USC–SIPI REPORT #333

Compression Artifact Removal and Inverse Halftoning Using Robust Nonlinear Filtering

by

Mei-Yin Shen

May 1999

Signal and Image Processing Institute
UNIVERSITY OF SOUTHERN CALIFORNIA
Department of Electrical Engineering-Systems
3740 McClintock Avenue, Room 400
Los Angeles, CA 90089-2564 U.S.A.

Dedication

To my parents, my brother Hung-I
and my husband Hao-Yang

Acknowledgements

I have been exceptionally fortunate in having Prof. C.-C. Jay Kuo as my research advisor. I would like to express my deep gratitude to him for his constant encouragement, valuable guidance and generous support throughout my study at USC.

I also gratefully acknowledge Prof. Alexander A. Sawchuk and Prof. Ulrich Neumann for serving as my dissertation committee and for their insightful comments on this thesis. I further thank Prof. Christos Kyriakakis and Prof. Zhen Zhang for their suggestion and encouragement in my qualifying exam.

I am grateful to my colleagues and friends, Dr. Jin Li, Dr. Liang-Jin Lin, Dr. Po-Yuen Chen, Dr. Wei-Feng Chen, Dr. Houng-Jyh Wang, Dr. Teng-Ying Kuo, Hwangjung Song, Te-Chung Yang and other members in Prof. Kuo's research group for their help and friendship in the last few years.

Finally, but the most important, I thank my family – my parents, my brother Hung-I, and especially my husband Hao-Yang. Their unconditional love and support give me the strength to overcome all the difficulties in my graduate studies and I dedicate all my achievement to them.

Contents

List Of Tables

List Of Figures

Abstract

In this dissertation, a versatile nonlinear filtering approach based on the framework of robust estimation has been developed to solve two inverse problems: compression artifact removal and inverse halftoning.

In the first part of this research, a low complexity postprocessing algorithm was proposed. We first formulate the artifact reduction problem as a robust estimation problem. Under this framework, the artifact-free image is obtained by minimizing a cost function that accounts for the smoothness constraint as well as image fidelity. Unlike the traditional approach that adopts gradient descent search for optimization, a set of nonlinear filters is used to calculate the approximate global minimum to reduce the computational complexity so that real-time postprocessing is possible. The proposed approach is generic and flexible. It can be applied to different compression schemes with minor variation. We have tested our algorithm on several standardized algorithms, including JPEG, JPEG 2000, and H.263, and demonstrated that our approach can alleviate compression artifacts efficiently with a low computational complexity.

The elegant framework of robust estimation makes the proposed nonlinear filtering technique suitable not only for compression artifact reduction but also for other inverse problems such as inverse halftoning. Inverse halftoning plays a key role for allowing a wide range of operations to be performed on halftones. In the second part of this research, a blind inverse halftoning algorithm of low complexity was presented. The cascade of lowpass and nonlinear filters in conjunction with the edge enhancement technique produces high quality inverse halftoned images at a low computational cost. The proposed inverse halftoning algorithm outperforms existing techniques in terms of PSNR measurement and subjective quality. In addition, the proposed method is more robust w.r.t. different error diffusion kernels.

Chapter 1

Introduction

1.1 Significance of the Research

In this research, a versatile nonlinear filtering approach based on the framework of robust estimation has been developed to solve two inverse problems: compression artifact removal and inverse halftoning. Solutions to these two problems have a great practical value, and have been extensively studied in the past. Although there have been quite a few ideas proposed before, only a couple of them are actually implemented in commercial applications due to the complexity concern. Therefore, to solve these two inverse problems in a cost-effective way is the main focus of this study.

The rapidly evolving computer and communication technologies have created a large amount of audio-visual applications that transform traditional media into new forms of computer-based multimedia. The major obstacle for multimedia applications is the vast amount of visual information required to be transmitted over networks with a low bandwidth, such as via modem connection or through wireless channels. This would not be possible without efficient compression techniques. In addition, to ensure interchangeability among different platforms, standardization of coding schemes is also necessary. Unfortunately, most existing coding algorithms and standards yield visually annoying artifacts for the low bit rate coding. In general, the appearance and severity of coding artifacts depend on several factors including transform filters, coding bit rates, quantization schemes, and image contents. Reduction of these artifacts can substantially improve the overall visual quality of decoded images.

In the first part of this research, we examine the issue of reducing compression artifacts for low bit rate coding. In particular, we emphasize artifact reduction for images or video coded by standardized coding algorithms, since commercial applications heavily rely on coding standards, and many software and hardware implementations have already been done based on these standard algorithms.

To improve the subjective quality of compressed image/video without changing the codec structure or coding bit rates, two strategies are commonly adopted. One is performed at the encoder through visually weighted quantization while the other is done at the decoder end by applying the postprocessing technique to the decoded image/video. The primary objective of this work is to develop efficient postprocessing techniques for the reduction of compression artifacts. Especially, we put heavy emphasis on low complexity techniques which can be practically implemented for real-time applications. Besides, we attempt to address fundamental issues to make our approach generally applicable to different compression schemes, and less sensitive to the change of the quantization or codec structure. These features make our work distinctive from previous research work that primarily works for a particular type of compression scheme with a high computational complexity.

Our solution to the artifact reduction problem begins with robust estimation, which is the theoretical foundation of many state-of-the-art image restoration techniques. Then, a nonlinear filtering approach is applied to reduce the computational burden. The elegant framework of robust estimation makes the proposed nonlinear filtering techniques suitable not only for compression artifact reduction but also for other inverse problems such as inverse halftoning.

Inverse halftoning algorithms recover the gray-scale images from their binary halftone images. There are two major applications for inverse halftoning algorithms: halftone manipulation and compression. Unlike the continuous tone images, simple image processing tasks may cause severe degradation on binary halftones so that halftone image manipulation is not a trivial problem. A better perceptual quality can be obtained by applying inverse halftoning before manipulation and re-halftoning afterwards. As to halftone compression, most current approaches compress halftone images directly using a dictionary of patterns. The coding efficiency can be significantly improved if a halftone image is converted to a grayscale image and then compress the grayscale image. Generally speaking, inverse halftoning plays a key

role for allowing a wide range of operations to be performed on halftones. In the second part of this research, we extend our nonlinear filtering techniques to the solution of the inverse halftoning problem.

For the rest of this introductory chapter, we will provide an overview of the background of this research, problems to be solved, proposed solutions, contributions that were made, and the outline of this dissertation.

1.2 Background

1.2.1 Compression Standards

The early effort to define the compression standard can be traced back to 1986, when the International Standard Organization (ISO) started a working group to develop the image compression standard. The first image coding standard, known as JPEG [18], was finalized in 1991. The JPEG standard can compress typical images to 1/10 to 1/20 of their original sizes without visible distortion. As to video compression, the MPEG-1 video compression standard [15], aiming at coding at rates of 1 – 1.5 Mbits/sec, was established in 1992. The second phase of the MPEG standard, which is known as MPEG-2 [16] and covers a wide range of applications from broadcast video to HDTV, was finalized in 1994. The MPEG algorithms are developed for the coding of full motion video in higher bit rates. Due to the long encoding delay and high overheads in MPEG bitstreams, the MPEG standards are not suitable for real-time audiovisual applications such as videoconferencing, and videotelephony. The first standard developed to meet this need was the CCITT H.261 standard [20]. Historically, H.261 was started long before JPEG, i.e. in 1984. The development has gone through several stages. In 1993, the H.261 standard was finalized to be a $p \times 64$ kbits/sec codec, with p between 1 to 30. A recent refinement of H.261 standard is called H.263, which aims at applications requiring a bit rate lower than 64 kbits/sec. Most of the above standards were initiated in early '90s based on technologies available at that time. As the blooming of the Internet and mobile applications, many existing standards no longer adequately meet the requirements of new media services. The need for faster, cheaper, and more efficient ways to

deliver audio-visual information to a remote site continues to drive new compression technologies and standards.

Currently, there are several on-going image/video standard activities. They are: JPEG 2000 [19], MPEG-4 [17], and H.263+/++ [9]. These new standards place great emphasis on the coding performance in low bit rates. In addition, extra functionalities, such as scalability, error resiliency, or even content-based interactivity are also considered in the new generation of standards. For the image compression standard, JPEG 2000 is intended to advance the current JPEG in order to meet the requirements of the applications in the next millenium. Specifically, the new standard will address areas that current JPEG fails to support, such as low bit-rate (below 0.25 bpp) compression performance, region of interest, and scalability, etc. Although the development of JPEG 2000 is in progress, a wavelet-based coding scheme has already been adopted as the baseline coding algorithm in a recent meeting due to its advantages over the DCT-based coding methods in terms of the rate-distortion trade-off and the visibility of compression artifacts. For video coding, MPEG-4 adopts a completely new concept — visual object coding instead of the traditional frame-based coding. This concept facilitates the access of visual objects in natural or synthetic videos to support content-based interactivity in various multimedia applications. The version one of MPEG-4 will be released in 1998. As mentioned earlier, the H.263 standard was proposed for video coding at extremely low rates (below 64kbits/sec). H.263+ is its near term refinement. The basic H.263 codec shares the major building blocks of the previous H.261 standard. Fifteen negotiable options are included to embrace technological advances in low bit-rate video coding. Currently, most streaming video coding schemes are based on H.263. The timeline of standard activities discussed above is depicted in Figure 1.1.

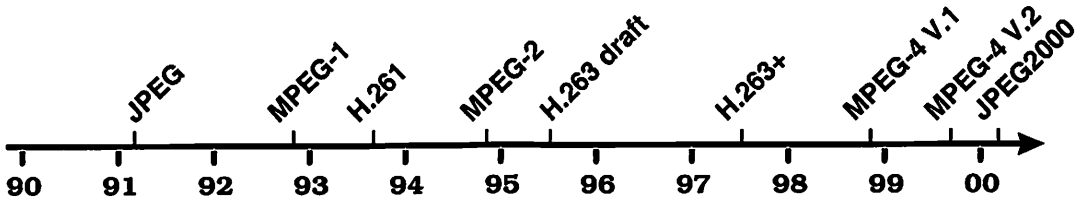


Figure 1.1: The timeline of standard activities.

1.2.2 Digital Halftoning

Digital halftoning is a process used to convert continuous tone images to bilevel images so that the binary images are perceptually similar to the original ones when observed from a distance. Halftone images are used in printing, facsimile, and many other applications. Popular halftoning techniques include ordered dithering, error diffusion, and optimization based techniques. These techniques differ in their processing structures and the characteristics of resulting bilevel images. For example, dithering [51] is a commonly used technique due to its low complexity. However, the spectra of dithered halftone images exhibit strong discrete components due to the periodicity of halftone patterns. There are many algorithms proposed to minimize the pattern visibility and maximize image details in dithered halftones [37, 53].

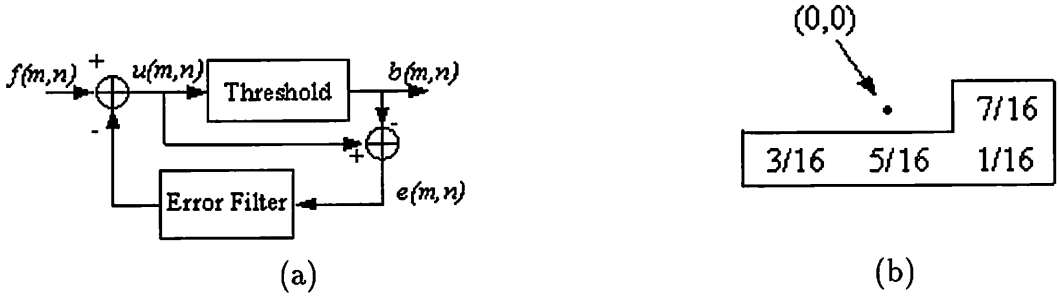


Figure 1.2: (a) Error diffusion halftoning process and (b) the Floyd-Steinberg kernel for error diffusion.

