compression. By allowing the introduction of small errors, one can expect that lossy compression would achieve higher compression ratio compared with that of lossless compression. Lossless compression is often demanded in some applications such as medical imaging and remote sensing, so as to avoid legal dispute over the significance of errors introduced into the imagery. Lossy compression works in the applications which need higher compression ratio with an allowable level of distortion.

Interest in image compression dates back more than 60 years. In 1948, Shannon and his students Oliver, Pierce published the paper about pulse code modulation (PCM) [1] for television signal, marking the beginning of the digital image compression technology. In the following, we will examine several lossless and lossy image compression techniques.

One simple and straightforward approach to lossless image compression is to directly use variable-length coding (VLC), which maps a source symbol to a variable number of bits. In general, the more probable symbol is represented with fewer bits

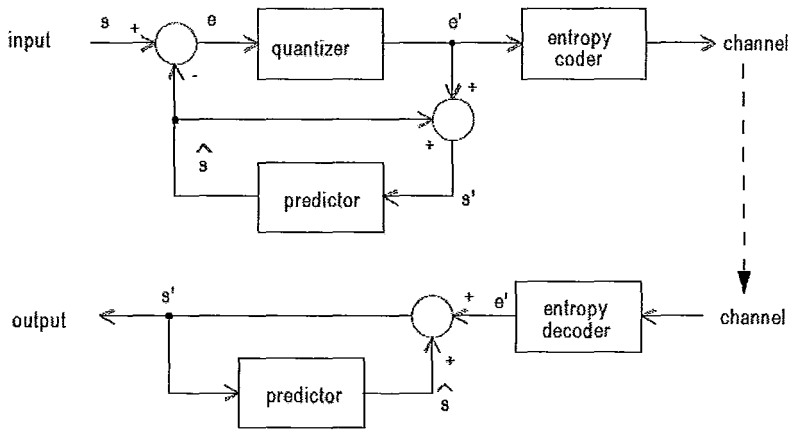


Figure 1.2: Schematic description of DPCM

(using a shorter codeword). There are many popular VLC algorithms such as Huffman coding, proposed by David Huffman in 1951 [2], arithmetic coding, developed by J.J. Rissanen and G.G. Langdon. [3], and Lempel-Ziv coding, proposed by Jacob Ziv and Abraham Lempel in 1977 [4].

The Differential pulse code modulation (DPCM) [5] is another approach for image compression by considering the correlation information instead of coding each pixel one by one. As shown in Fig. 1.2, DPCM is to do the prediction based on its previous pixels, and only encode and transmit the difference between the prediction value and the currently encoding pixel value. A DPCM coder that was originally developed for continuous-tone images is the Context Adaptive Lossless Image Compression (CALIC) proposed by Wu and Memon [6]. The CALIC algorithm selects the predictor from a number of pre-determined configurations according to the gradient of spatially adjacent pixels. More details concerning CALIC will be presented in Chapter 2.

Typically, lossless coding can achieve a compression ratio from 1.5:1 to 3:1 for

most natural images. To achieve further reduction of the data size , lossy coding methods apply quantization to the original samples or coefficients based on some transformation of the original signal. One popular lossy coding scheme for digital images is transform coding [7, 8] based on transforming methods such as Discrete Cosine Transform (DCT), Discrete wavelet transform (DWT). We divide an image into non-overlapping blocks in a block-based transform coding. For every block, we transform the original pixel values into a set of transform coefficients using a unitary transform. Then we quantize and encode the transformed coefficients. The transform should be designed not only to compact the structure of image signal into only a few coefficients, but to reduce the correlation among the pixels to be coded. In this way we could increase the compression ratio. Joint Photographic Experts Group (JPEG) compression standard, proposed in 1992 [9], is a representative transform coding approach using Discrete Cosine Transform (DCT).

Another popular approach to decompose an image signal into transformed coefficients is based on the wavelet transform [10], which could decompose the image into the sub-image signals with different spatial resolutions, frequencies, and directions. It works on the overall image, which makes it possible to remove the global relevance of the image signals, and distribute the quantization error into the whole image, avoiding the "mosaic" block effect. Some well-known implementations of the wavelet-based compression algorithms are Set Partitioning In Hierarchical Trees (SPIHT) proposed by Said and Pearlman [11], Embedded Zerotree Wavelet (EZW) proposed by Shapiro [12], Embedded Block Coding with Optimized Truncation (EBCOT) proposed by Taubman and Zakhor [13]. The wavelet transform has been adopted in the image compression standard JPEG 2000 (J2K) [14], which can be shown in Fig. 1.3.

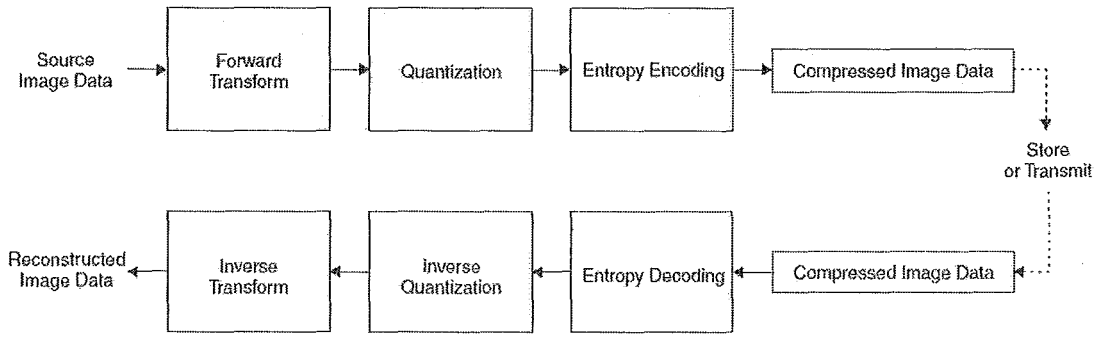


Figure 1.3: Schematic description of J2K

1.2 Problem and Motivation

This work is motivated by a new, important application of image compression: ultra high throughput imaging. As high-end semiconductor sensors become more and more sophisticated, the obtainable solutions of digital cameras in space, time, spectrum and amplitude are much less constrained by the speed, density, sensitivity, and signal-to-noise ratio of image sensors than by the write speed of on-camera mass storage device (e.g., SSD) that records the data, which is much slower than what is required by many high throughput imaging applications. In fact, modern CCD and CMOS technologies are able to capture color (three spectral bands) videos at more than 2M pixels/frame, more than 7000 frame/second, and 12 bits/sample, creating a raw data throughput of 21,000 MB/second or at least 7,000MB/second (if color is sampled in mosaic). To circumvent the severe shortfall in memory throughput, high-end camera manufacturers offer users trade-offs between spatial and temporal resolutions. For instance, the high-speed camera Phantom v710 of Vision Research can offer a spatial resolution of 1280×800 at 7530 frame/second, but it has to reduce the spatial resolution to 128×128 when operating at 215600 frames/second.

Indeed for modern cameras, compression is less critical for storage as the cost per megabyte steadily decreases, but it is the only means to relieve the throughput bottleneck created by grossly inadequate memory write speed. The problem becomes particularly challenging if the compression solution has to ensure high fidelity (e.g., applications in medicine, space, sciences and precision engineering), and the camera has to operate with limited power budget (e.g., cameras in outer space). In addition, the encoder has to run in real time, otherwise computation time simply nullifies the benefit of compression by blocking the output flow. Given all the constraints, an obvious way out is to use an asymmetric image codec that delegates the pursue of high compression performance and thus shifts the associated computation burdens to the decoder.

To meet the design goal of light-duty encoder, the approach is, in conventional wisdom, distributed source coding (DSC) [15, 16, 17, 18], or compressive sensing (CS) as many researchers recently advocated [19, 20]. In this thesis we demonstrate, however, that L_∞-constrained predictive coding [6], i.e., DPCM coupled with uniform scalar quantizer (DPCM+SQ), can offer a more compelling and advantageous solution than DSC and CS for low-complexity, high-throughput, and energy-efficient encoding of images. Arguably, the simplest image encoder is DPCM with uniform scalar quantization of prediction residuals without entropy coding. If the quantizer step size is $2\tau+1$, then this simple scheme can achieve, by setting reproduction values to the mid points of quantizer cells, an error bound of τ on each single pixel, i.e., $\|\mathbf{I}-\hat{\mathbf{I}}\|_\infty \leq \tau$, where $\mathbf{I}$ and $\hat{\mathbf{I}}$ are respectively the original and decompressed image. Therefore, DPCM+SQ was studied as an L_∞-constrained near-lossless (for sufficiently small τ) compression technique [21, 22, 23], mostly motivated by high-fidelity compression applications

(e.g., medical imaging and remote sensing). But a useful property of the DPCM+SQ architecture was overlooked. That is, the knowledge $\|\mathbf{I} - \hat{\mathbf{I}}\|_\infty \leq \tau$, which is freely available to the decoder, can be exploited by a soft decoding process to optimally recover the original image $\mathbf{I}$ in L_2 sense from the L_∞-decoded image $\hat{\mathbf{I}}$. By adopting a pixel-by-pixel hard-decision decoder, all published L_∞-constrained image compression methods forfeit the benefits of $\|\mathbf{I}-\hat{\mathbf{I}}\|_\infty \leq \tau$ and other prior information on the input image $\mathbf{I}$, and hence fall significantly short in the L_2 rate-distortion performance. If one can unlock the power of prior knowledge inherent to L_∞-constrained compression and eliminate the coding loss of hard-decision decoder, then the DPCM+SQ scheme becomes a viable compression solution for overcoming the write-out bottleneck of high-resolution and high-speed cameras, and for many other compression scenarios of resource-deprived encoders.

All existing L_∞-constrained near-lossless image coding techniques incur larger L_2 distortion, or lower PSNR, than lossy image compression techniques without L_∞ error constraint. Wu and Bao tried to improve the L_2 performance of predictive near-lossless image coding. They studied the adverse effect of residue quantization on the robustness of the predictor, and proposed adaptive context modeling techniques to detect and correct prediction biases caused by quantization errors. By incorporating bias cancelation into the prediction loop, Wu and Bao improved the PSNR results of L_∞-constrained predictive coding [24]. However, this work made the coding gain at the expense of greatly increased encoder complexity, and hence not suited for solving the bottleneck problem of high throughput cameras.

Instead of fine tuning the encoder to gain coding efficiency, we shift the task of removing redundancy to the decoder and keep the encoder complexity at minimum.