Another important halftone technique is error diffusion, which was first introduced by Floyd and Steinberg in [11]. The block diagram of an error diffusion system is shown in Figure 1.2 (a). Generally speaking, error diffusion can be considered as a 2D extension of delta-sigma modulation. The output of error diffusion is produced by thresholding the modified input signal which is defined to be the input signal subtract the weighted sum of past quantization errors. The simplest error diffusion kernel is the Floyd-Steinberg kernel as shown in Figure 1.2 (b). The advantage of this approach is its ability to shape the spectrum of noise introduced in quantization so that the resulting halftone has better image quality. Due to the nonlinearity of the system, error diffusion is difficult to analyze theoretically. However, it provides a good trade-off between the complexity and the performance and many researchers have been devoted to the theoretical and practical understanding of the

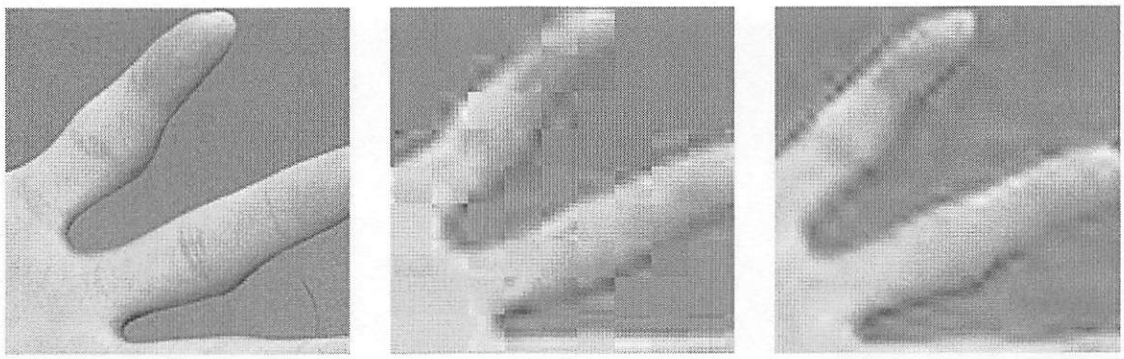
error diffusion technique and the improvement of halftone results. An example of an improved error diffusion technique is the multiscale error diffusion proposed by Katsavounidis and Kuo [22].

Recent halftone research emphasizes the generation of high quality halftones, and several optimization-based techniques have been proposed [43, 62, 55]. In an optimization approach, an error metric that measures the distortion between continuous tone and halftone images is selected. The result of the halftone image is found by minimizing the distortion. The definition of the distortion measure is extremely important in this approach, since it directly impacts the quality of output halftones.

1.3 The Problems

1.3.1 Compression Artifacts

The development of international compression standards usually goes through several times of competition and verification before the convergence of the final algorithms. This ensures the technology behind the standard is truly start-of-the-art at that time. Clearly, these standard algorithms can successfully compress an image or video to a certain rate without perceptible artifacts. However, the problem arises as the compression ratio is pushed beyond the target bit rates. The perceived quality of decompressed images or video is severely degraded with the presence of compression artifacts. Generally, the severity of artifacts is determined by the compression ratio and the image content. The more an image is compressed, the severer the artifacts occur after decompression. On the other hand, images with simple content usually have fewer artifacts than the complex one when being coded at the same bit rate. Besides, the appearance of artifacts is dependent on compression techniques. For example, blocking and ringing effects are major artifacts perceived in DCT-based algorithms (e.g. JPEG, MPEG, and H.26x) while ringing is the only objectionable artifact in wavelet-based methods (e.g. JPEG 2000). Examples of compression artifacts generated by DCT- and wavelet-based coding methods are shown in Figure 1.3.



(a) (b) (c)

Figure 1.3: Examples of compression artifacts: (a) the original image, (b) the JPEG coded image at 0.26 bpp, and (c) the wavelet-coded image at 0.125 bpp.

1.3.2 Halftone Conversion and Compression

The need for inverse halftoning arises in many applications that require conversion, manipulation, or compression on halftone images. Much better results can be obtained if theses operations are applied to grayscale images than directly to halftone images.

As mentioned before, a simple image processing operation may introduce a large distortion on the halftone image. Examples of a direct manipulation of halftones are shown in Figure 1.4. It is well known that a good halftone image should have a *blue noise* characteristics [52]. That is, the black and white dots should spread out as uniformly as possible. From Figure 1.4 (b) and (c), we see clearly the quality degradation caused by scaling and rotating a halftone image.

Another reason for inverse halftoning is the need of halftone conversion. The recorded image resolution, measured by dots per inch, is usually different from the resolution of the printing or the display device. Rehalftoning is necessary to convert a halftone image between different resolutions. The rehalftoning process is essentially inverse halftoning followed by forward halftoning.

Rehalftoning is also useful in the compression of halftone images. Halftone compression is a key issue in the emerging JBIG2 standard, which is the coding standard for the next-generation fax machines, printers and scanners. In JBIG2, halftone images may be either compressed directly or converted to grayscale images for compression. The performance of JPIG2 lossless compression is high for clustered dot

halftones (5:1) but low for error diffused halftones (2:1). A good inverse halftone algorithm to improve the coding gain for error diffused halftones is very desirable.

1.4 Proposed Solutions

1.4.1 Postprocessing of the Decoded Data

To reduce visual artifacts, compression algorithms along are not sufficient. Additional techniques, such as the encoder-based visually weighted quantization and/or the decoder-based postprocessing, are needed. The decoder-based postprocessing techniques have been gained a lot of attentions, since that they can be easily incorporated with current standards.

Various postprocessing techniques have been developed in the last decade, particularly for reducing blocking effects in JPEG compression. These techniques can be roughly categorized into two classes: image enhancement and image restoration. The former includes several adaptive filtering or interpolation techniques applied in various domain (e.g. spatial, DCT, or wavelet). The later contains many classical image restoration algorithms.

Applications that require real-time processing are strongly in favor of the enhancement approach due to its low complexity. However, its major drawback is that most techniques of this class limit their applicability to a particular type of artifacts, say, blocking effects only. In addition, to avoid smearing edges and other image details, they are heavily relied on explicit edge extraction techniques. Consequently, the enhancement approach is in general not robust to the choice of different transform and quantization schemes.

On the other hand, image restoration techniques have also been extended to solve the compression-artifact problem. Restoration techniques usually consist of two stages. First, a cost function is formulated based on the prior knowledge about artifacts as well as the image smoothness. Second, the estimate of an artifact-free image as the solution of this cost function is calculated usually by iterative optimization methods. The advantage of this approach is in its superior artifact-reduction performance. However, it is mainly suitable for off-line processing, where

constraints on the computational complexity and the memory bandwidth can be relaxed.

Previous work put much emphasis either on the low complexity or the high artifact-removal performance. Nevertheless, none of the work provides a good trade-off between the complexity and the performance. In this research, we try to bridge the gap between enhancement and restoration approaches, and balance the requirements of low complexity and high artifact reduction performance. Our goal is to develop an efficient postprocessing technique that is not only applicable to various types of compression schemes but also practical for real-time applications. We proceed to achieve this goal by examining a class of nonlinear filters and establishing the foundation on robust estimation theory.

With our approach, we first formulate the artifact reduction problem as a robust estimation problem. Under this framework, the artifact-free image is obtained by minimizing a cost function that accounts for the smoothness constraint as well as image fidelity. Unlike the traditional approach that adopts gradient descent search for optimization, a set of nonlinear filters is used to calculate the approximate global minimum to reduce the computational complexity so that real-time postprocessing is possible. The proposed approach is generic and flexible. It can be applied to different compression schemes with minor variation. We have tested our algorithm on several standardized algorithms, including JPEG, JPEG 2000, and H.263, and demonstrated that our approach can alleviate compression artifacts efficiently with a low computational complexity.

1.4.2 Inverse Halftoning

The forward halftoning process is a many-to-one mapping. Therefore, the inverse halftoning process is often cast as an ill-posed image recovery problem that has no unique solution. Previous work can be classified into frame-based and scan-based approaches depending on their operations and storage requirements. The frame-based approach requires the buffering of the entire frame for inverse halftoning. The resulting methods are basically modified image restoration techniques. The obtained continuous tone image is a solution that best fits the prior knowledge. However, frame-based methods are slow and memory-extensive. Also, they are under the

assumption that the halftoning scheme and the halftone kernels are known, which is not realistic in practical applications.

The scan-based approach is performed locally. The corresponding methods are essentially adaptive lowpass filters which smooth halftone noise while preserving edge details. Since the halftone noise is high frequency blue noise, a natural solution is to use a lowpass filter to reduce it. Most scan-based methods use linear filters for smoothing. They have the advantages of low complexity and low memory. They are however less robust due to the pre-determined passband of lowpass filters.

Based on the above discussion, we conclude that there are two key elements in all inverse halftone algorithms, i.e. lowpass filtering and edge enhancement. The former is important in reducing halftone noise. The latter is used for edge sharpening since sharpness of edges can be significantly reduced in the forward halftoning process. Neither filtering nor restoration can well reconstruct the original edge quality so that additional enhancement to recover edges is needed.

In this work, we propose a new approach in which lowpass and nonlinear filters are cascaded for noise smoothing. This combined filter provides more robust smoothing than linear adaptive filtering. In addition, an edge enhancement algorithm is applied to recover the original edge quality. Since our inverse halftoning technique is a filtering-based approach, it has less complexity and memory requirements than frame-based techniques. Although the proposed method is primarily designed for error diffused halftone images, it can also be applied to other types of halftone images with some slight modification.

1.5 Contributions of the Research

1.5.1 Compression Artifact Removal

Our ultimate goal is to develop an unified framework for the postprocessing of all low-bit-rate compressed multimedia data. This research is one step toward this goal. Our contributions that advance existing postprocessing techniques are summarized as follows.

- **Robustness.** Robust estimation is applied to a generic image model for the estimation of an artifact-free image. This approach inherits the richness of underlying robust estimation theory. It can smooth compression artifacts while preserving the quality of edges without the use of explicit edge-extraction techniques. In addition, the data fidelity constraints are imposed to avoid the distortion introduced by postprocessing filters.

- **Flexibility.** Nonlinear filtering is adopted instead of the gradient descent search used in other restoration-based work. This filtering approach makes the computation tractable for both convex and non-convex potential functions. With the flexibility of choosing non-convex potential functions, the proposed technique is able to achieve both artifact reduction (image smoothing) and image enhancement (edge recovery, or even edge sharpening) at the same time.

- **Efficiency.** It is shown by experimental results that the proposed postprocessing technique can efficiently reduce the noticeable compression artifacts existing in the decompressed image and video.

- **Complexity.** In the restoration-based approach, the solution is obtained by computationally intensive iterative techniques. Besides, the data fidelity is often defined in the transform domain so that additional forward and backward transforms are required in each iteration. To overcome these problems, we define our models in the space domain only, and adopt the nonlinear filters to reduce the computational burden. In real-time postprocessing of video, the time constraint is even more critical. A table-lookup approach is proposed for the deblocking purpose to further reduce the complexity.

- **Memory.** The memory requirement for our proposed technique is low. Since all postprocessing filters do not use information outside a small local window, the processing can be done without a large buffer. This is important for hardware devices that have a memory constraint.

- **Compatibility.** All of postprocessing algorithms proposed in this research can be implemented as post-filters for current compression standards. There is no need to change the codec structure or the bitstream format.

1.5.2 Inverse Halftoning

The proposed inverse halftoning algorithm outperforms existing techniques in the following aspects.

- **Better objective and subjective quality.** Results of the proposed method rank the top among all others in terms of the PSNR measure. The reconstructed continuous tone images are more visually pleasant than those obtained from other methods.

- **More efficient to suppress noise.** In the linear adaptive filtering approach, the cut-off frequency of pre-designed filters cannot be arbitrarily low in order to preserve important details. This results in small grainy noise in flat or slow varying regions of the reconstructed gray-scale image. This type of noise is very difficult to remove by linear filtering. Bayesian-based algorithms can reduce this type of noise at the price of a high complexity. In contrast with these algorithms, the propose technique can reduce this type of noise by simply increasing the window size of the robust nonlinear filter.

- **Low complexity and low memory.** The proposed technique is scan-based so that it can be performed locally. The computational complexity and memory requirements are slightly more compared to linear adaptive filtering, but much less than the frame-based approach.

1.6 Outline of Thesis

The rest of this proposal is organized as follows. A thorough survey of related work in the field is presented in Chapter 2. The basic idea of applying robust estimation theory to artifact reduction and inverse halftoning is described in Chapter 3. Postprocessing of wavelet-based coding (adopted in the JPEG-2000 standard) with the proposed nonlinear filtering technique is given in Chapter 4. A real-time postprocessing technique for the H.263 codec and low-bit-rate JPEG is detailed in Chapter 5. A fast blind inverse halftoning technique is presented in Chapter 6. Finally, concluding remarks and future extensions are stated in Chapter 7.

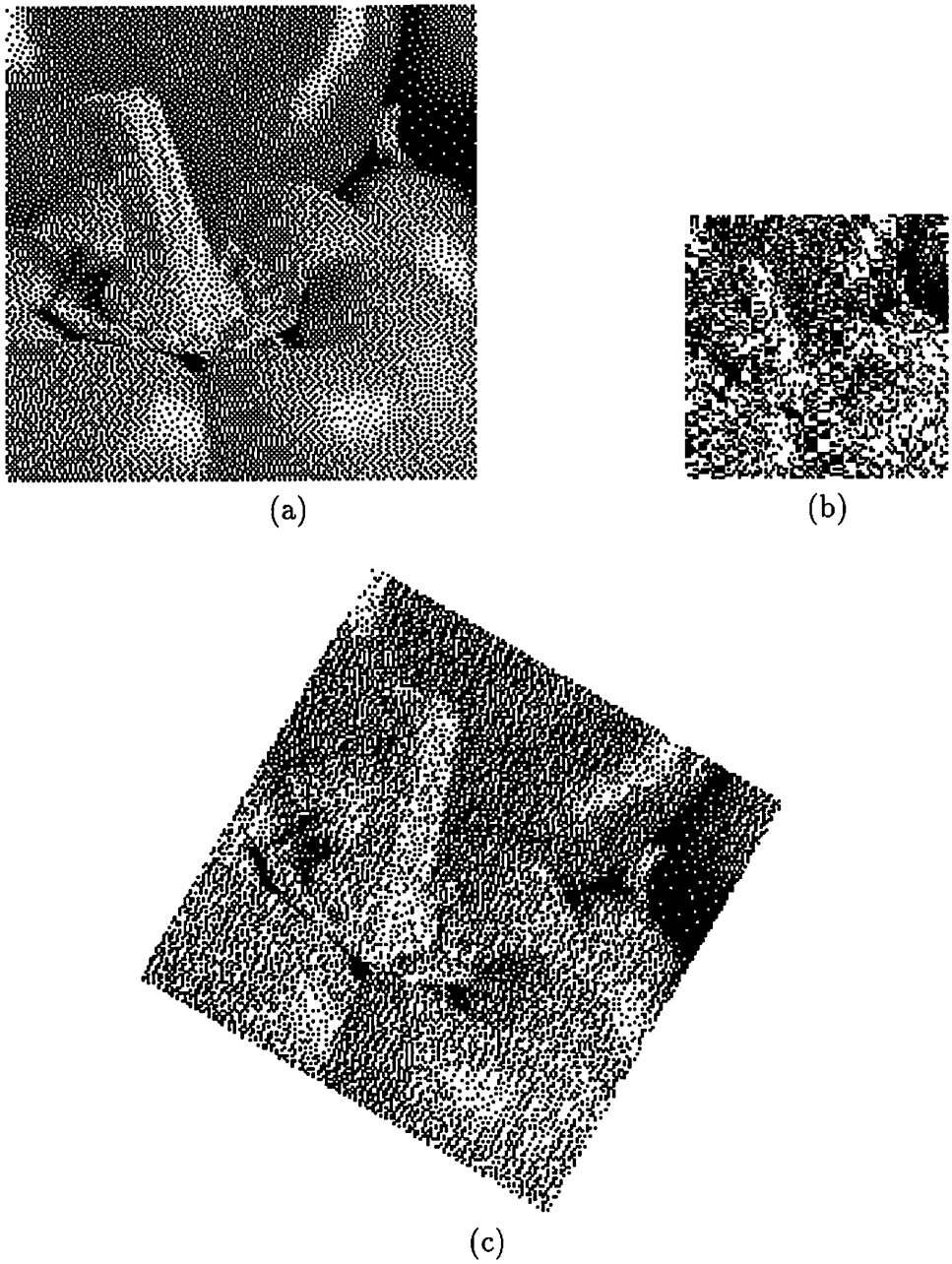


Figure 1.4: Examples of halftone manipulations: (a) the original halftone, (b) the resized halftone, and (c) rotation by 30 degree.

Chapter 2

Previous Work

This chapter reviews previous work in image/video postprocessing and inverse halftoning techniques. We intend to cover the entire spectrum of research in these fields by providing a complete survey of various algorithms ranging from the simplest filtering method to the most sophisticated restoration approach.

2.1 Review of Postprocessing Techniques

Early work in postprocessing research is motivated by the need of reducing blocking artifacts in JPEG compression. A wide range of linear- and rank-based filtering techniques [24, 25, 34, 44, 45, 47] are proposed for this purpose. These filtering techniques often incorporate additional classification and edge-extraction schemes to avoid blurring the important image detail. However, to correctly estimate the edge information is a very difficult task for very high compression, where the actual edge information is somewhat scrambled.

To overcome the robustness problem of the filtering approach, the postprocessing research turned to restoration techniques which generally provide a more robust and consistent artifact-reduction performance. Image restoration is a relatively mature field, and all its major techniques, such as the minimum mean square error estimation, the maximum a posteriori (MAP) estimation, and several constraint-based techniques can be extended to this new application, i.e. compression artifact reduction.

Wavelet related research has been very active since early '90s. Recently, some wavelet-based algorithms have been proposed for the deblocking purpose. There are

other types of postprocessing techniques, such as AC prediction [40] recommended in the JPEG standard, VQ enhanced decoding, and perceptually optimized postprocessing techniques, *etc.*

While there is a large amount of work to reduce blocking artifacts in JPEG compression, only few methods have been studied for the postprocessing of wavelet compressed images. Ringing around edges is clearly perceived in the wavelet-based compression. In comparison with JPEG ringing artifacts that are confined within the block partitioning boundaries, the ringing effect in wavelet compression tends to cover a wider range of pixels. Consequently, it is difficult to remove with the short-tap FIR filters and all previous work about wavelet postprocessing are restoration-based techniques.

Over the years, many ideas have been developed for postprocessing of compression artifacts in still images, little work has been done in the quality improvement of compressed video. Current standards, such as H.263 and MPEG-4, have defined postprocessing filters for the deblocking, or even the deringing (e.g. in MEPG-4) purpose. The reason to define these filters in the standard is that they can be used not only as pure post-filters but also as loop filters used in the motion compensation loop in both the encoder and the decoder to further improve the quality of predicted frames.

2.1.1 Postprocessing via Image Enhancement

- **Filtering**

Since blocking and ringing artifacts resulting from block DCT coding are high frequency artifacts, a straightforward solution is to apply lowpass filtering to the region where artifacts occur. A space-invariant filtering method to reduce blocking artifacts was first proposed by Reeves and Lim [45]. However, filtering without considering the local image statistics often causes the loss of high frequency details such as edges. Therefore, a number of adaptive spatial filtering techniques [24, 25, 34, 44, 47] have been proposed to overcome this problem.

With the concern about the complexity and the memory requirement, adaptive spatial filtering provides a practical method to reduce blocking and ringing effects in

video compression. Several algorithms [3, 30, 59] were proposed along this direction. The post-filters defined in H.263 and MPEG-4 belong to this family.

The adaptive post-filtering has the advantage of a low computational complexity but the adaptation to local statistics relies on edge-exaction and classification schemes, which are usually not robust at low bit rates.

- **AC Prediction**

Niss [40] proposed a technique to reduce the blocking effect in JPEG by predicting the low-frequency AC coefficients from DC coefficients in the given block and its 8 neighboring blocks. A quadratic surface, given by

$$P(x,y) = a_1x^2y^2 + a_2x^2y + a_3xy^2 + a_4x^2 + a_5xy + a_6y^2 + a_7x + a_8y + a_9 \quad (2.1)$$

is used to model the shape of a local surface. Coefficients a_i, $1 \leq i \leq 9$, can be determined by finding the best fit to the surface determined by nine DC values in an array of 3×3 blocks. Once coefficients $a_1, \cdots, a_9$ are determined, low-frequency AC coefficients required to reproduce the quadratic surface can be calculated by a set of equations represented in terms of these nine DC values. To avoid the introduction of a large distortion value, predicted AC coefficients are confined to their original quantization intervals. Based on this framework, Lakhani [27] proposed a set of improved prediction equations by using a smaller block of 4×4 pixels, instead of the original size of 8×8. The performance is improved at the cost of a higher computational complexity. This approach requires the least amount of computation among all methods, and it can be carried out on the fly during decoding. However, the quadratic surface assumption only works for the smooth area. The prediction is generally not valid in the edge or texture areas. Moreover, at very low bit rate coding, the modification of only five lowest AC coefficients is not sufficient for a good deblocking performance.

- **Wavelet-based Deblocking**

The wavelet-based denoising algorithm was first proposed by Donoho [8] to reduce the white Gaussian noise in noisy signals. Gopinath *et al.* [12] applied Donoho's algorithm directly to deblocking. There is however a problem. As point out by

Coifman and Donoho in [5], de-noising with traditional (i.e. orthogonal and maximally decimated) wavelet transforms exhibits visual artifacts, such as the Gibbs phenomenon in the neighborhood of discontinuities that corresponds to the ringing effect in the image case. They suggested to use a shift-invariant (or equivalently undecimated) wavelet transform. Based on the suggestion of Coifman, the overcomplete wavelet representation is adopted by most researchers under this framework. Besides, with the observation that the blocking effect is a structured noise, similar methods are proposed by Lu [31] and Xiong *et al.* [58] to handle the denoising problem. They applied Mallat's multiscale edge representation [36] to identify edge positions, and blocking artifacts are reduced by thresholding wavelet coefficients at block boundaries, except pixels where an edge is located.