An obvious cause for the inferior L_2 performance of the existing L_∞-constrained coding methods is that the strong knowledge $|I_{i,j} - \hat{I}_{i,j}| \leq \tau$, $\forall(i,j)$ (where $I_{i,j}$ and $\hat{I}_{i,j}$ denote the original, respectively reconstructed value at pixel position (i,j)) is totally ignored and wasted. Another prior information about the image, which is also wasted by hard decision decoding, is that natural images can be satisfactorily modeled as Markov random fields (MRF). To capitalize on the above sources of prior domain knowledge, the decoder needs to go beyond the current practice of hard-decision pixel-by-pixel decompression of an L_∞ coded image. Therefore, we adopt a soft-decision decoding approach, starting from the decompressed image $\hat{\mathbf{I}}$ as an initial estimate of $\mathbf{I}$, and proceed to compute an improved estimate $\mathbf{X}$ of the original image $\mathbf{I}$ by solving the following inverse problem

$$\begin{aligned} &\min_{\mathbf{X}}\{\|\mathbf{X} - \mathbf{X}_{\mathbb{M}}\|_F + \lambda\|\mathbf{X} - \hat{\mathbf{I}}\|_F\} \\ &\text{subject to } \|\mathbf{X} - \hat{\mathbf{I}}\|_\infty = \tau \end{aligned} \tag{1.1}$$

where $\mathbf{X}_{\mathbb{M}}$ stands for an estimate of $\mathbf{X}$ that is generated by an image model $\mathbb{M}$ and $\|\bullet\|_F$ denotes some norm. In the objective function $\|\mathbf{X} - \mathbf{X}_{\mathbb{M}}\|_F$ is the regularization term and $\|\mathbf{X} - \hat{\mathbf{I}}\|_F$ is the fidelity term. The constraint $\|\mathbf{X} - \hat{\mathbf{I}}\|_\infty = \tau$ helps to confine the solution space of the underlying optimization problem.

With the above observations and motivation we propose a soft decoding technique of reestimating the L_∞-decoded image $\hat{\mathbf{I}}$. This is cast as an inverse problem of L_2 restoration of L_∞-coded images, in which the regularization term is based on a prior image model of piecewise autoregressive process, and convex constraints $\forall(i,j)$ $\|I_{i,j} - \hat{I}_{i,j}\| \leq \tau$ are imposed to confine the solution space. The piecewise autoregressive image model, via its parameters, offers an adaptive representation of natural images;

the model parameters are estimated locally from denoised samples of $\hat{\mathbf{I}}$ by solving an L_2-minimization problem. Once the PAR model is constructed, soft decoding can be performed efficiently by constrained linear least-square estimation.

The proposed soft decoding strategy can improve the PSNR of L_∞-coded image $\hat{\mathbf{I}}$ by up to 2dB. For bit rates above 1.2 bpp it can even outperform competitive encoder-centralized image codecs, such as JPEG 2000, in PSNR while achieving much tighter L_∞ error bounds as well. In contrast, DSC and CS, despite years of intensive research, still have large gaps in rate-distortion performance against traditional centralized compression methods. The performance advantage of the proposed asymmetric image codec of DPCM+SQ is commendable, considering the fact that it has a substantially lower encoder complexity than DSC and CS. It can compress an image of N pixels in $O(KN)$ operations, where K is the order of the predictor. Both DSC and CS have an encoder complexity of $O(N^2)$, because they generate code streams by computing $O(N)$ linear combinations of the N pixels. In addition, the new strategy of L_2 soft decoding of L_∞-constrained code stream offers a side benefit of tight L_∞ bound.

1.3 Organization

The rest of this thesis is structured into four chapters. In Chapter 2, we explain the necessity for the L_∞-Constrained Near-Lossless compression and describe the features of Near-Lossless CALIC algorithm. In Chapter 3, for the regularization term in (1.1), we assume an image model $\mathbb{M}$ of 2D piecewise autoregressive model and discuss the reason why we choose it. Then we present the mathematical solutions for the piecewise autoregressive model. In Chapter 4, by using the piecewise autoregressive image model as the prior information for the inverse problem (1.1), we built the

least-square optimization problem to further improve the decompressed image in L_2 sense. Since the L_∞-constrained near-lossless image coding techniques inherently have a strong form of side information that can be exploited by the decoder, we can then develop a constrained least-square problem with the assistance of the known tight error bound on each pixel. This approach can reduce the L_2 distortion and at the same time still maintain a known bound on the maximum possible error. The simulation results and discussion are presented in the end. In Chapter 5, we conclude this thesis.

Chapter 2

L_∞-Constrained Near-Lossless Compression

In this chapter, we explain the necessity for the L_∞-Constrained Near-Lossless compression. Then Near-Lossless CALIC, which is a popular L_∞-Constrained Near-Lossless Compression algorithm, is introduced. After observing the drawbacks of the Near-Lossless CALIC, we use the median filter to remove the salt and pepper noise.

2.1 Review of Near-Lossless Image Compression

Many important applications in medicine, sciences, space exploration, precision engineering, etc., have very stringent quality requirements on image compression algorithms. For users of these areas, images after being fetched and decompressed, are subject to rigorous computer analysis and hence the reconstruction quality is far more than just pleasing the eyes as in entertainment and consumer applications. The

ideal solution in such cases should be mathematically lossless image coding. However, after decades of extensive research, achievable lossless compression ratios remain stubbornly low, ranging from 1.5:1 to 3:1 depending on the image contents and imaging modalities. Even prohibitively expensive code optimization in minimum description principle cannot reduce lossless image bit rates of CALIC, a benchmark of practical good lossless image codecs, by more than five percent [25].

It may appear that we might as well do away with compression if no loss is allowed, given the fact that the storage cost per megabyte has been continuously decreasing. But high fidelity image compression is still important for high-end applications for a different reason: saving the I/O bandwidth between the camera sensor and memory. For modern professional digital cameras, such as those of ultra high spatial resolution or/and very high frame rate, and those in functional medical imaging, the I/O throughput becomes the system bottleneck: how quickly the image data can be off loaded from the sensor array and stored into memory? Compression, if done in real time, can increase the effective data throughput and thus enhance camera capability.

When an image $\mathbf{I}$ has to be compressed to a lower bit rate than lossless coding allows, an alternative is near-lossless or L_∞-constrained image coding. The L_∞ constraint is imposed on the compressor such that the compression error is tightly bounded for each single pixel, i.e., $\|\mathbf{I} - \hat{\mathbf{I}}\|_\infty \leq \tau$, where $\hat{\mathbf{I}}$ is the decompressed image and τ is a small positive integer. The L_∞-constrained coding strategy achieves the best of the both worlds: significantly higher compression ratio of lossy compression and a predetermined minmax fidelity that is close to lossless compression. In contrast, lossy image codecs designed under the L_2 criterion can incur large errors on some pixels that are statistical outliers. Such large individual errors, although with

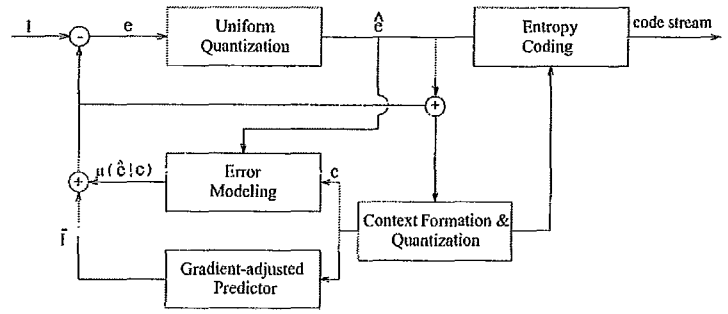
negligible contribution to PSNR, can be disastrous in some applications. For instance, in a satellite image a boat can be only few pixels in size; it can be easily removed by a codec that is optimal in L_2 sense but preserved by an L_∞-constrained codec.

A number of L_∞-constrained image compression techniques were proposed. Among them the simplest one is uniform scalar prequantization of pixel values in $\mathbf{I}$; the prequantized image $\hat{\mathbf{I}}$ is then losslessly coded, typically by predictive coding schemes like CALIC [6], or by reversible integer wavelet transforms [26, 27] if progressive transmission is desired. This prequantization and lossless coding strategy bounds the maximum error by $\|\mathbf{I}-\hat{\mathbf{I}}\|_\infty \leq \tau$, if the quantization step size is $2\tau+1$. An alternative strategy, which delivers higher coding efficiency, is to perform uniform scalar quantization of residuals in a close loop of predictive coding. The majority of near-lossless image compression techniques belong to the category of L_∞-constrained predictive coding, e.g. the schemes by Chen and Ramabadran [21] and by Ke and Marcellin [28], and near-lossless CALIC.

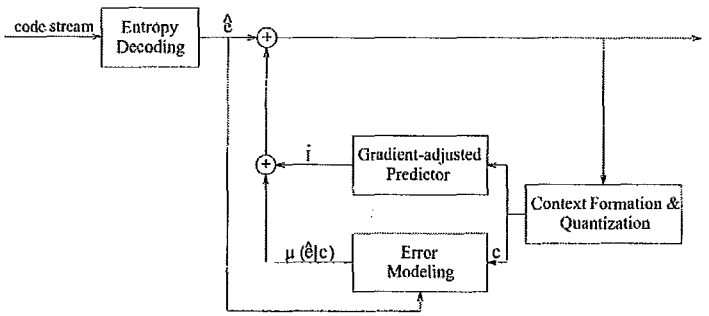
2.2 Near-Lossless CALIC

A schematic description of L_∞-constrained CALIC is given in Fig. 2.1. the system has five integrated components: gradient-adjusted prediction (GAP), context formation and quantization, prediction residue quantization, context modeling of quantized residues, and entropy coding of quantized residues. For any image $\mathbf{I} \in \mathbb{Z}^{M\times N}$, if the L_∞ bound we are given for near-lossless CALIC is τ, the encoding process is the following [6]:

First we estimate a predictor $\bar{I}$ of the next pixel. In Fig. 2.2, we list the pixel X which is currently being encoded, and its previous pixels: NN, NNE, NW, WW,



(a) ENCODER



(b) DECODER

Figure 2.1: Schematic description of L_∞-constrained CALIC

NE, W, N. To this end the following quantities are computed

$$\begin{aligned} d_h &= |\hat{I}_W - \hat{I}_{NW}| + |\hat{I}_N - \hat{I}_{NW}| + |\hat{I}_{NE} - I_N| \\ d_v &= |\hat{I}_W - \hat{I}_{NW}| + |\hat{I}_N - \hat{I}_{NN}| + |\hat{I}_{NE} - \hat{I}_{NNE}| \end{aligned} \tag{2.1}$$

where $\hat{I}_W, \hat{I}_N, \hat{I}_{WW}, \hat{I}_{NW}, \hat{I}_{NN}, \hat{I}_{NE}, \hat{I}_{NNE}$ are the reconstructed pixel values at position W, N, WW, NW, NN, NE, NNE respectively.

After the calculation in Eq. 2.1, the prediction scheme is explained as follows: If $d_h >> d_v$, which means there are a large amount of horizontal variations, N is picked as the initial prediction of I_X: $\bar{I} = I_N$

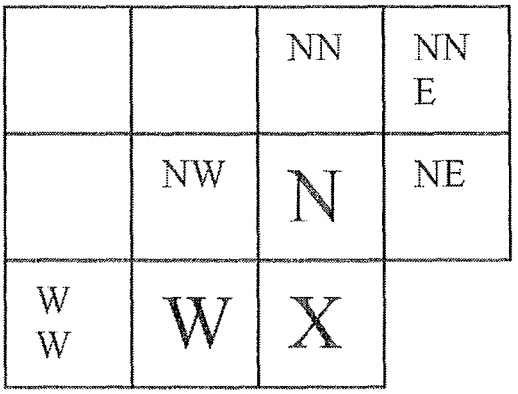


Figure 2.2: CALIC prediction table

If $d_v >> d_h$, which means there are a large amount of vertical variations, W is picked as the initial prediction of I_X: $\bar{I} = I_W$

If the differences are moderate or small, which means there is not any obvious variation in this context, the initial prediction value is a weighted average of neighboring pixels.

The reason why the prediction process is designed in this way is because a given pixel generally has a value close to one of its neighbors in an image, while which neighbor has the closest value depends on the local structure of the image. Next, to reduce the possible prediction biases in different contexts, we adjust the prediction value $\bar{I}$:

$$\tilde{I} = \bar{I} + \mu(\hat{e}|c) \tag{2.2}$$

where $\mu(\hat{e}|c)$ is the conditional sample mean of the quantized residues in the current modeling context c and $\hat{e}$ is the quantized prediction residue which will be clarified shortly. Now we have a context-based adaptive nonlinear prediction value $\tilde{I}$.

Then the prediction residue $e = I_X - \tilde{I}$ can be uniformly quantized with the quantization step size $2\tau+1$ and the reproduction value is the midpoint. The mapping

is given in Eq. 2.3.

$$\hat{e} = \begin{cases} (2\tau+1)\lfloor (e+\tau)/(2\tau+1) \rfloor & if \quad e > 0 \\ (2\tau+1)\lfloor (e-\tau)/(2\tau+1) \rfloor & if \quad e \leq 0 \end{cases} \tag{2.3}$$

Finally, at the last step the quantized prediction residue $\hat{e}$ is sent to the decoder after the entropy coding. While the decoder is just the reverse process of the encoder, we could have the reconstructed image $\hat{\mathbf{I}} = \tilde{\mathbf{I}} + \hat{\mathbf{e}}$, where $\hat{\mathbf{I}}, \tilde{\mathbf{I}}, \hat{\mathbf{e}} \in \mathbb{Z}^{M \times N}$ are the decompressed image, prediction image and quantized prediction residue image respectively.

From Eq. 2.3, it is obvious that $||\mathbf{e} - \hat{\mathbf{e}}||_\infty \leq \tau$, then the L_∞ bound for the near-lossless CALIC is

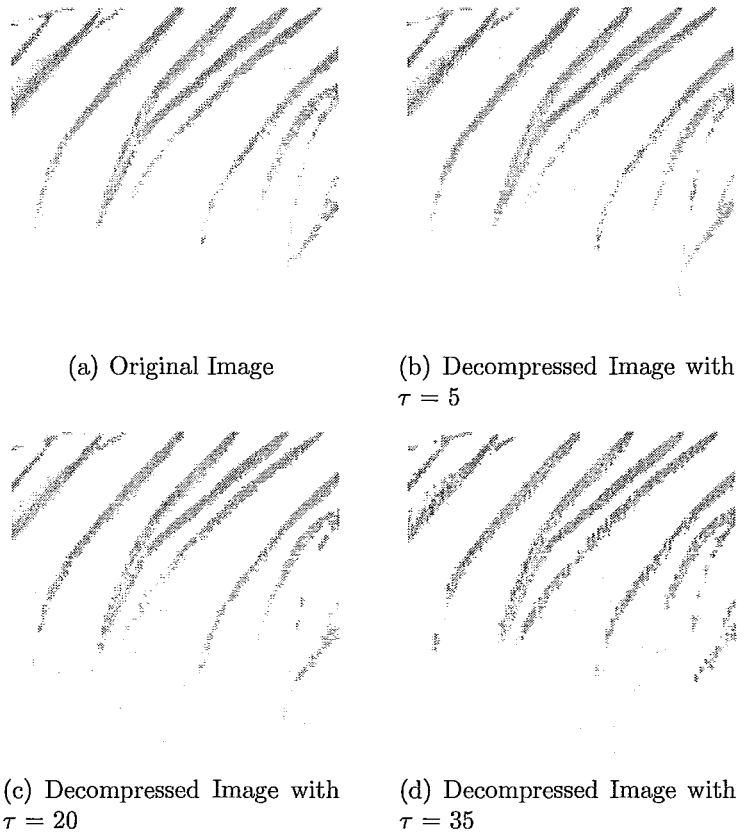
$$\begin{aligned} ||\hat{\mathbf{I}} - \mathbf{I}||_\infty &= ||\tilde{\mathbf{I}} + \hat{\mathbf{e}} - \mathbf{I}||_\infty \\ &= ||\hat{\mathbf{e}} - \mathbf{e}||_\infty \\ &\leq \tau \end{aligned} \tag{2.4}$$

In this way we guarantee the L_∞ bound for the near-lossless CALIC is τ.

2.3 Compression Noise

Fig. 2.3 lists the 8-bit graylevel decompressed images from near-lossless CALIC with the L_∞ bound τ being 5,20,35. We can find random scattering of black and white pixels all over the image, just like the salt and pepper noise [29]. These artifacts become more and more obvious with the increment of τ.

The compression artifacts, as shown in Fig. 2.3, appear as scattering of black



(a) Original Image

(b) Decompressed Image with $\tau = 5$

(c) Decompressed Image with $\tau = 20$

(d) Decompressed Image with $\tau = 35$

Figure 2.3: Compression artifacts in decompressed image with different L_∞ bounds

and white pixels randomly distributed over the image. In order to improve the image quality, a median filter is next applied. This is because the median filter is an effective method that can remove isolated noises without blurring sharp edges [30, 31], specifically, as shown in Fig. 2.4, the median filter replaces a pixel by the median of all pixels in the neighborhood:

$$Y_{i,j} = \text{Median} \quad \{X_{m,n} \quad (m,n) \in P_{i,j}\} \tag{2.5}$$

where $\mathbf{X} \in \mathbb{R}^{M \times N}$, $\mathbf{Y} \in \mathbb{R}^{M \times N}$, $\mathbf{X}, \mathbf{Y}$ are the input and output images. $P_{i,j}$ represents

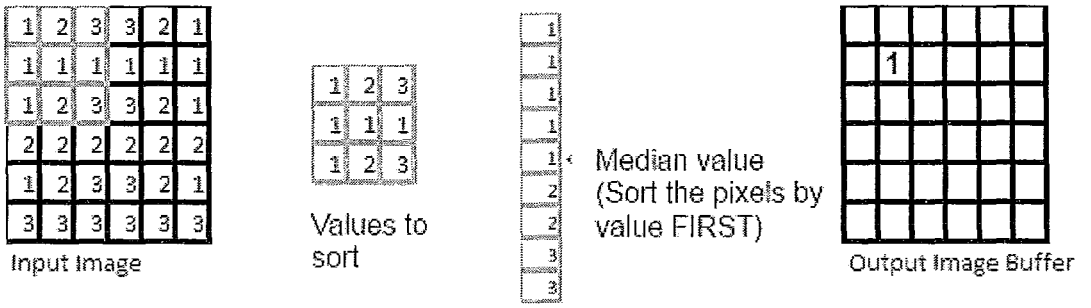


Figure 2.4: Median Mask

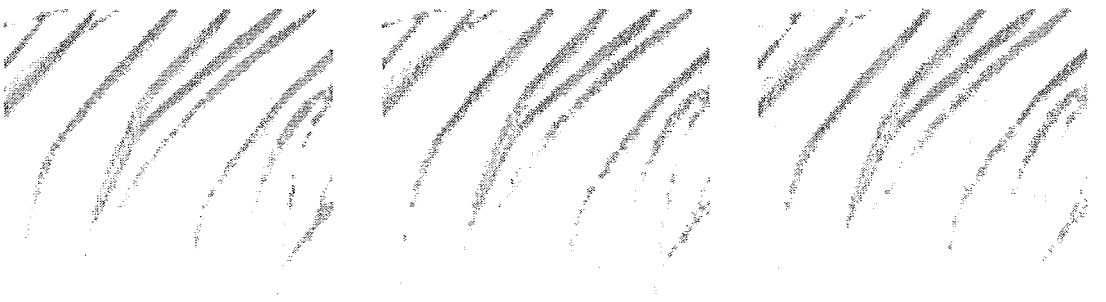
a neighborhood centered around location (i, j) in the image.

In Fig. 2.5, it can be easily found that the median filter suppresses most of the artifacts while preserving the edges effectively. Also, the PSNR values between the decompressed images and the original images have been greatly increased after the effect of median filter. For instance, for Fig. 2.5 (a), the PSNR value is 27.09 dB between the original image and the near-lossless CALIC decompressed image, the median filter improves the result up to 27.89 dB, which is 0.8 dB gain. Thus the median filter is very effective in removing the salt and pepper noise resulting from the Near-Lossless CALIC.

(a) Original Image

(b) Decompressed image with artifacts PSNR = 27.09 dB

(c) After median filter PSNR = 27.89 dB



(d) Original Image

(e) Decompressed image with artifacts PSNR = 28.34 dB

(f) After median filter PSNR = 28.89 dB

Figure 2.5: Median filter effect

Chapter 3

Piecewise Autoregressive Image Model

We need an appropriate mathematical model to fit the image for our proposed soft-decision decoding algorithm, which can be formulated as an inverse problem in (1.1). In this chapter, we introduce and discuss the piecewise autoregressive (PAR) image model, which can be learned as the prior domain knowledge from the image. Then we present a new adaptive weighted least square solution to learn the PAR image model.

3.1 Piecewise Autoregressive Image Model

From 1990s, image modeling has been a challenging research topic in image processing areas. For both natural and computer synthesized images, the structure of local image waveform varies spatially over the image, which results in the non-stationarity of the second-order statistics of image signal. Therefore, modeling of the non-stationary

image waveform needs to be highly adaptive to the varying local pixel structures. At the end of 1990s, Wu *et al.* had a measured success in the research on predictive lossless image coding [32]. In that work, on the assumption of piecewise stationarity of image signals in a local window, the image signal is modeled as a piecewise 2D autoregressive (PAR) process. The model parameters are adaptively estimated from pixel samples of a moving local window (the stationary area) on a pixel-by-pixel basis using the least square approach across the image. This assumption of PAR image model has been widely used in [33, 34, 35, 36, 37].

For any image $\mathbf{X} \in \mathbb{Z}^{M \times N}$, we use (i, j) to represent the pixel location. The PAR image model is

$$X_{i,j} = \sum_{(m,n) \in S_{i,j}} \alpha_{m,n}^{(i,j)} X_{i+m,j+n} + n_{i,j} \tag{3.1}$$

where $X_{i,j}$ is predicted by its neighbor pixels. $\boldsymbol{\alpha}^{(i,j)}$ is the PAR parameter vector for pixel position i, j, $\boldsymbol{\alpha}^{(i,j)} \in R^{|S_{i,j}| \times 1}$, $\boldsymbol{\alpha}^{(i,j)} = \{\alpha_{m,n}^{(i,j)}, \quad (m,n) \in S_{i,j}\}$. $S_{i,j}$ specifies the support window of the PAR model centered at pixel position (i, j). $n_{i,j}$ is a random perturbation independent of spatial location (i, j) and the image signal, and it accounts for both fractal like fine details of image and measurement noise. The validity of the PAR model depends on a mechanism that adjusts the model parameters $\boldsymbol{\alpha}^{(i,j)}$.