Although the reported deblocking performance seems good. This approach still has some problems. First, the complexity of the overcomplete wavelet transform is $O(N \log N)$ for a 1-D signal with N samples. For a 2-D image of size $N \times N$, the complexity is $O(N^2 \log N)$, which is too high for real-time applications. Second, to extend this approach to deringing is not straightforward. The performance relies on the robustness of edge-detection algorithms. Third, additional artifacts may occur due to improper thresholding of edges. Finally, the memory requirement is another major concern. It requires at least $O(N^2 \log N)$ for the storage of overcomplete wavelet coefficients.

- **Other Techniques**

There are other types of postprocessing techniques, such as the VQ enhanced decoding proposed Wu and Gersho [56] and the perception-based deblocking proposed by Macq *et al.* [35]. In the former approach, the inverse DCT operation in the JPEG decoder is replaced with a nonlinear interpolation obtained by table lookup. The later considering about masking effect [6] in human visual system (HVS). The artifact-free image is obtained by cancellation of the unmasked noise.

2.1.2 Postprocessing via Image Restoration

The image restoration approaches treat artifact removal as an ill-posed image recovery problem, they can be generally classified into the following three categories:

1. **criterion-based methods.** The solution is found by some predefined optimality criterion. For images, the commonly used criterion is MSE. Examples of this approach include the minimum least squares (MLS), the minimum mean square error (MMSE), or the maximum a posteriori (MAP) estimation.

2. **constrained optimization-based methods.** This class of methods optimizes an optimality criterion subject to constraints on the solution. A well-known method of this type is constrained least squares (CLS).

3. **constraint-based methods.** One example in this class is projection onto convex sets (POCS).

The above techniques applied to compression artifact reduction have been widely studied. Since details about these techniques can be found in image restoration literatures, we only review methods that are closely related to our work.

• **Minimum Mean Square Error (MMSE) Estimation**

Let $\mathbf{y}$ be the decompressed image of the original image $\mathbf{x}$. The degradation due to compression can be formulated as

$$\mathbf{y} = \mathbf{x} + \mathbf{e}, \tag{2.2}$$

where $\mathbf{e}$ is the compression error. We would like to find an estimate $\hat{\mathbf{x}}$ of the original image so that the cost

$$\mathbf{E}\{\|\hat{\mathbf{x}} - \mathbf{x}\|^2\} \tag{2.3}$$

is minimized. The best estimate $\hat{\mathbf{x}}$ can be obtained by using a linear or a nonlinear estimator acting on the observed image $\mathbf{y}$. In the case that linear-based estimators are used, this approach become the classical Wiener filters. Nakajima *et al.* [38] first applied this approach to reduce coding artifacts in MPEG coded video. A similar work for deblocking of JPEG compressed images was proposed by Hong *et al.* [14]. Both of them obtained the error variance empirically by training the test data. Choy *et al.* [4] applied this technique in the transform domain with the assumption that quantization noise is uniformly distributed.

There are two problems with this approach. First, for the pure space domain processing [38, 14], the formulation in (2.3) has the underlying assumption that

the error is with the Gaussian distribution. This is clearly not a correct model for JPEG blocking artifacts. Second, the linear-based estimator does not work well at the discontinuity region, since it assumes the Gaussian distribution for the samples as well. This becomes a more serious problem when the linear estimator is applied to the transform coefficients as in [4]. The reason why their algorithm still works is due to the additional quantization constraint imposed on estimated transform coefficients.

- **Maximum A Posteriori (MAP) Estimation**

In MAP-based restoration, the criterion is to maximize the probability density of the original image conditioned on the observation. Let $\mathbf{y}$ be the observed image that contains compression artifacts. The original image $\mathbf{x}$ can be estimated by a MAP estimate $\hat{\mathbf{x}}$ via

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}} Pr(\mathbf{x}|\mathbf{y}).$$

By applying Bayes' rule and the log-likelihood function, we obtain

$$\hat{\mathbf{x}} = \arg\max_{\mathbf{x}}\{\log Pr(\mathbf{y}|\mathbf{x}) + \log Pr(\mathbf{x}) - \log Pr(\mathbf{y})\}, \tag{2.4}$$

$$= \arg\max_{\mathbf{x}}\{\log Pr(\mathbf{y}|\mathbf{x}) + \log Pr(\mathbf{x})\}. \tag{2.5}$$

Since $Pr(\mathbf{y})$ is constant with respect to the optimization parameter $\mathbf{x}$, the term $\log Pr(\mathbf{y})$ can be dropped. The first term of (2.5) models the data fidelity which is the different between the resulting estimate and the original image. The second term models the smoothness of the estimate based on the prior assumption about the original image. Stevenson [48, 41] first applied this Bayesian framework for reduction of coding artifacts. In his approach, a Huber function is used to regularize the image smoothness. Considering the difficulty of modeling data fidelity in spatial domain, the first term of (2.5) is imposed by projecting the estimate onto the original quantization bins in the transform domain. The estimate that best fits the model in (2.5) can be determined by several iterative techniques, including the gradient projection method [48, 41], or the iterative conditional mode (ICM) [33]. Li and Kuo [28] proposed a multiscale MAP method, in which the decoded image was enhanced from coarse to fine scales. Postprocessing at coarse scales improves the global appearance of the image and reduces long range artifacts such as ringing,

while postprocessing at finer scales preserves the sharpness of edges. It achieved a better performance than the single-scale MAP method in terms of ringing artifact reduction and computational speed. Another type of MAP-based formulation is to apply the line process to model the discontinuity in images. It was proposed by Özcelik *et al.* [42]. Unfortunately, the line process is not convex so that the solution cannot be computed by using the deterministic technique. The stochastic mean field annealing was used in [42] for optimization.

With a proper model, the MAP-based postprocessing algorithms can reduce all kinds of artifacts in a single framework. It is the state-of-the-art postprocessing technique from the viewpoint of artifact reduction performance. The major drawback with this approach is the high computational complexity and the large memory requirement.

- **Projection Onto Convex Sets (POCS)**

In POCS approach, a constraint set is defined as a closed convex set whose members are consistent with a priori knowledge of the original image. For m known properties, we can construct m closed convex sets, $C_i, i = 1, 2, ..., m$. A feasible solution, which is in the intersection of all sets, can be found via an iterative process. That is, we start from an initial guess, and project the result onto each constraint set in sequence until it converges to the final solution.

In the context of image postprocessing, two constraint sets are commonly used in POCS. One is the quantization constraint set (C_q) determined from the known quantization intervals of each DCT coefficient in the transform domain. The other is the smoothness constraint set that captures the smoothness properties of the original image (C_s) in the spatial domain. Let P_q and P_s be the two projection operators which project a solution onto C_q and C_s, respectively. Then, the artifact-free image can be obtained by performing the iteration

$$\mathbf{f}_{k+1} = P_q P_s \mathbf{f}_k, \tag{2.6}$$

where $\mathbf{f}_k$ is the postprocessed image at the kth iteration. The projection operator P_q simply restricts each DCT coefficient to its original quantization interval. The definition of P_s is more critical for noise smoothing. The concept of POCS for

artifact reduction was first proposed by Rosenholtz and Zakhor [46]. However, the low-pass filter used in their experiments cannot result in a closed convex set so that their algorithms did not converge to a single solution after iterations. Yang *et al.* [60] proposed the first POCS-based postprocessing algorithm that follows the strict definition of "convex constraint sets". He further improved it by considering spatial adaptivity [61]. Kwak and Haddad [26] improved the complexity of Yang's algorithm by canceling out the IDCT and DCT pairs in POCS iterations. There are other postprocessing techniques mentioned in the literature under the name of POCS. However, they are actually not POCS according to the strictest definition.

All POCS methods described above aim at reducing the blocking effect. If there are other types of artifacts present in the image, additional constraint sets are required and the corresponding computational complexity increases as well. Also, it is challenging to determine the proper constraint for non-blocking artifacts. For example, the constraint set for the ringing artifact is very difficult to define.

2.1.3 Summary

We summarize some of the trade-offs of these postprocessing techniques in Table 2.1.

2.2 Review of Inverse Halftoning Techniques

Several inverse halftoning algorithms have been proposed in the past ranging from simple lowpass filtering to sophisticated iterative optimization techniques. Since order dither and error diffusion will result in different halftone artifacts, they must be treated differently when performing inverse halftoning. Here we emphasize on the review of inverse halftoning for error diffusion, although some of them may also be used for ordered dither halftone with slight modification. We categorized these inverse halftone into frame-based and scanned-based approaches according to their operation and buffer requirements.

2.2.1 Frame-based Approaches

- **MAP Estimation**

Stevenson [50] applied the MAP estimation technique to the inverse halftone problem. The formulation is the same as his work for compression artifact reduction, i.e. (2.5), in the previous section. By assuming the Huber-Markov random field model(HMRF) for the ideal image distribution, the inverse halftoning problem can be reduced to a well-posed constrained optimization algorithm that minimizes the following cost function:

$$\sum_{c \in C} V_c(z) = \sum_i \sum_j \sum_{m=0}^{3} \rho_T(\mathbf{d}^t_{i,j,m} z), \tag{2.7}$$

where c is a clique that defines the neighboring pixels, z is the pixel value, $\rho(\cdot)$ is the Huber potential function , and T is the coefficient in the Huber function that controls the amount of smoothing. To perform better smoothing in flat and slow varying areas, $\mathbf{d}^t_{i,j,m}$ is defined as an finite difference approximation of second-order derivative [1]. The minimization of (2.7) can be obtained by steepest descent search. The global optimization only ensures the result close to the prior image model. However, an image may contain features that cannot be modeled by HMRF. Consequently, an additional projection operation is needed in each iteration to avoid some features being over smoothed. The projection operation ensures that the resulting grayscale image produces the same halftone as the original one. This scheme however requires the knowledge of the halftone scheme and the halftone kernel. This is one of the major drawbacks of this scheme. This method also suffers from a high computational complexity and a high memory requirement.

- **POCS**

Since inverse halftoning does not have a unique solution, POCS-based algorithms [13, 1] incorporate all available information to improve the quality of the solution. A feasible solution is obtained by alternative projection between two or more convex constraint sets as formulated in (2.6). The definition of the constraint set is critical to the performance of POCS algorithm.

Most commonly used constraint sets are defined in the space and the frequency domains. The simplest frequency-domain constraint set is a set that contains all

[1] This is different from artifact reduction case that only first-order residue is used

band-limited images. This is based on the assumption that halftone noise is high frequency so that an image containing only low frequency components must be free from halftone noise. The projection to this set can be easily performed by lowpass filtering. The halftone image itself provides another spatial constraint. This constraint is defined as a set contains all continuous-tone images that can produce the same halftone as the given one. The implicit assumption of this projection is the halftone process must be known beforehand. Projection to the original halftone in the space domain usually requires iterative methods such as the constrained least-square technique. The projection operation is done by iteratively modifying values in the resulting grayscale image so that it can produce exactly the same halftone as the original one. However, these constrained optimization techniques always face some numerical problems such as the convergence rate and stability.

There are other constrained sets such as edge information in the wavelet scale-space domain. However, projection onto this domain is even more difficult to apply, and the quality improvement may not be significant. Except the frequency projection, projection onto other spaces are very difficult and often iterative. If the projection operator is improperly performed, the performance will degrade significantly. Clearly, POCS has a very high complexity and is not practical for many applications.