Here, the model parameter vector $\boldsymbol{\alpha}^{i,j}$ can change from pixel to pixel. The PAR model is chosen for the following reasons. Firstly, MRF is a common image model that has been proven effective in a wide range of applications. Secondly, the PAR model, while being a special form of MRF, has a generality afforded by the property that the Gauss-Markov process of the form $X_i = -\sum_{t=1}^{k} a_t X_{i-t} + n_i$, where n_i's are i.i.d. $\sim \mathcal{N}(0, \sigma^2)$, is the maximum entropy rate stochastic process, if a_t's are

chosen such that $\mathbb{E}\{X_i X_{i+t}\} = a_t$, for $1 \leq t \leq k$ and all i. Thus, the PAR model class is versatile, expressive, and capable of fitting image waveforms ranging from smooth shades, periodic textures to transients like edges. Thirdly, the piecewise linearity of the PAR model makes the soft-decision decoding problem convex, and hence, computationally amenable.

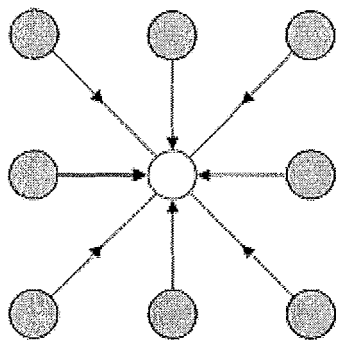


Figure 3.1: 8-order AR parameters

To keep the stationarity of the PAR parameters, one can expect that the size of PAR support window should not be too large. Thus we only consider the 8-connected neighbors of a PAR model, in other words $|S_{i,j}| = 8$. And as shown in Fig. 3.1, we define a pixel vector $\mathbf{X}^8_{i,j}$ as the eight disjoint neighbors of $X_{i,j}$,

$$\mathbf{X}^8_{(i,j)} = (X_{i-1,j-1}, X_{i-1,j}, X_{i-1,j+1}, X_{i,j-1}, X_{i,j+1}, X_{i+1,j-1}, X_{i+1,j}, X_{i+1,j+1})^T \quad (3.2)$$

Using the new-defined notations, we rewrite Eq. 3.1 as

$$X_{i,j} = \sum_{0 \leq t \leq 7} \alpha_t^{(i,j)} X^8_{(i,j)\diamond t} + n_{i,j} \quad (3.3)$$

where $\alpha_t^{(i,j)}$ denotes the $(t+1)$-th component of PAR parameter vector $\alpha^{(i,j)}$, $X^8_{(i,j)\diamond t}$

denotes the $(t+1)$-th component of vector $\mathbf{X}^8_{(i,j)}$,

Mathematically, the structure of image signals can be learnt by fitting the PAR model to pixels in another local window $W_{i,j}$. In other words, we assume the statistical stationarity of the image signal in window $W_{i,j}$. To best fit the structure of the image in the L_2 sense, the PAR model parameter vectors $\boldsymbol{\alpha}^{i,j} = (\alpha_0^{(i,j)}, \alpha_1^{(i,j)}, ..., \alpha_7^{(i,j)}$ should satisfy:

$$\boldsymbol{\alpha}_{i,j} = \arg\min_{\boldsymbol{\alpha}} \{ \sum_{(m,n)\in W_{i,j}} (X_{m,n} - \sum_{0\le t\le 7} \alpha_t X^8_{(m,n)\diamond t})^2 \} \tag{3.4}$$

where the minimization is over all 8 dimensional vector $\boldsymbol{\alpha} = (\alpha_0, \alpha_1, ..., \alpha_7) \in \mathbb{R}^8$.

The minimization problem in Eq. (3.4) can be written as:

$$\boldsymbol{\alpha}_{i,j} = \arg\min_{\boldsymbol{\alpha}} \|A\boldsymbol{\alpha} - \mathbf{Y}\|_2^2 \tag{3.5}$$

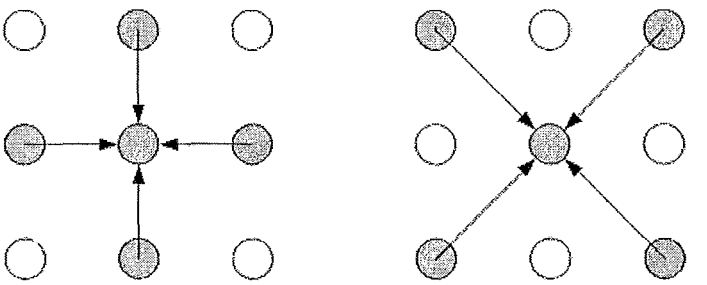
where $\mathbf{Y} \in \mathbb{R}^{|W_{i,j}|\times 1}$, whose elements are $X_{m,n}$ with $(m,n) \in W_{i,j}$. The matrix $A \in \mathbb{R}^{|W_{i,j}|\times 8}$, and the i-th row of matrix A contains the 8-neighbors of pixel Y_i. The least square problem (3.5) has a closed-form solution as:

$$\boldsymbol{\alpha}^{(i,j)} = (A^T A)^{-1} A^T \mathbf{Y} \tag{3.6}$$

Despite of its widely usage, the PAR model suffers two shortcomings when applied in the problem discussed in the thesis. First, if the size of the moving window is kept small to satisfy the local stationarity requirement, then the number of training samples is limited, and the PAR model estimated will face the dilemma of model overfitting. Mathematically, to solve an array of equations, the number of independent observations is required to meet the number of unknown variables. In other words, the number of observed pixels needs to be large enough for a robust estimate of the

PAR model. However, owing to the piecewise stationarity nature of image signals, a 2D PAR model holds only within a small local window and therefore insufficient observation data can be provided. In [36], the limitation of piecewise stationarity of image signals against the minimum required number of observations is well balanced by choosing a low-order PAR model and a moderate-sized window. Second, the moving window based PAR parameters estimation is problematic if the window contains pixels of different stochastic properties. For example, the local stationarity assumption is obviously violated if the window is acrossed by an edge separating two different areas. We will propose our solutions to the shortcomings in the following two sections.

3.2 Overfitting Problem



(a) axial direction (b) diagonal direction

Figure 3.2: PAR models in two different directions

In order to avoid the potential pitfall of model overfitting, our solution is to increase the number of PAR model constraint equations, so we use two separate PAR models of order 4 rather than a unified model of order 8. The 4-connected neighbors of a pixel $X_{i,j}$ in the axial direction are labeled $X^{+}_{(i,j)\diamond t}$, t=0,1,2,3; similarly, the

4-connected neighbors of a pixel $X_{i,j}$ in the diagonal direction are labeled $X^{\times}_{(i,j)\diamond t}$, t=0,1,2,3, as shown in Fig. 3.2, the two models act on two disjoint sets of neighbors of $X_{i,j}$,

$$\begin{aligned} \mathbf{X}^{+}_{(i,j)} &= (X_{i,j-1}, X_{i-1,j}, X_{i,j+1}, X_{i+1,j})^T \\ \mathbf{X}^{\times}_{(i,j)} &= (X_{i-1,j-1}, X_{i-1,j+1}, X_{i+1,j+1}, X_{i+1,j-1})^T \end{aligned} \tag{3.7}$$

If we assume that $\mathbf{a}^{(i,j)} = (a_0^{(i,j)}, a_1^{(i,j)}, a_2^{(i,j)}, a_3^{(i,j)})^T$ and $\mathbf{b}^{(i,j)} = (b_0^{(i,j)}, b_1^{(i,j)}, b_2^{(i,j)}, b_3^{(i,j)})^T$ are the vectors of PAR coefficients in the axial and diagonal directions respectively, then this fitting process is formulated by solving the following two least-squares problems:

$$\begin{aligned} \mathbf{a}^{(i,j)} &= \arg \min_{\mathbf{a}\in\mathbb{R}^{4\times 1}} \Big\{ \sum_{(m,n)\in W_{i,j}} (X_{m,n} - \sum_{0\le t\le 3} a_t X^{+}_{(m,n)\diamond t})^2 \Big\} \\ \mathbf{b}^{(i,j)} &= \arg \min_{\mathbf{b}\in\mathbb{R}^{4\times 1}} \Big\{ \sum_{(m,n)\in W_{i,j}} (X_{m,n} - \sum_{0\le t\le 3} b_t X^{\times}_{(m,n)\diamond t})^2 \Big\} \end{aligned} \tag{3.8}$$

where the first minimization is over all $\mathbf{a} = (a_0, a_1, a_2, a_3)^T$ and the second is over all $\mathbf{b} = (b_0, b_1, b_2, b_3)^T$.

Further, we can rewrite the above problem as

$$\begin{aligned} \mathbf{a}^{(i,j)} &= \arg \min_{\mathbf{a}\in\mathbb{R}^{4\times 1}} ||A^{+}\mathbf{a} - \mathbf{Y}||_2^2 \\ \mathbf{b}^{(i,j)} &= \arg \min_{\mathbf{b}\in\mathbb{R}^{4\times 1}} ||A^{\times}\mathbf{b} - \mathbf{Y}||_2^2 \end{aligned} \tag{3.9}$$

where $\mathbf{Y}$ is composed of the pixels in the moving window $W_{i,j}$, and $Y \in \mathbb{R}^{|W_{i,j}|\times 1}$. The matrix $A^{+} \in \mathbb{R}^{|W_{i,j}|\times 4}$, and the i-th row of A^{+} contains the 4-axial-neighbors of

pixel Y_i; similarity, $A^{\times} \in \mathbb{R}^{|W_{i,j}| \times 4}$, and the i-th row of $A^{\times}$ contains the 4-diagonal-neighbors of pixel Y_i. Thus, the closed form solution for the two-directional PAR model parameters $\mathbf{a}$ and $\mathbf{b}$ are:

$$\begin{aligned} \mathbf{a}^{(i,j)} &= (A^{+T}A^{+})^{-1}A^{+T}\mathbf{Y} \\ \mathbf{b}^{(i,j)} &= (A^{\times T}A^{\times})^{-1}A^{\times T}\mathbf{Y} \end{aligned} \tag{3.10}$$

3.3 Weighted PAR model (WPAR)

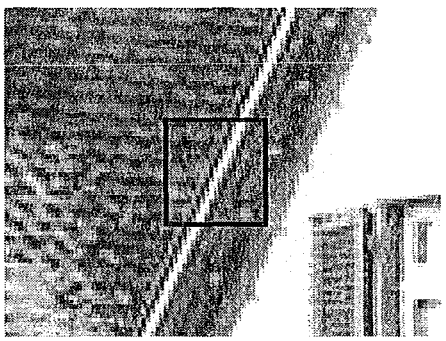


Figure 3.3: Example of non-stationary window

As shown in Fig. 3.3, it is obviously not reasonable to assume PAR model being stationary for the windows across the edge along the roof. For each pixel $X_{m,n}$ in the window $W_{i,j}$, we consider a 3×3 patch centered at $X_{i,j}$ as its feature vector. If we denote by $\mathbf{E}^k$ the feature vector of the $(k+1)$-th pixel in the window $W_{i,j}$, where $0 \leq k < |W_{i,j}|$, and by $\mathbf{E}^p$ the feature vector for the current pixel we are processing, the distance between any two feature vectors in the local window is defined as

$$D^{p,q} = ||\mathbf{E}^p - \mathbf{E}^q||_2 \tag{3.11}$$

where $0 \leq p < |W_{i,j}|, 0 \leq q < |W_{i,j}|$.