- **Wavelet Denoising**

Xiong *et al.* [57] proposed an inverse halftoning algorithm using wavelets. This method is similar to his work for JPEG deblocking [58] with slight modification. The block diagram of this wavelet-based scheme is shown in Figure 2.1. Conventional inverse halftoning techniques only use information of the low-pass band (i.e. $S_1 f$) of a halftone image. In this approach, the information in all frequency bands are used to recover the edge information.

Because halftone noise has very low SNR, some halftone patterns are still visible in $S_1 f$ after one step wavelet decomposition. To reduce visible noise in $S_1 f$, two more scales wavelet decomposition are performed. Then, cross-scale correlations among higher frequency subbands $W_j^O f's (j = 2, 3, O \in \{H, V\})$ are used to identify the edge location. Then, for those coefficients not at edge positions, their values are set to zero. After reconstruction, $S_1 f$ is free from halftone patterns. However, $W_1^H f$

and $W_1^V f$ also contain halftone noise. A Gaussian filter with $\sigma^2 = 2$ is applied for smoothing in these subbands. The inverse halftoning result is reconstructed from smoothed $S_1 f$, $W_1^H f$ and $W_1^V f$.

This method can be performed with or without knowing the halftone kernel, and the PSNR value has been reported to be the best. However, there are problems with this approach. First of all, the reconstruction procedure is not specified in the original paper, which is critical. In Mallat's miltiscale edge detection method, the wavelet is defined as a quadratic spline with compact support. Since the decomposition is not orthogonal, the reconstruction from wavelet maxima requires iterative projection. Another issue is the memory requirement. Since the wavelet decomposition is an over-complete one, each subband has the same size as the original image. Thus, at least nine floating-point copies of the original image have to be stored at the same time in memory. This is not practical for low cost devices.

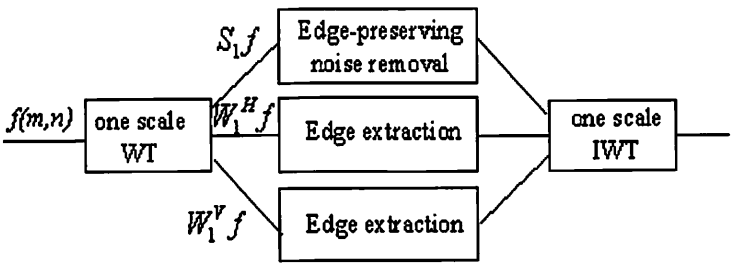


Figure 2.1: Block diagram of wavelet-based inverse halftoning scheme.

2.2.2 Scan-based Approaches

- **Adaptive Smoothing**

A linear adaptive filter for inverse halftoning is proposed by Kite *et al.* in [23]. A separable 7×7 FIR filter, which is parameterized by the strength of horizontal and vertical gradients, is applied to each pixel in the halftone image. Filter coefficients of a 1-D 7-tape filter can be found by normalized the vector,

$$[x_2 - x_1 + 2, x_2, x_1, 4, x_1, x_2, x_2 - x_1 + 2],$$

where x_2 and x_1 are chosen such that the filter satisfies the passband and stopband specification. In their implementation, x_1 is determined by the gradient estimation, and the value of x_2 can be found by

$$x_2 = 0.4631x_1^3 - 2.426x_1^2 + 4.660x_1 - 3.612.$$

For robustness, the gradient is estimated across scales. The edge is preserved due to a smoothing operation that is in parallel to edges while noise is reduced. A strong filter can be applied when edges are present.

- **Nonlinear Denoising**

A nonlinear de-noising method is proposed in [7]. A lowpass filter is first applied to the given halftone and then the edge information is extracted from the filtered result by using the bandpass filter and binary operations. The edge locations of the lowpass filtered image are enhanced with the computed edge map to get the final grayscale image. This method can deliver comparable PSNR at a fractional time of frame-based approaches.

2.2.3 Summary

The comparison of different inverse halftone schemes is summarized in Table 2.2. From these reported data, all methods have comparable PSNR results, so that there is no clear winner. We can only comment that there is a trade-off between complexity and robustness. Besides, some schemes assume that the halftone kernel is known, which may or may not be reasonable depending on applications.

Technique	Suitable Artifact Type	Artifact Reduction Performance	Complexity	Memory	Robustness
Adaptive Filtering	blocking ringing	moderate	low	low	moderate
AC Prediction	blocking	moderate	low	low	poor
Wavelet De-noising	blocking	good	high	high	moderate
LMMSE	blocking	moderate	moderate	moderate	moderate
MAP	blocking, ringing	good	high	high	good
POCS	blocking, ringing*	good**	high	high	good

* with additional constraints

** if the constraint is properly formulated

Table 2.1: Comparison of postprocessing techniques.

Technique	Blind Inverse Halftoning	Noise Reduction Performance	Complexity	Memory	Robustness
MAP	No	good	high	high	good
POCS	No	good*	high	high	good
Wavelet	Yes	good	median	high	moderate
Adaptive filtering	Yes	good	low	low	moderate
Nonlinear Denoising	Yes	good	low	low	moderate

* if the constraint is properly formulated

Table 2.2: Comparison of inverse halftoning techniques.

Chapter 3

Robust Estimation and Artifact Reduction

Robust estimation, which describes how to fit a large group of data to a predetermined cost function, has been applied to computer vision and image restoration applications in the context of segmentation, surface fitting, and noise reduction for many years. In this work, we apply robust estimation to compression artifact reduction, particularly for the smoothing of blocking and ringing artifacts that exhibit in the DCT- and wavelet-based compression.

The proposed postprocessing algorithm is inspired by the excellent artifact-reduction performance of MAP-based techniques (or, more generally, Bayesian-based restoration). However, they are not practical for real-time applications due to the computational complexity. To reduce the complexity, we introduce a non-iterative filtering method to obtain the suboptimal solution. Basically, our approach includes the following three steps.

- Derive a generic model that accounts for both distortion and image smoothness for various type of compression schemes.
- Identify parameters of the model according to dominating artifacts.
- Obtain a suboptimal solution of this model by using the nonlinear filter.

This chapter is organized as follows. The theoretical background is first provided in Section 3.1. Then, a generic model used to solve the artifact reduction problem is given in Section 3.2. A nonlinear filtering algorithm is proposed to get the suboptimal solution. The filter and its efficiency are detailed in Section 3.3. The application to

inverse halftoning problem is described in Section 3.4 followed by concluding remarks in Section 3.5.

3.1 Background

3.1.1 Robust Estimation

It is well-known that least-squares estimators are sensitive to the presence of outliers that violate the underlying assumption about the distribution of data. Many robust estimation techniques have been proposed in the last three decades to overcome this problem. One popular robust estimation technique is known as the M-estimator. To explain the idea of M-estimator, we consider the following simple example. Let v_1, v_2, $\cdots$, v_k be k samples and r_i with $1 \leq i \leq k$ be the residual of the ith samples (i.e. the difference between the ith samples and its fitted value). The standard least-squares estimation minimizes $\sum_i r_i^2$, which is however sensitive to the outliers. The reason is that the outliers contribute "too much" to the overall solution. To handle this problem, the M-estimator tries to reduce the effect of outliers by replacing the quadratic residual function r_i^2 with other potential functions $\rho(\cdot)$ (or called the cost functions). In other words, we want to solve a new problem of the following form:

$$\min \sum_i \rho(r_i). \tag{3.1}$$

The influence function $\psi(\cdot)$, which is defined as the first derivative of $\rho(\cdot)$, measures the influence of a datum on the value of estimate. The robustness of an M-estimator is determined by the shape of its influence function. If we define a weight function $\omega(x) = \psi(x)/x$, Eq. (3.1) can be rewritten as

$$\min \sum_i \omega(r_i^{(k-1)}) r_i^2, \tag{3.2}$$

where $\omega(r_i^{(k-1)})$ indicates the weight in the kth iteration. This shows that the M-estimator is actually equivalent to the weighted least-squares formulation.

It is often required that the potential function be a symmetric, positive-definite function with a unique minimum at zero and an increasing slope less than that of the

quadratic function. Any function that satisfies these requirements can be a potential function to be used in the M-estimator. However, the M-estimator is robust only when its influence function is bounded and $\rho(\cdot)$ should contain a strictly convex part. A few commonly used potential functions are listed Table 3.1 and they are graphically depicted in Figure 3.1.

Type	$\rho(x)$
L^2	x^2
L^1	$\|x\|$
"Fair"	$\gamma\|x\| - \gamma^2 \log(1 + \frac{\|x\|}{\gamma})$
Huber	$\rho(x) = \begin{cases} x^2 & ,\|x\| \leq \gamma \\ \gamma^2 + 2\gamma(\|x\| - \gamma) & ,\|x\| > \gamma \end{cases}$
L^γ	$\|x\|^\gamma$,$1 \leq \gamma \leq 2$
Truncated L^2	$\min\{\gamma x^2, 1\}$
Lorentzian	$\log(1 + \frac{1}{2}(\frac{x}{\gamma})^2)$.
Tukey	$\rho(x) = \begin{cases} \frac{\gamma^2}{6}(1 - [1 - (\frac{x}{\gamma})^2]^3) & ,\|x\| \leq \gamma \\ \frac{\gamma^2}{6} & ,\|x\| > \gamma \end{cases}$

Table 3.1: Commonly Used M-estimators.

The above potential functions differ in convexity, boundedness, differentiability, etc. It is important to choose a potential function so that the estimate will agree with the prior expectations and requirements. The are few comments about these functions.

- L^2 estimators are not robust since the influence function is not bounded.
- L^1 estimators are not robust since the absolute value is not strictly convex.
- L^γ is robust only if $1 < \gamma < 2$, particularly with $\gamma = 1.2$. However, difficulties in computation may be encountered.
- Huber function provides a good trade-off between robustness and complexity and is therefore recommended for almost all situations. This function still suffers from a minor numerical problem due to its discontinuous second derivative at $\pm\gamma$.