If $0 \leq k < |W_{i,j}|$, we consider the weight for the $(k+1)$-th element in the local window $W_{i,j}$ as:

$$w^{p,k} = e^{-\frac{D^{p,k}}{2}} \tag{3.12}$$

By defining the diagonal weighting matrix $w \in \mathbb{R}^{|W_{i,j}| \times |W_{i,j}|}$,

and $w = diag\{w^{0,p}, ..., w^{p-1,p}, 1, w^{p+1,p}, ..., w^{|W|-1,p}\}$.

we can rewrite the least-square problem (3.9) by adding the weighting matrix,

$$\begin{aligned} \mathbf{a}^{(i,j)} &= \arg \min_{\mathbf{a} \in \mathbb{R}^{4\times 1}} \|wA^{+}\mathbf{a} - w\mathbf{Y}\|_2^2 \\ \mathbf{b}^{(i,j)} &= \arg \min_{\mathbf{b} \in \mathbb{R}^{4\times 1}} \|wA^{\times}\mathbf{b} - w\mathbf{Y}\|_2^2 \end{aligned} \tag{3.13}$$

where A^{+},$A^{\times}$, and $\mathbf{Y}$ have been defined in (3.9). The closed form solution for the WPAR model parameters $\mathbf{a}^{(i,j)}, \mathbf{b}^{(i,j)}$ is

$$\begin{aligned} \mathbf{a}^{(i,j)} &= ({A^{+}}^T w^T w A^{+})^{-1} {A^{+}}^T w^T w\mathbf{Y} \\ \mathbf{b}^{(i,j)} &= ({A^{\times}}^T w^T w A^{\times})^{-1} {A^{\times}}^T w^T w\mathbf{Y} \end{aligned} \tag{3.14}$$

In Eq. (3.14), we present the mathematical solutions to the two 4-order WPAR parameters $\mathbf{a}^{(i,j)}$ and $\mathbf{b}^{(i,j)}$ for near-lossless CALIC decompressed image.

Chapter 4

Soft-Decision Decoding Algorithm with PAR modeling

After learning the PAR image model, coupled with the side information exploited from the L_∞-constrained near-lossless compression algorithm, we adopt a soft-decision decoding approach, starting from the decompressed image $\hat{\mathbf{I}}$ as an initial estimate of $\mathbf{I}$, and proceed to compute an improved estimate $\mathbf{X}$ of the original image $\mathbf{I}$ by solving a constrained convex optimization problem. We also interpret the power of the L_∞ constraints in terms of reducing the L_2 distortions in a mathematical way. Then we present and discuss our experiments results.

4.1 Formulation of the Optimization Problem

Assume $\mathbf{I}$ is the original image, $\hat{\mathbf{I}}$ is the decompressed image after near-lossless CALIC. If we assume the PAR model would keep consistent between $\mathbf{I}$ and $\hat{\mathbf{I}}$, then the image restoration can be accomplished by the approach that reestimates the decompressed

image reconstructed by the codec based on the PAR image model, in other words, it takes the decompressed image $\hat{\mathbf{I}}$ as an intermediate result and computes an improved estimate $\mathbf{X}$ of the original image $\mathbf{I}$.

The re-estimation method proceeds in two steps: it first estimates the PAR model parameters adaptively for each spatial location (i, j), using samples of a moving window $W_{i,j}$ in image $\hat{\mathbf{I}}$; then it improves the decoded result from $\hat{\mathbf{I}}$ to $\mathbf{X}$ by solving the following linear least-squares estimation problem:

$$\begin{aligned}\min_{\mathbf{X}}\{&\xi^{+}_{i,j}\sum_{(m,n)\in W_{i,j}}[X_{m,n}-\sum_{0\le\ t\le 3}a_t^{(m,n)}X^{+}_{(m,n)\diamond t}]^2\\&+\xi^{\times}_{m,n}\sum_{(m,n)\in W_{i,j}}[X_{m,n}-\sum_{0\le\ t\le 3}b_t^{(m,n)}X^{\times}_{(m,n)\diamond t}]^2\\&+\lambda\sum_{(m,n)\in W_{i,j}}[X_{m,n}-\hat{I}_{m,n}]^2\}\end{aligned} \tag{4.1}$$

where $\mathbf{X}\in\mathbb{R}^{|W_{i,j}|\times 1}$, the elements of $\mathbf{X}$ are from the pixels of local window $W_{i,j}$. $\mathbf{X}^{+}_{(i,j)}$ and $\mathbf{X}^{\times}_{(i,j)}$ represent the 4-order neighbor pixels of $X_{i,j}$, which are defined in (3.7). $\mathbf{a}^{(m,n)}=(a_0^{(m,n)},a_1^{(m,n)},a_2^{(m,n)},a_3^{(m,n)})$ and $\mathbf{b}^{(m,n)}=(b_0^{(m,n)},b_1^{(m,n)},b_2^{(m,n)},b_3^{(m,n)})$ are the vectors of AR coefficients defined in (3.10), for corresponding pixel (m,n). The corresponding PAR models are learnt from the decompressed image. $\xi^{+}_{i,j}$ and $\xi^{\times}_{i,j}$ are the two corresponding least squares weights to be clarified shortly, and the parameter λ should be chosen according to the bit rate. The first two terms in the cost function (4.1) describe the fitting error for the PAR models, and the third term behaves as the fidelity item with respect to the decompressed image $\hat{\mathbf{I}}$ in L_2 sense.

The underlying weights $\xi_{i,j}^{+}$ and $\xi_{i,j}^{\times}$ can be set as:

$$\xi_{i,j}^{+} = \frac{e_{i,j}^{\times}}{e_{i,j}^{\times} + e_{i,j}^{+}}; \qquad \xi_{i,j}^{\times} = \frac{e_{i,j}^{+}}{e_{i,j}^{\times} + e_{i,j}^{+}} \tag{4.2}$$

where $e_{i,j}^{+}$ and $e_{i,j}^{\times}$ are the squared errors associated with the solutions of the two 4-order PAR model, with local window centered in position (i,j),which are the optimal weights if the two PAR models are independent.

$$\begin{aligned} e_{i,j}^{+} &= \sum_{(m,n)\in W_{i,j}} [\hat{I}_{m,n} - \sum_{0\le t\le 3} a_t^{(m,n)} \hat{I}_{(m,n)\Diamond t}^{+}]^2 \\ e_{i,j}^{\times} &= \sum_{(m,n)\in W_{i,j}} [\hat{I}_{m,n} - \sum_{0\le t\le 3} b_t^{(m,n)} \hat{I}_{(m,n)\Diamond t}^{\times}]^2 \end{aligned} \tag{4.3}$$

L_∞-constrained near-lossless image coding techniques inherently have a strong form of side information that can be exploited by the decoder to greatly reduce the L_2 distortion of a decompressed image. From (2.4), we know for near-lossless CALIC that the L_∞ constraint is imposed on the compressor such that the compression error is tightly bounded for each single pixel, i.e., $\|\mathbf{I} - \hat{\mathbf{I}}\|_\infty \le \tau$, where τ is called the L_∞ bound for the compression. It is reasonable to expect that we can reduce the L_2 distortion of a decompressed image with assistance of the known tight error bound on each pixel, which we will discuss in the next section. In this way, we can rewrite

(4.1) as

$$\begin{aligned}\min_{\mathbf{X}}\{&\xi^{+}\sum_{(m,n)\in W_{i,j}}[X_{m,n}-\sum_{0\le t\le 3}a_t^{(m,n)}X^{+}_{(i,j)\Diamond t}]^2\\&+\xi^{\times}\sum_{(m,n)\in W_{i,j}}[X_{m,n}-\sum_{0\le t\le 3}b_t^{(m,n)}X^{\times}_{(m,n)\Diamond t}]^2\\&+\lambda\sum_{(m,n)\in W_{i,j}}[X_{m,n}-\hat{I}_{m,n}]^2\}\\\text{subject to}\quad&|X_{m,n}-\hat{I}_{m,n}|\le\tau,\quad(m,n)\in W_{i,j}\end{aligned}\tag{4.4}$$

where τ is the L_∞ bound for the codec. $\mathbf{X}, \mathbf{X}^{+}_{(i,j)}, \mathbf{X}^{\times}_{(i,j)}, \mathbf{a}, \mathbf{b}, \xi^{+}, \xi^{\times}, \lambda$ are defined in the same way as that of (4.1).

We can rewrite Eq. (4.4) as

$$\begin{aligned}&\mathbf{X}=\arg\min_{\mathbf{X}}\|C\mathbf{X}-\mathbf{d}\|^2\\&\text{subject to}\quad\|\mathbf{X}-\hat{\mathbf{I}}\|_\infty\le\tau\end{aligned}\tag{4.5}$$

where C and $\mathbf{d}$ are composed of $\mathbf{a}$, $\mathbf{b}$, λ, and the decoded image $\hat{\mathbf{I}}$.

Due to the use of constraints in (4.1) and the L_∞ bound in near-lossless coding, we have both $|X_{i,j}-\hat{I}_{i,j}|\le\tau$ and $|I_{i,j}-\hat{I}_{i,j}|\le\tau$, and consequently

$$|X_{i,j}-I_{i,j}|\le 2\tau.\tag{4.6}$$

In other words, the proposed soft decoding technique can greatly reduce L_2 distortion and at the same time still maintain a known bound on the maximum possible error.

Moreover, for every pixel $I_{i,j}$ we obtain an even tighter error bound:

$$\begin{aligned} |X_{i,j} - I_{i,j}| &= |\hat{I}_{i,j} - I_{i,j} + X_{i,j} - \hat{I}_{i,j}| \\ &\leq \tau + |X_{i,j} - \hat{I}_{i,j}| \end{aligned} \tag{4.7}$$

which can be very important information for critical applications in medicine, security, remote sensing, reconnaissance, etc., where decompressed images are subject to rigorous computer analysis, not just to please the eyes as in entertainment and consumer applications.

It is also interesting to note that a proper shrinkage of the L_∞ constraint in the optimization problem of (4.5), namely

$$\|\mathbf{X} - \hat{\mathbf{I}}\|_\infty = \beta\tau, \quad 0 \leq \beta \leq 1 \tag{4.8}$$

can reduce both the L_2 and L_∞ errors of $\mathbf{X}$, i.e., achieve higher PSNR value and tighter maximum error bound (replacing 2τ with a smaller value $(1+\beta)\tau$ in (4.6)).

The proposed soft decoding process can be carried out iteratively, if one replaces the intermediate image $\hat{\mathbf{I}}$ by the newly estimated image $\mathbf{X}$. In order to ensure maximum error bound, the constraints in (4.5) should not change in successive iterations.

4.2 Mean Squared Error (MSE) Analysis

In this section, we conduct the MSE analysis for our proposed soft-decision decoding algorithm presented in the previous section. This provides a way to better interpret the power of the L_∞ constraint in reducing the MSE, or equivalently, achieving higher PSNR value than the case where this strong side information is wasted.

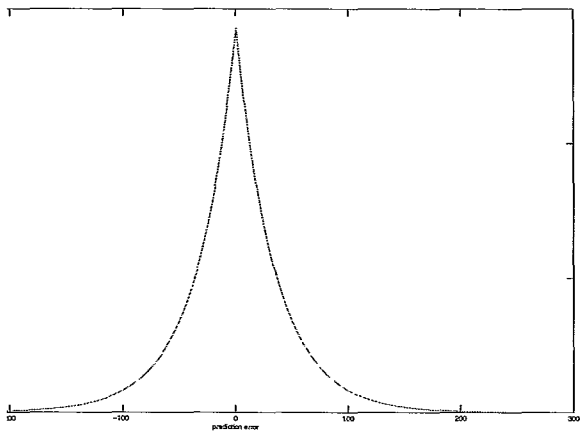


Figure 4.1: CALIC Error Prediction Distribution

To prevent notations clutter, single index is used to label 2D pixel locations. More specifically, let $\mathbf{I}^P = [I_1^P, I_2^P, \cdots, I_N^P]^T$, and $\mathbf{e} = [e_1, e_2, \cdots, e_N]^T$ be the predicted image and the residue image, respectively, where N is the number of pixels in the image. Obviously, $\mathbf{e} = \mathbf{I} - \mathbf{I}^P$. Let e denote the error random variable taking the values $e_1, e_2, ..., e_N$. Based on practical observation, it is assumed that e, follows the Laplacian distribution (shown in Fig. 4.1)

$$f_e(x) = \frac{1}{2b} e^{\frac{-|x|}{b}} \tag{4.9}$$

where $b > 0$. On the other hand, for the L_∞-decoded image $\hat{\mathbf{I}} = [\hat{I}_1, \hat{I}_2, \cdots, \hat{I}_N]^T$, we have $\hat{I}_i = I_i^P + Q(e_i)$, where the quantization function $Q(\cdot)$ is defined as

$$Q(x) \triangleq \sum_{j=1}^{M} \hat{e}_j \mathbf{1}_{\mathcal{S}_j}(x) \tag{4.10}$$

with $M \geq 0$, $\hat{e}_1, ..., \hat{e}_M \in \mathbb{Z}$, and

$$\mathcal{S}_j = \{e | \hat{e}_j - \tau < e \leq \hat{e}_j + \tau\} \tag{4.11}$$

being the jth quantization region and $\mathbf{1}_{\mathcal{A}}(x)$ being the indicator function returning 1 if $x \in \mathcal{A}$, and 0 otherwise.

Let $\breve{\mathbf{e}} = [\breve{e}_1, \breve{e}_2, \cdots, \breve{e}_N]^T = \mathbf{X} - \mathbf{I}^P$ be the residue image associated with the improved estimate $\mathbf{X}$. Let $\breve{e}$ denote the random variable taking values $\breve{e}_1, \breve{e}_2, ..., \breve{e}_N$

Denote by $f_{\breve{e}}(x|e \in \mathcal{S}_j)$ the probability density function (p.d.f) of $\breve{e}$ conditioned on $e \in \mathcal{S}_j$, and by $f_{\breve{e},e}(x, y|e \in \mathcal{S}_j)$ the joint p.d.f of $\breve{e}$ and e conditioned on $e \in \mathcal{S}_j$, respectively.

The MSE of $\mathbf{X}$ with respect to $\mathbf{I}$ can then be calculated as

$$\begin{aligned} \text{MSE}(\mathbf{X}) &= (\mathbf{I} - \mathbf{X})^T(\mathbf{I} - \mathbf{X}) \\ &= (\mathbf{e} - \breve{\mathbf{e}})^T(\mathbf{e} - \breve{\mathbf{e}}) \\ &= \sum_{i=1}^{N} e_i^2 - 2\sum_{i=1}^{N}(e_i\breve{e}_i) + \sum_{i=1}^{N} \breve{e}_i^2 \end{aligned} \tag{4.12}$$

As assumed above, all the random variables e_i's follow the same Laplacian distribution given in (4.9), and hence,

$$\begin{aligned} \sum_{i=1}^{N} e_i^2 \approx & N\mathbb{E}[e^2] \\ = & 2Nb^2 \end{aligned} \tag{4.13}$$

For the second term of (4.12), we have

$$\begin{aligned}\sum_{i=1}^{N}(e_i\check{e}_i) &\approx N\mathbb{E}[e\check{e}] \\ &= N\sum_{j=1}^{M}P(e\in\mathcal{S}_j)\mathbb{E}[e\check{e}|e\in\mathcal{S}_j] \\ &= N\sum_{j=1}^{M}P_j\int_{-\infty}^{\infty}\int_{\mathcal{S}_j}xyf_{\check{e},e}(x,y|x\in\mathcal{S}_j)dxdy \qquad (4.14)\end{aligned}$$

where $P_j = P(e \in \mathcal{S}_j) = \int_{\mathcal{S}_j} f_e(x)dx$.

Similarly, we get

$$\begin{aligned}\sum_{i=1}^{N}(\check{e}_i^2) &\approx N\mathbb{E}[\check{e}] \\ &= \sum_{j=1}^{M}P_j\int_{-\infty}^{\infty}x^2 f_{\check{e}}(x|e\in\mathcal{S}_j)dx \qquad (4.15)\end{aligned}$$

Therefore,

$$\text{MSE}(\mathbf{X}) = 2Nb^2 - N\Big\{\sum_{j=1}^{M}P_j\int_{-\infty}^{\infty}\Big[2\int_{\mathcal{S}_j}xyf_{\check{e},e}(x,y|e\in\mathcal{S}_j)dy - x^2 f_{\check{e}}(x|e\in\mathcal{S}_j)\Big]dx\Big\} \qquad (4.16)$$

For the constrained optimization problem in (4.5), the conditional p.d.f $f_{\check{e}}(x|e \in \mathcal{S}_j)$ would take non-zero values only within $\mathcal{S}_j$, due to the L_∞ constraint imposed. In contrast, in the case that this constraint is ignored, $f_{\check{e}}(x|e \in \mathcal{S}_j)$ can possibly take non-zero values over the entire range $[-255, 255]$. This implies that the conditional

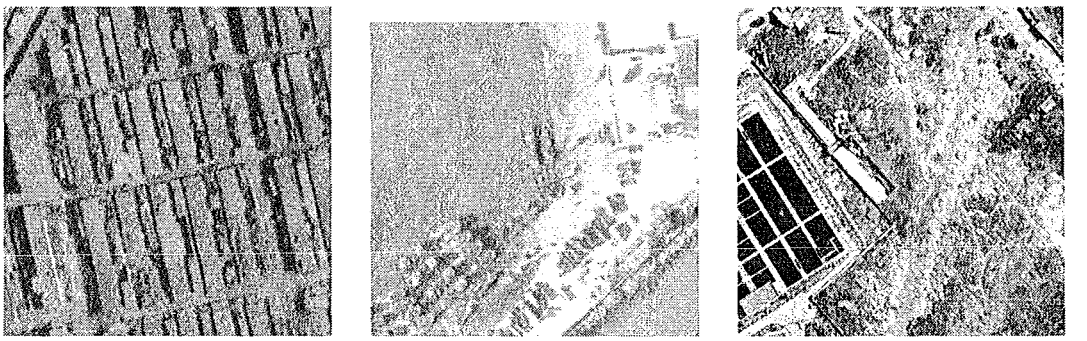
variance of $\breve{e}$ becomes smaller with the L_∞ constraint. In addition, the L_∞ constraint confining the solution space would potentially make the covariance between $\breve{e}$ and e larger than that of the case without this constraint. Based on these observations, we can find, from (4.16), that the MSE of $\mathbf{X}$ with respect to $\mathbf{I}$ can be reduced by utilizing the L_∞ constraint as strong side information in solving the inverse problem.

On the other hand, the above MSE analysis is also useful to guide us on selecting an appropriate shrinkage factor β in (4.8). When β is sufficiently close to 1, it is reasonable to assume that the conditional p.d.f $f_{\breve{e}}(x|e \in \mathcal{S}_j)$ and the conditional joint p.d.f $f_{\breve{e},e}(x, y|e \in \mathcal{S}_j)$ are truncated versions of their counterparts without shrinkage. The problem of finding an optimal β can then be converted to minimizing MSE($\mathbf{X}$) with respect to β. Empirically, we found that $\beta = 0.7$ improves the PSNR of $\mathbf{X}$, and at the same time, reduces the L_∞ error bound. One approach is that we can assume that $f_{\breve{e}}(x|e \in \mathcal{S}_j)$ and $f_{\breve{e},e}(x, y|e \in \mathcal{S}_j)$ are respectively truncated Gaussian distribution and truncated joint Gaussian distribution, whose parameters can be estimated from training data. However, the further discussion of optimal β is beyond the scope of this thesis, and hence, we omit the details here.

4.3 Experimental Results and Remarks

In this section, we present the experiment results to demonstrate the effectiveness of the asymmetric high-fidelity image compression algorithm. In our experiments, we compare the compression distortions in both L_2 and L_∞ sense for three algorithms: near-lossless CALIC, J2K, and the proposed soft-decision decoding technique initialized by near-lossless CALIC. Without loss of generality to the results, we selected

a large set of typical test images with different context and structure. Since near-lossless CALIC is mainly used in the applications which set stringent quality requirements on the distortion such as remote sensing, medicine, we choose the test images mainly from these areas. Our five sample images 'Satellite','Remote-Sea','Remote-Farm','MRI-Brian', and 'Lena'. are presented in Fig. 4.2.



(a) Satellite (b) Remote-Sea (c) Remote-Farm

(d) MRI-Brain (e) Face

Figure 4.2: Five Test Images in the test set

Tables 4.1 - 4.3 list the experimental results for these five images with $\tau = 1$, $\tau = 3$, and $\tau = 5$, respectively. The PSNR value and the maximum possible error are compared with the same bit rate for these three algorithms. G_1 is defined as the PSNR difference between our proposed method and near-lossless CALIC, while G_2 is defined as that between our proposed method and J2K. It can be seen that the PSNR

Table 4.1: Performance of Different Methods For $\tau = 1$.