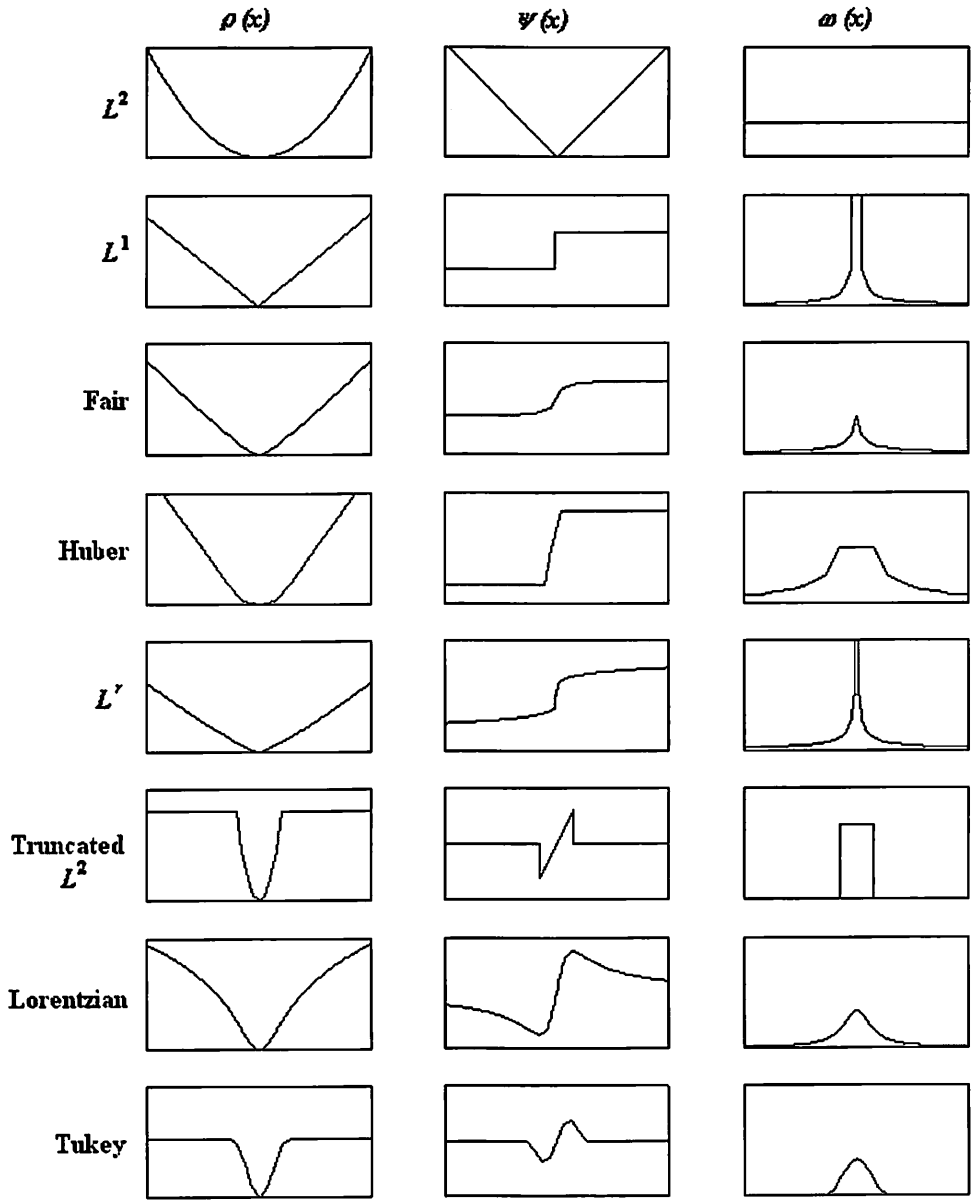


Figure 3.1: The qualitative shapes of some potential functions

- The influence functions of L^γ, "Fair", and Huber are bounded. The line process models, such as Truncated L^2 and Lorentzian, reduce or even cut off the influence from a large gross error and can avoid the oversmoothing problem. However, these functions contain concave parts which may trap the solution to a local minimum. To avoid this situation, the computation of these estimators usually requires the stochastic relaxation technique.

3.1.2 Generic Image Model

To solve the artifact-reduction problem with robust estimation framework, a probabilistic model is required to characterize both source signal and the noise. Markov random field (MRF) theory that provides a theoretic basis for modeling the joint probability to the prior contextual constraints is a commonly adopted model for many vision and image restoration problems. It is often incorporated with MAP estimation criteria and Gibbs distribution assumption to form a cost function for optimizing the solution of a given the problem. Based on this approach, our problem can be formulated as follows.

To reconstruct the artifact-free image $\mathbf{u}$ from a decompressed image $\mathbf{d}$ that contains compression artifacts, the MAP estimate $\hat{\mathbf{u}}$ is given by

$$\hat{\mathbf{u}} = \arg\max_{\mathbf{u}}\{\log Pr(\mathbf{u}|\mathbf{d})\}, \tag{3.3}$$

$$= \arg\max_{\mathbf{u}}\{\log Pr(\mathbf{u},\mathbf{d}) - \log Pr(\mathbf{d})\}, \tag{3.4}$$

where $\log Pr(\mathbf{d})$ can be dropped in determining $\hat{\mathbf{u}}$ since it is a constant indepedent of $\mathbf{u}$. With the assumption of the Gibbs random field, we have

$$Pr(\mathbf{u},\mathbf{d}) = k\exp\{-E(\mathbf{u},\mathbf{d})\},$$

where k is a normalization constant and $E(\mathbf{u},\mathbf{d})$ is an energy function. The maximization of $Pr(\mathbf{u},\mathbf{d})$ is equivalent to the minimization of $E(\mathbf{u},\mathbf{d})$, i.e.

$$\hat{\mathbf{u}} = \arg\min_{\mathbf{u}} E(\mathbf{u},\mathbf{d}). \tag{3.5}$$

The energy function in the Gibbs random field regularizes the smoothness of neighbor pixels in an image, which is analogous to the potential function that constrains data smoothness in statistics. Thus, the image smoothing problem can be solved by using the M-estimator if we formulate the energy function of the Gibbs random field as one the potential function listed in Table 3.1. To be more explicit, assume that the decompressed image is an $m \times n$ image of sites S. Then, there is a set of neighbors $c \in C$ for each site (or pixel) $s \in S$. For the 1st order MRF, C can be defined as the four-nearest-neighbor (or the eight-nearest-neighbor) clique. Theoretically, there is no need to restrict the clique type, or the order of MRF. In fact, higher order MRF is able to model more complicated image details rather than the 1st order, which can only model the step edge. However, for the complexity reason, the 1st order MRF with 4- or 8-neighbor clique is often used. Therefore, the energy function in Eq. (3.5) can be expressed as

$$E(\mathbf{u}, \mathbf{d}) = E_d(\mathbf{u}, \mathbf{d}) + E_s(\mathbf{u}), \tag{3.6}$$

$$= E_d(\mathbf{u}, \mathbf{d}) + \sum_{s \in S} \sum_{c \in C} \rho_s(u_s - u_c), \tag{3.7}$$

where the data fidelity term E_d ensures that the estimate is faithful to the original image while the spatial smoothness term E_s is used to model the transition between adjacent pixels. The potential function $\rho_s(x)$ characterizes the behavior of the estimator, or equivalently the smoothing capability. The optimal solution of Eq. (3.5) provides the best trade-off between image soothing and detail preserving. In the next section, we explain how to model E_d and select $\rho_s(x)$ for a particular compression scheme in detail.

3.2 Artifact Reduction via Robust Estimation

The formulation in Eq. (3.7) is general for many applications. For reducing compression artifacts, there are two major issues. One is how to select potential functions that adapt to the image content. The other is how to impose noise constraints according to the severity of artifacts.

3.2.1 Selection of Potential Functions

It is important to choose a potential function so that the result of the M-estimator will agree with the prior expectations and requirements. For example, if v_1, v_2, $\cdots$, v_k represent the intensities of adjacent pixels in an image, the result of the M-estimator with a quadratic potential function is equivalent to the output of the mean filter. When L^1 potential function is applied, the result is equivalent to output of the median filter. It is well known that the mean filter achieves a strong smoothing effect while the median filter removes the outliers effectively. The Huber, L^γ and Lorentzian are proposed to achieve a better tradeoff between noise smoothing and detail preservation. All the above potential functions have no cut-off point in their influence functions. That is, they did not reduce the influence from a large error completely. For some applications that require to recover the edges, a potential function that contains a strictly nonconvex part, such as the truncated L^2, Lorentzian and Tukey, must be used. To be more general, the necessary condition for a potential function to preserve or recover an edge is:

$$\lim_{x \to \infty} |\psi(x)| \leq C, \tag{3.8}$$

where C is a constant. That is, the influence function should be bounded. When $C = 0$, smoothing is prohibited at the large discontinuities so that edge can be recovered. While $C > 0$, limited (bounded) smoothing is allowed so that the edge is preserved. A detailed comparison about several potential functions and their applicability to image smoothing and edge recovery can be found in [39, 29].

For applying the robust estimation technique to compression artifact reduction in DCT-based coding, The potential function $\rho_s(x)$ is preferred to be L^2 in the planar regions where severe blocking artifacts are usually perceived, and for the remaining part of the image, the Huber function (or other function of a similar type) can be used to achieve both smoothing and detail preserving. Since the Huber function is quadratic when $|x| \leq \gamma$, another alternative is to apply the Huber function for the entire image with a larger γ in the planar region and a smaller γ in other areas.

For wavelet-based coding, the ringing artifacts are associated with edges and often with a smaller energy. They are ideal for the robust estimation framework.

That is, any potential function which is able to smooth small perturbations (artifacts) while reducing the influence from large gross errors (edges) can be applied to wavelet-based compression. In general, "Fair" and Huber functions are good choices for natural images. The non-convex potential functions such as Loretzian is better for synthetic images that are often composed of homogeneous regions separated by edges.

3.2.2 Modeling the Noise Constraints

The energy function defined in Eq. (3.6) consists of two terms, $E_s(\mathbf{u})$ and $E_d(\mathbf{u}, \mathbf{d})$. As discussed previously, the robust techniques can be used in the term $E_s(\mathbf{u})$ to model the spatial relationship among a neighboring set of pixels in images. However, a natural image contains all kinds of features such as texture, which may violate the smoothness assumption imposed by $E_s(\mathbf{u})$. Therefore, an addition noise constraint, $E_d(\mathbf{u}, \mathbf{d})$, is necessary to regulate the distortion introduced by the smoothing estimators.

The data fidelity term, $E_d(\mathbf{u}, \mathbf{d})$, corresponds to $\log Pr(\mathbf{d}|\mathbf{u})$ in the MAP estimation criterion. It is used to model the additive noise. In other applications, such as the Gassian noise removal or the surface fitting problems, noise is often assumed to be with the i.i.d. Gaussian distribution. This results in a quadratic form in the data fidelity term. However, the quadratic cost function is not an accurate model of noise introduced by image compression due to the following reasons.

- The compression noise is resulted from the quantization of transform coefficients. Thus, its appearance in the space domain is affected by several factors such as transform filters and quantization schemes, *etc.*
- The compression noise is signal dependent. In image compression, visually important information is coded with more accuracy and, consequently, noise is affected by content even with the same compression method.
- Compression noise may be structured. It occurs when an image is divided into blocks and each block is quantized separately.

It is clear that modeling the compression noise in the space domain is a difficult problem. Stevenson [48] proposed a transform domain projection method to

enforce the data fidelity without modeling the noise explicitly. He used the quantized transform domain representation $\mathbf{y}$ to replace the observation $\mathbf{d}$ in the original MAP formulation. The representation $\mathbf{y}$ is defined by applying transform H and quantization Q to the original image $\mathbf{x}$, i.e.

$$\mathbf{y} = Q[H\mathbf{x}], \tag{3.9}$$

and its relation to $\mathbf{d}$ is

$$\mathbf{d} = H^{-1}Q^{-1}[\mathbf{y}]. \tag{3.10}$$

The data fidelity is ensured by only adopting the estimate that is in the solution space U_T, where U_T is defined as a set of images which compress to $\mathbf{y}$. (i.e., $U_T = \{\mathbf{u} : \mathbf{y} = Q[H\mathbf{u}]\}$). With this assumption, the MAP estimation in Eq. (3.5) can be reformulated as

$$\hat{\mathbf{u}} = \arg\min_{u \in U_T} \sum_{s \in S} \sum_{c \in C} \rho_s(u_s - u_c). \tag{3.11}$$

It is more appropriate to call this approach the *constrained minimization* technique rather than the *joint optimization* in the original MAP estimation. It is desirable to avoid the problem of compression noise modeling. The major drawback of this approach is that it requires a pair of transforms (e.g. DCT/IDCT) in each iteration of gradient search. The computational cost is too high for real-time applications.