Image	rate	CALIC		J2K		proposed			
		PSNR	$\|e_q\|_\infty$	PSNR	$\|e_q\|_\infty$	PSNR	$\|e_q\|_\infty$	G_1	G_2
Satellite	2.25	49.91	1	48.68	5	50.47	2	0.56	1.79
Remote-Sea	2.41	49.91	1	48.58	4	50.28	2	0.37	1.7
Remote-Farm	2.33	49.88	1	49.14	4	50.40	2	0.52	1.26
MRI-Brian	2.23	51.84	1	49.76	5	52.05	2	0.21	0.29
Lena	3.01	49.92	1	48.00	5	50.00	2	0.08	2.00

Table 4.2: Performance of Different Methods For $\tau = 3$

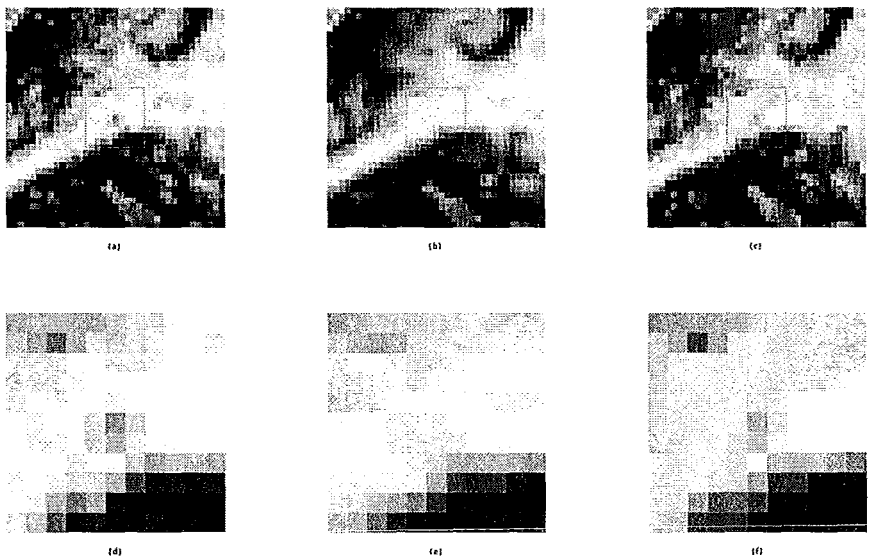
Image	rate	CALIC		J2K		proposed			
		PSNR	$\|e_q\|_\infty$	PSNR	$\|e_q\|_\infty$	PSNR	$\|e_q\|_\infty$	G_1	G_2
Satellite	1.42	42.32	3	43.44	9	43.97	6	1.65	0.53
Remote-Sea	1.55	42.40	3	43.67	9	43.75	6	1.35	0.08
Remote-Farm	1.43	42.39	3	44.24	9	44.24	6	1.85	0
MRI-Brian	1.58	44.15	3	44.20	12	44.35	6	0.2	0.15
Lena	1.96	42.17	3	42.55	9	44.79	6	0.62	0.24

Table 4.3: Performance of Different Methods For $\tau = 5$

Image	rate	CALIC		J2K		proposed			
		PSNR	$\|e_q\|_\infty$	PSNR	$\|e_q\|_\infty$	PSNR	$\|e_q\|_\infty$	G_1	G_2
Satellite	1.10	38.44	5	40.95	13	40.64	10	2.20	-0.31
Remote-Sea	1.18	36.76	5	40.04	13	40.20	10	3.44	0.16
Remote-Farm	1.08	38.83	5	41.82	13	41.10	10	2.27	-0.72
MRI-Brian	1.26	40.42	5	40.79	14	40.90	10	0.48	0.11
Lena	1.46	38.37	5	39.60	14	39.54	10	1.17	-0.06

gain against the near-lossless CALIC (G_1) can be up to 3.44 db, and the value of G_1 tends to increase for larger τ. Meanwhile, we still maintain a L_∞ bound of 2τ. On the other hand, compared with J2K, our proposed method achieves higher PSNR values when the bit rate is above 1.2 bpp, and the value of G_2 can be very significant in high rate region. For instance, for image 'Satellite', our approach achieves 0.53 db gain in PSNR with respect to J2K at bit rate 1.42 bpp, while this value is further improved to be 1.79 db at bit rate 2.25. In other words, in the relative high rate region, our proposed soft-decision decoding strategy equipped with an image prior and L_∞ side information can achieve the best of the both worlds: significantly higher compression ratio of lossy compression and a predetermined minmax fidelity that is close to lossless compression. When the bit rate becomes lower than 1.2 bpp, the PSNR performance of our proposed scheme begins to be inferior to J2K. However, it should be noted that in low bit rate, lossy image codecs designed under the L_2 criterion can incur large errors on some pixels that are statistical outliers, and hence, increases the L_∞ bound. Such large individual errors, although with negligible contribution to PSNR, can be disastrous in some critical scenarios, e.g., medical and security applications. In contrast, our proposed method can still guarantee a known tight bound of 2τ. This maximum error bound can be further reduced through using the shrinkage as mentioned in Section 4. Taking the example of the image 'Satellite', we can achieve 0.27 db gain when $\tau = 8$ and also reduce the maximum error bound to be 13, compared with the case without shrinkage.

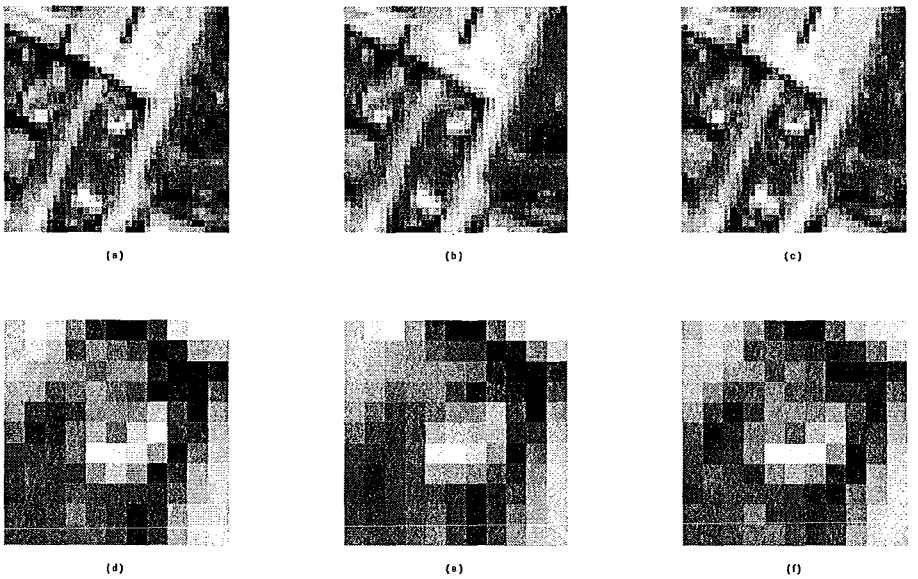
Furthermore, our proposed algorithm possess an additional advantage: it better preserves the original pixel patterns, even in the rate region below 1.2 bpp. To demonstrate this phenomenon, we in Fig. 4.3 show the decoded images using our



(a) Original Image (Center = 118), J2K(Center = 180), Proposed(Center = 134)

Figure 4.3: (a) Original image; (b) J2K decoded image; (c) Soft-decoded image; (d) Enlarged version of the red box of (a); (e) Enlarged version of the red box of (b);(f) Enlarged version of the red box of (c).

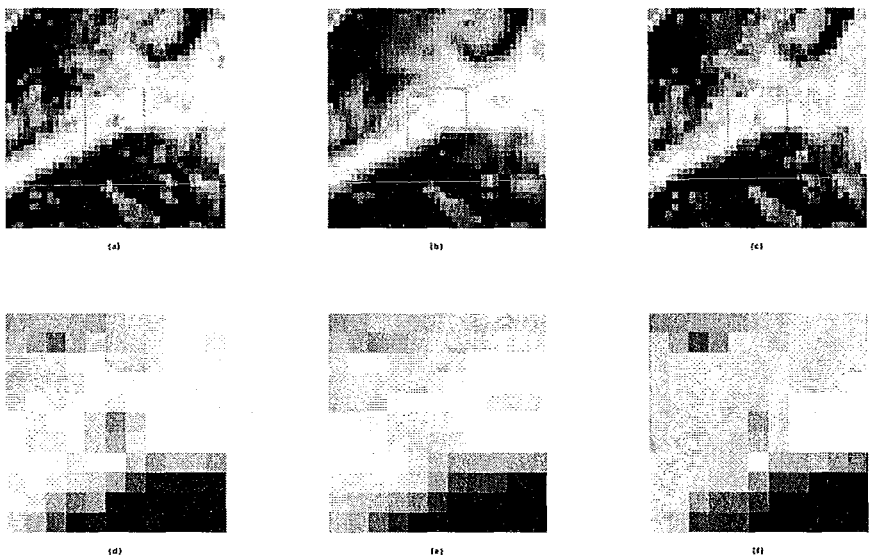
proposed method and J2K. The bit rate is 1.04 bpp, and J2K achieves 0.1 db gain in PSNR than our method. However, from Fig.4.3 (a)-(c), we can see that our method can better preserve the fine structures compared with J2K, even though our method is outperformed by J2K in PSNR at this bit rate. This property can be better verified by observing the blocks highlighted, which are enlarged in Fig.4.3 (d)-(f). The black dots within this block is completely removed in the J2K decoded image, while well preserved in our soft-decoded image. In addition, our proposed method tends to keep the relative pixel ranking among the neighboring pixels unchanged. For instance, in the center part of Fig.4.3 (d)-(f), the horizontal trend of the pixels is bright→dark→bright, which is broken in the J2K decoded image, while still appears in our soft-decoding image. This kind of information can be very important in object



(a) Original Image (Center = 128), J2K(Center = 148), Proposed(Center = 69)

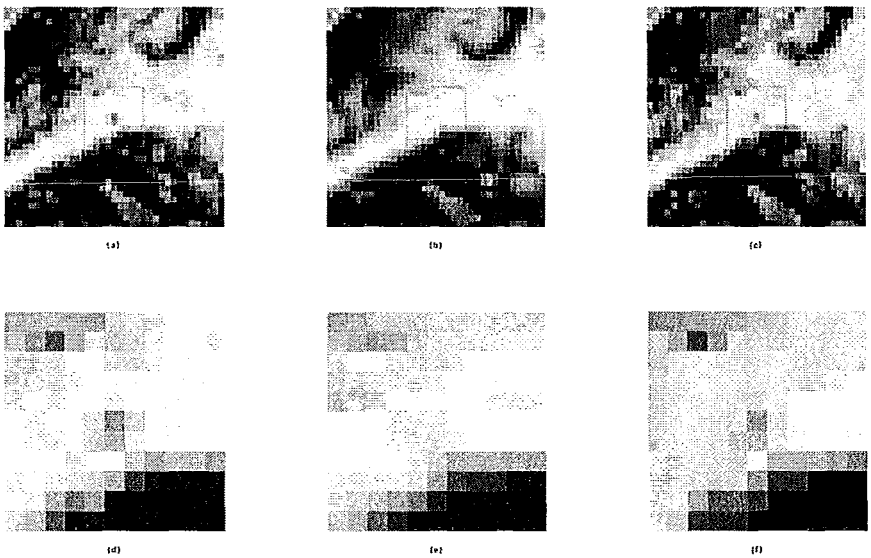
Figure 4.4: (a) Original image; (b) J2K decoded image; (c) Soft-decoded image; (d) Enlarged version of the red box of (a); (e) Enlarged version of the red box of (b);(f) Enlarged version of the red box of (c).

identification and localization, especially considering those applications in which the objects interested are composed a few pixels only.