In this work, we follow Stevenson's constrained minimization approach, but define the noise constraints in the space domain to avoid the use of DCT/IDCT transforms. Let the estimate of particular site be $\hat{u}_s$. The standard decoded value d_s will be updated only if

$$|\hat{u}_s - d_s| \leq \sigma. \tag{3.12}$$

That is, we only search the estimates in the solution space U_S so that the Eq. (3.5) can be written as a constrained optimization problem

$$\hat{\mathbf{u}} = \arg\min_{u \in U_S} \sum_{s \in S} \sum_{c \in C} \rho_s(u_s - u_c), \tag{3.13}$$

where $U_S = \{\mathbf{u} : |\mathbf{u} - \mathbf{d}| \leq \sigma\}$. The idea behind this approach is that although we do not know the exact probabilistic model about the compression noise, it is assumed to be bounded by a certain value since lossy image compression intends to

reduce the distortion as much as possible. With this assumption about noise, the average error energy introduced by the smoothing estimator is less than σ^2. The value of σ determines the trade-off between the artifact smoothing and the distortion introduced by the estimator.

3.3 Approximation of Global Solution

The model proposed in the previous section unifies the robust estimation with MAP probability criteria so that artifact reduction and image enhancement can be solved in one framework. In previous work, the solution of this constrained optimization problem was often computed with deterministic relaxation techniques such as gradient descent search and iterated conditional mode(ICM), or stochastic-based simulated annealing. However, all these methods are iterative and impractical for real-time applications. In addition, searching for global optimum solution inherits a large memory requirement, since the entire image has to be buffered. To reduce the computational complexity as well as the storage, a nonlinear filtering technique is proposed to find the approximating global minimum with a low computational cost.

3.3.1 Nonlinear Filtering

Here, we propose two nonlinear filtering techniques which can generate an approximate global minimum. The performances of nonlinear filtering and gradient projection are also compared.

Filter-A: Given a set of samples x_1, x_2, ..., x_N and the potential function ρ_s as discussed in the previous section, the filter output is defined as

$$\hat{x}_A = \arg\min_{x_j} \sum_{i=1}^{N} \rho_s(x_i - x_j). \tag{3.14}$$

Instead of searching the global minimum over the entire space, this filter perform the optimization over the finite set of samples within a local window. When the potential function is the absolute value of residues. This filter becomes a median filter. The are few properties for this filter:

- With N samples in the filter, the complexity of of this filter is always on the order $O(N^2)$ or less, which is low compared with iterative gradient descent search that complexity is unpredictable.

- For any potential function $\rho(\cdot)$, the associated filter output $\hat{x}_A$ is bounded by

$$x_{(2)} \leq \hat{x}_A \leq x_{(N-1)} \tag{3.15}$$

  where the subscript (i) denotes the ith order statistic of the sample.

Filter-B: Given a set of samples x_1, x_2, ..., x_N and the potential function ρ_s. Let

$$r_j = \sum_{i=1}^{N} \rho_s(x_i - x_j).$$

Let us rank all r_j, $1 \leq j \leq N$, and values x_j associated with the k smallest r_j are denoted by a new array a_j, $1 \leq j \leq k$. The value k is taken less than $\lceil (1+N)/2 \rceil$. Finally, the resulting filter output is

$$\hat{x}_B = \frac{1}{k} \sum_{i=1}^{k} a_j \tag{3.16}$$

Again, the rank process rejects outliers. Then, we approximate the global minimum by taking the mean of local sample points that are identified as inliers.

3.3.2 Justification

The efficiency of proposed filters can be explained with a simple example. Let $\{x_1, ..., x_9\} = \{35, 35, 34, 30, 27, 24, 25, 28, 30\}$ be 9 sample points. The M-estimate with the Lorentzian potential function ($\gamma = 0.5$) is 31. The results of **Filter-A** and **Filter-B** ($k = 2$) are both 30. Optimization of this cost function cannot be calculated by the gradient descent method, since it can be trapped at the local minimum. **Filter-A** approximates the global minimum by simply taking the sample with the lowest sum of costs, while **Filter-B** averages the K samples that have lower sums of costs. The relation between proposed filters and the M-estimator is graphically depicted in Figure 3.2.

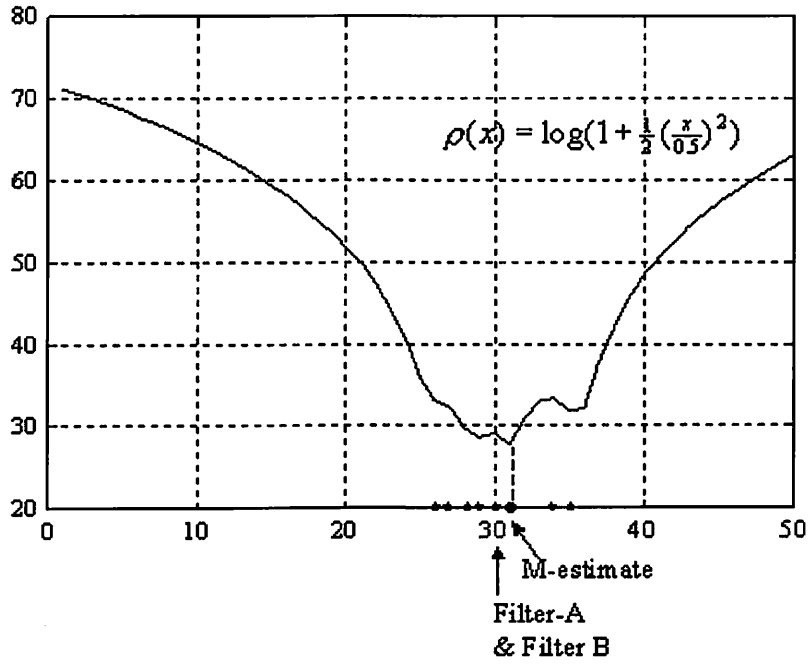


Figure 3.2: Illustration of the relation between proposed filters and M-estimator.

To compare the performance visually, we apply the proposed filters to a 1-D signal, which is extracted from the 300th line of the compressed "Target" image. Fig. 3.3 shows the result smoothed by proposed filters and the gradient descent approach. We compare results of one iteration of nonlinear filtering and 100 iterations of the gradient descent search. The length for both Filters A and B is 5. The Huber function is used as the cost function. It is clear that the nonlinear filter is more efficient in outlier rejection as smoothing the data. Filter B is more robust than Filter A. Fig. 3.4 is the image representation of the filtering result. The gradient projection concurrently optimizes all N samples and usually converges slower than the proposed two filters, because a large gradient (contributed from the outlier) will result in a small step size to reduce the convergence rate. On the contrary, the nonlinear approach searches the local neighborhood for the best solution that results in the smallest energy. Due to the high correlation in neighboring pixels of a natural image, the nonlinear filter is able to approximate the optimal solution in an efficient way. The length N of the filters is a design issue. The longer the filter length, the smoother the estimate. A longer filter length produces a more robust result with respect to small local variations at the price of a higher computational cost.

3.4 Halftone Noise Reduction via Nonlinear Filtering

The nonlinear filters described in Section 3.3.1 can also be applied to the inverse halftoning problem in reducing halftone noise. However, the proposed nonlinear filter is a *location filter*. That is, the output is selected from the neighboring pixel values. Thus, it cannot be directly applied to the smoothing of the halftone image. To overcome this problem, we can apply a simple lowpass filter to get an initial estimate of continuous-tone image and then use nonlinear filtering to reduce halftone noise in the lowpass filtered halftone image.

Fig. 3.5 shows an error-diffused image obtained by the Floyed-Steinberg kernel, and the result of the lowpass filtered image. The binary patterns in the error-diffused image are spread out spatially, and the resulting image consists of small clusters of "peaks and valleys" over the grayscale background. Since this grainy noise is generally several pixels wide, it cannot be removed by further lowpass filtering due to its excessive blurring to the resulting image. On the other hand, the median filter dose not work well in this situation either. The proposed robust nonlinear filter fits this job, since it is designed to smooth small perturbation without blurring edges. We propose a cascade filter as shown in Fig. 3.6 to reduce halftone noise in the inverse halftoning process.

3.5 Conclusion

In this chapter, we proposed a new postprocessing technique for compression artifacts reduction. We first formulate the problem in a stochastic framework as the restoration approach does, and then find a sub-optimal solution by using a nonlinear filtering technique of low complexity. The data fidelity constraints are imposed purely in the space domain to avoid the computational cost associated with the forward and inverse transforms. This hybrid restoration/filtering approach can avoid the numerical difficulty involved in the optimization of a non-convex potential function, and give us more freedom in selecting different types of potential functions to

achieve the desired postprocessing properties such as noise smoothing, edge recovery, etc. In addition, the nonlinear filters based on robust estimation can smooth the noise while preserve the quality of edges without performing edge extraction explicitly. The robust nonlinear filtering technique can also be applied to inverse halftoning for the smoothing of halftone noise.

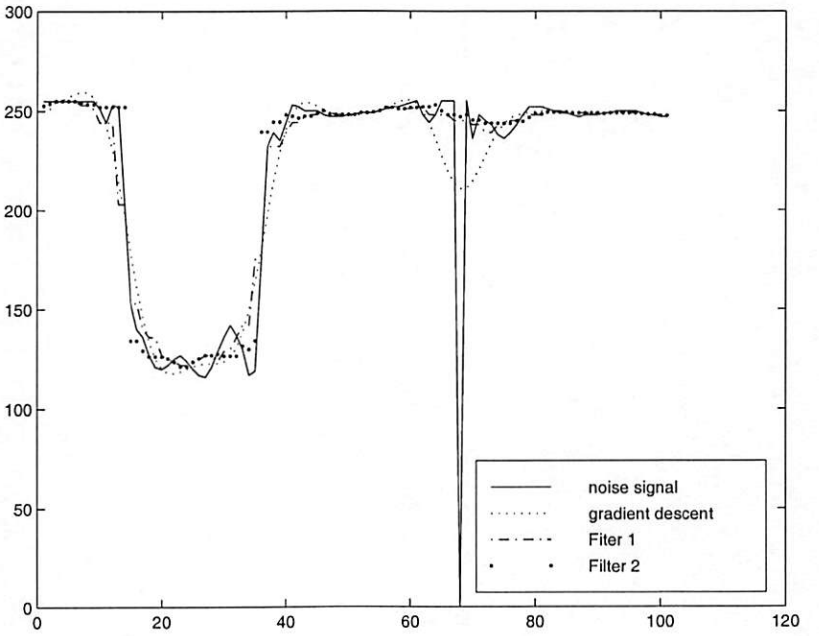


Figure 3.3: 1-D signal filtered by the proposed nonlinear filters.