(a) Original Image (Center = 149), J2K(Center = 93), Proposed(Center = 150)

Figure 4.5: (a) Original image; (b) J2K decoded image; (c) Soft-decoded image; (d) Enlarged version of the red box of (a); (e) Enlarged version of the red box of (b);(f) Enlarged version of the red box of (c).



(a) Original Image (Center = 29), J2K(Center = 101), Proposed(Center = 17s)

Figure 4.6: (a) Original image; (b) J2K decoded image; (c) Soft-decoded image; (d) Enlarged version of the red box of (a); (e) Enlarged version of the red box of (b);(f) Enlarged version of the red box of (c).

Chapter 5

Conclusions

In this thesis we investigated high-fidelity image compression for high-throughput and energy-efficient cameras. We demonstrate that L_∞-constrained predictive coding, i.e., DPCM coupled with uniform scalar quantizer, can offer a more compelling and viable compression solution than DSC and CS for high-fidelity, high-throughput and energy-efficient cameras, if soft decoding is carried out via image modeling and constrained optimization. In L_∞-constrained compression, the compression error is tightly bounded for each single pixel, i.e., $\|\mathbf{I}-\hat{\mathbf{I}}\|_\infty \leq \tau$, where $\hat{\mathbf{I}}$ is the decompressed image and τ is a small positive integer. This can be implemented at very low encoder complexity by embedding a fixed-rate uniform scalar quantizer of cell size $2\tau+1$ into the prediction loop of the DPCM scheme, with the quantizer reproduction value being the mid point of the quantizer cell.

The main innovation of this work is to exploit the L_∞ bound $\|\mathbf{I}-\hat{\mathbf{I}}\|_\infty \leq \tau$ as strong side information about the original image $\mathbf{I}$ and reestimate $\mathbf{I}$ from the hard decoded image $\hat{\mathbf{I}}$. Somewhat surprisingly, although the knowledge that for every pixel in $\hat{\mathbf{I}}$ the decoded value deviates from the original by no more than τ is available to

the decoder free of any extra bits, it was not used until now. To unlock the power of the side information inherent to L_∞-constrained compression, we perform an iterative soft decoding of $\hat{\mathbf{I}}$ using a prior image model of piecewise autoregressive process with the convex constraint $\|\mathbf{I} - \hat{\mathbf{I}}\|_\infty \leq \tau$. In each iteration it is an inverse problem that can be efficiently solved by constrained linear least-square estimation technique.

As an L_∞-constrained coding scheme, the proposed compression system has an added benefit of maintaining a tight error bound on each pixel. By the original design of minmax distortion, the hard decision decoder of an L_∞-coded image achieves $\|\mathbf{I} - \hat{\mathbf{I}}\|_\infty \leq \tau$. While the least-square soft decoder aims to reduce the average distortion in L_2 metric, it can still ensure a good L_∞ bound $\|\mathbf{I} - \mathbf{X}\|_\infty \leq (1+\beta)\tau$, $0 \leq \beta \leq 1$, where $\mathbf{X}$ is the soft decoded (reestimated) image. This property is important for many applications of stringent quality requirements on image compression algorithms, when decompressed images are subject to rigorous computer analysis, not just to please the eyes as in entertainment and consumer applications. The L_∞-constrained L_2 soft decoding strategy achieves the best of the both worlds: high compression ratio of lossy compression and a predetermined minmax fidelity that is close to lossless compression. A pure L_2 distortion criterion can incur large errors on some pixels that are statistical outliers. Such large individual errors, although with negligible contribution to PSNR, can be disastrous in some scenarios.

The proposed soft decoding strategy can improve the PSNR of L_∞-coded image $\hat{\mathbf{I}}$ by up to 2dB, and for bit rates above 1.2 bpp it can even outperform competitive encoder-centralized image codecs, such as J2K. By shifting the computation burdens and the pursue of high compression performance to the decoder, we provide an asymmetric compression solution for high-fidelity applications, in which the encoder

has low complexity and low energy consumption with all the heavy computations performed by the decoder for high-coding efficiency.

Bibliography

[1] B. M. Oliver, J. R. Pierce, and C. E. Shannon. The philosophy of pcm. *Proceedings of the IRE*, 36(11):1324–1331, 1948.

[2] D. A. Huffman. A method for the construction of minimum-redundancy codes. *Proceedings of the IRE*, 40(9):1098–1101, 1952.

[3] J. Rissanen and G. G. Langdon. Arithmetic coding. *IBM J. Res. Dev.*, 23(2):149–162, 1979.

[4] J. Ziv and A. Lempel. A universal algorithm for sequential data compression. *IEEE Transactions on Information Theory*, 23(3):337–343, 1977.

[5] C. C. Cutler. Differential quantization of communication signals. http://www.freepatentsonline.com/2605361.html(2605361), July 1952.

[6] X. Wu and N. Memon. Context-based, adaptive, lossless image coding. *IEEE Transactions on Communications*, 45(4):437–444, 1997.

[7] H. Kramer and M. Mathews. A linear coding for transmitting a set of correlated signals. *IRE Transactions on Information Theory*, 2(3):41–46, 1956.

[8] J. Huang and P. Schultheiss. Block quantization of correlated gaussian random variables. *IEEE Transactions on Communications Systems*, 11(3):289–296, 1963.

[9] G. K. Wallace. The jpeg still picture compression standard. *IEEE Transactions on Consumer Electronics*, 38(1), 1992.

[10] M. Barlaud, P. Mathieu, and M. Antonini. Wavelet transform image coding using vector quantization. In *Proc. Sixth Multidimensional Signal Processing Workshop*, pages 103–104, 1989.

[11] A. Said and W. A. Pearlman. A new, fast, and efficient image codec based on set partitioning in hierarchical trees. *IEEE Transactions on Circuits and Systems for Video Technology*, 6(3):243–250, 1996.

[12] J. M. Shapiro. Embedded image coding using zerotrees of wavelet coefficients. *IEEE Transactions on Signal Processing*, 41(12):3445–3462, 1993.

[13] D. Taubman and A. Zakhor. Multirate 3-d subband coding of video. *IEEE Transactions on Image Processing*, (5):572–588, 1994.

[14] M. Charrier, D. S. Cruz, and M. Larsson. Jpeg2000, the next millennium compression standard for still images. In *Proc. IEEE Int Multimedia Computing and Systems Conf*, volume 1, pages 131–132, 1999.

[15] D. Slepian and J. Wolf. Noiseless coding of correlated information sources. *IEEE Transactions on Information Theory*, 19(4):471–480, 1973.

[16] S. S. Pradhan and K. Ramchandran. Distributed source coding: symmetric rates and applications to sensor networks. pages 363–372, 2000.

[17] Zixiang Xiong, A. D. Liveris, and S. Cheng. Distributed source coding for sensor networks. *IEEE Signal Processing Magazine*, 21(5):80–94, 2004.

[18] B. Girod, A. Aaron., S. Rane, and D. Rebollo-Monedero. Distributed video coding. *Proceedings of the IEEE*, 93(1):71–83, 2005.

[19] G. Cormode and S. Muthukrishnan. Combinatorial algorithms for compressed sensing. In *Proc. 40th Annual Conf. Information Sciences and Systems*, pages 198–201, 2006.

[20] R. Robucci, J. D. Gray, L. K. Chiu, J. Romberg, and P. Hasler. Compressive sensing on a cmos separable-transform image sensor. *Proceedings of the IEEE*, 98(6):1089–1101, 2010.

[21] K. Chen and T. V. Ramabadran. Near-lossless compression of medical images through entropy-coded dpcm. *IEEE Transactions on Medical Imaging*, 13(3):538–548, 1994.

[22] Ligang Ke and M. W. Marcellin. Near-lossless image compression: minimum-entropy, constrained-error dpcm. *IEEE Transactions on Image Processing*, 7(2):225–228, 1998.

[23] B. Aiazzi, L. Alparone, and S. Baronti. Near-lossless image compression by relaxation-labelled prediction. *Signal Process.*, 82(11):1619–1631, 2002.

[24] X. Wu and P. Bao. L_∞-constrained high-fidelity image compression via adaptive context modeling. *IEEE Transactions on Image Processing*, 9(4):536–542, 2000.

[25] X. Wu, G. Zhai, X. Yang, and W. Zhang. Adaptive sequential prediction of multidimensional signals with applications to lossless image coding. *IEEE Transactions on Image Processing*, (99):1, 2010. Early Access.

[26] A. Zandi, J. D. Allen, E. L. Schwartz, and M. Boliek. Crew: Compression with reversible embedded wavelets. pages 212–221, 1995.

[27] A. Said and W. A. Pearlman. An image multiresolution representation for lossless and lossy compression. *IEEE Transactions on Image Processing*, 5(9):1303–1310, 1996.

[28] L. Ke and M. W. Marcellin. Near-lossless image compression: minimum-entropy, constrained-error dpcm. In *Proc. Conf. Int Image Processing*, volume 1, pages 298–301, 1995.

[29] D. S. Kalivas and A. A. Sawchuk. Motion compensated enhancement of noisy image sequences. In *Proc. Int Acoustics, Speech, and Signal Processing ICASSP-90. Conf*, pages 2121–2124, 1990.

[30] P. Wei, J. Li, D. Lu, and G. Chen. A fast and reliable switching median filter for highly corrupted images by impulse noise. In *Proc. IEEE Int. Symp. Circuits and Systems ISCAS 2007*, pages 3427–3430, 2007.

[31] R. H. Chan, C. W. Ho, and M. Nikolova. Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization. *IEEE Transactions on Image Processing*, 14(10):1479–1485, 2005.

[32] X. Wu, E. U. Barthel, and W. Zhang. Piecewise 2d autoregression for predictive image coding. In *Proc. Int. Conf. Image Processing ICIP 98*, pages 901–904, 1998.

[33] X. Li and M. T. Orchard. Edge directed prediction for lossless compression of natural images. In *Proc. Int. Conf. Image Processing ICIP 99*, volume 4, pages 58–62, 1999.

[34] Xin Li and M. T. Orchard. New edge-directed interpolation. *IEEE Transactions on Image Processing*, 10(10):1521–1527, 2001.

[35] X. Zhang and X. Wu. Image interpolation by adaptive 2-d autoregressive modeling and soft-decision estimation. *IEEE Transactions on Image Processing*, 17(6):887–896, 2008.

[36] X. Wu and X. Zhang. Model-based non-linear estimation for adaptive image restoration. In *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing ICASSP 2009*, pages 1185–1188, 2009.

[37] X. Wu, X. Zhang, and X. Wang. Low bit-rate image compression via adaptive down-sampling and constrained least squares upconversion. *IEEE Transactions on Image Processing*, 18(3):552–561, 2009.