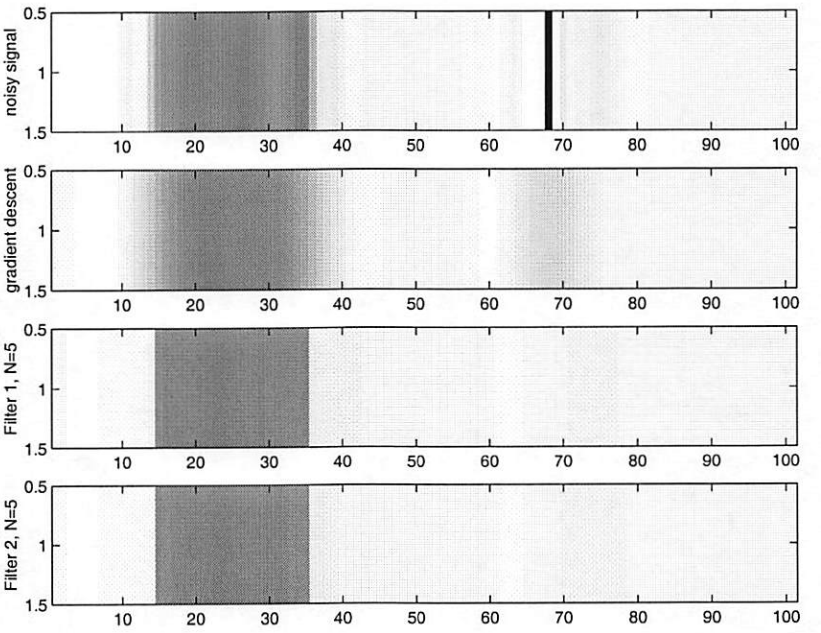
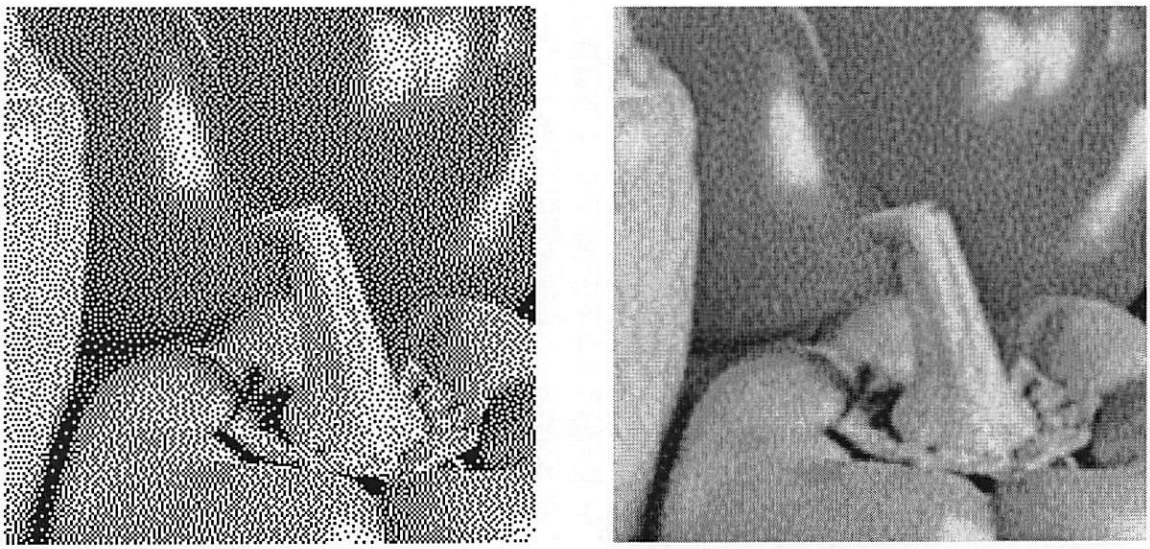


Figure 3.4: Image representation of filter results.



(a) (b)

Figure 3.5: (a) Floyd-Steinberg error-diffused halftone image. (b) lowpass filtered halftone.

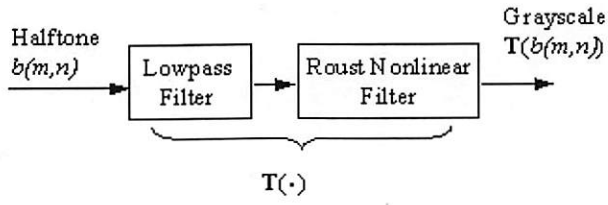


Figure 3.6: The proposed filter for inverse halftoning.

Chapter 4

Postprocessing of Wavelet-Coded Images

A direct application of the proposed technique is postprocessing of wavelet-coded images. As mentioned before, the primary artifact of the low-bit-rate wavelet coding is ringing artifacts around sharp edges, and it has been demonstrated in the last chapter that nonlinear filters based on robust estimation can smooth the small perturbation around discontinuity without blurring it. This is ideal in reducing compression artifacts resulted from wavelet-based compression. In this chapter, we perform our experiments on several wavelet codecs, proposed by JPEG-2000 participants. We also compare our results with other work in this area.

4.1 Introduction

The wavelet-based coding scheme has advantages over the block DCT-based coding method in terms of the rate-distortion tradeoff performance and visibility of coding artifacts, and has been adopted as the baseline scheme of the new JPEG-2000 standard. However, reconstructed images with wavelet coding still suffer from various coding artifacts when the compression ratio becomes too high. The appearance and severity of coding artifacts depend on several factors including transform filters, coding bit rates, quantization schemes, and image contents. Typical artifacts in wavelet-based coding include the ringing effect, graininess, and blotchiness. Among these artifacts, the ringing effect around sharp edges caused by the coarse quantization of high frequency components is the most objectionable. When coded at very low bit rates, all subbands are severely quantized. The loss of low and mid-frequency

components results in blotchiness in homogenous regions and granular noise in texture regions. In this section, we apply the postprocessing technique described in Section 3.2 to various wavelet-based coding algorithms which are proposed by the JPEG 2000 participants.

Work in postprocessing for subband/wavelet coding was first reported by Rourke and Stevenson in [49]. They applied a constrained optimization technique as formulated in Eq. (3.11) for ringing artifacts reduction. A similar approach was proposed by Luo *et al* in [32]. Li and Kuo [28] pointed out that the approach proposed in [49] was not effective to remove the wide spread ringing artifact in wavelet coding. A multiscale postprocessing were proposed in [28] to overcome this problem. Postprocessing operations applied at coarser scales reduce long-range artifacts while preserving the sharpness of edges at the fine scale. All these methods are essentially restoration approach which suffer from high computational complexity.

4.2 Filter Description

To reduce ringing artifacts in low-bit-rate wavelet coding, **Filter-A** described in Section 3.3.1 is applied to the entire decompressed image. That is, for a given central pixel x, the corresponding result of **Filter-A** is denoted with $\hat{x}$. Then, the final artifact-free estimate y, which is used to replace x, is obtained by

$$y = x + \mathbf{c}(d, Th1), \tag{4.1}$$

where $d = \hat{x} - x$, and $\mathbf{c}(d, Th1)$ is a clipping function to ensure data fidelity and defined as

$$\mathbf{c}(d, Th1) = Sign(d) \cdot Max(0, abs(d) - Max(0, 2(abs(d) - Th1))). \tag{4.2}$$

The clipping function controls the amount of distortion introduced by the smoothing operation and prevents features such as isolated pixels and the corner from being removed by using nonlinear filters. The shape of $\mathbf{c}(d, Th1)$ is depicted in Figure 4.1. It is clear that the result of nonlinear filtering is effective only when d is smaller than $2Th1$. The value $Th1$ determines the amount of smoothing and should

vary according to the quantization parameter. The up-down shape of the clipping function is used to ensure the continuity of the output value y when $\hat{x}$ has a small perturbation abound the value $Th1$.

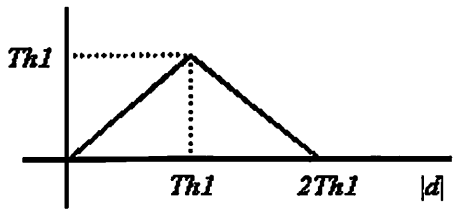


Figure 4.1: The clipping function.

Filter-A performs the optimization over the finite set of local samples to obtain the estimate. The set of neighboring pixels can be defined to account for the geometric relationship as defined in many order statistics filters. Examples of filtering windows are given in Figure 4.2.

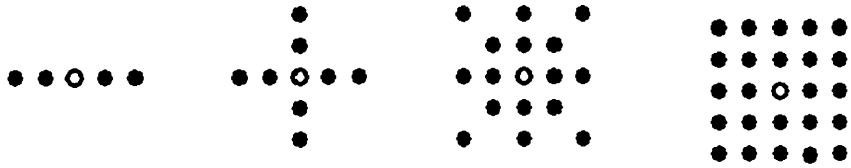


Figure 4.2: Several window shapes for nonlinear filters.

4.3 Experimental Results

We apply the proposed postprocessing filter to four JPEG-2000 test images, i.e. Target, Bike, Woman and Cafe, at various bit rates (0.25, 0.125, and 0.0625bpp) with different wavelet-based coding algorithms, including JPEG-2000 VM/TCQ, JPEG-2000 VM/SQ, USC/MTWC, and TI/EZW. These test images are defined to embrace images of different types such as natural (Woman, Cafe), synthetic (Target), and hybrid (Bike). The sizes of these images are 2048×2560 for Woman, Cafe and Bike, and 512×512 for Target. In our experiments, a plus-shaped window with size 9 is used to find the estimate of the central pixel. Threshold value $Th1$ is set to

12, 10, 8 at bit rates equal to 0.0625, 0.125, 0.25, respectively. For all test images, *Truncated* L^2 is used and γ is set to 1/256. The PSNR performance is shown in Tables 4.1 - 4.2. Although PSNR cannot reflect visual quality, it still indicates how much distortion introduced by the proposed technique. We have the following observations from PSNR results.

- For images contain more edges such as Target and Bike, the PSNR improvement is substantial. This is reasonable since the proposed method is efficient in reducing ringing artifacts which are often associate with edges.
- The Cafe and Woman are natural images. The former contains more texture while the later has more homogeneous regions. Although ringing effects are less visible in these two cases due to the masking effect, the proposed filter still performs well and no significant feature is smoothed out after filtering.
- In the case of severe artifacts, the proposed filter may be applied iteratively. Usually, good results can be produced after 2 or 3 iterations.
- The distortion introduced by the proposed filter is bounded, since PSNR converges after a few iterations. The number of iterations required for convergence depends on the potential function and its parameter.

For subjective evaluation, the Target image encoded by JPEG-2000 VM/SQ at 0.125bpp and its postprocessing result are shown in Figs. 4.3– 4.4. JPEG-2000 also provides the visual-weighted quantization, which is the encoder-based quality enhancement technique. The result of JPEG-2000 with visual-weighted coding is shown in Figure 4.5. The artifacts are still noticeable even with the encoder-based enhancement when coded at such a low rate. The postprocessing of this image is shown in Figure 4.6. We see that ringing artifacts are smoothed while sharp discontinuities and other features such as isolated pixels and single pixel-wised lines are still well preserved by the proposed filters.

We also compare our method with other previous work. Figure 4.7(a) shows a synthetic image compressed by the wavelet codec at 0.25 bpp. Figure 4.7(b) is the result by using the Huber MRF (HMRF) with a single scale. Figure 4.7(c) is obtained by using the multiscale HMRF method described in [28]. Figure 4.7(d)

shows the result of the proposed filtering technique with a nonconvex potential function (Truncated L^2). This potential function has a strictly nonconvex parts that can preserve edges well. This property is clearly observed from the sharp discontinuities in the reconstructed image. Figure 4.8 is an example of a natural image. Figure 4.8(d) is the result of the nonconvex potential function which preserves sharper edges than the HMRF-based approach. As shown in these examples, we see that the proposed nonlinear filtering can quickly converges to the artifact-free estimate. It is still possible to achieve both artifact reduction and edge enhancement with nonconvex potential functions.

4.4 Conclusion

Experimental results demonstrated the excellent artifact reduction performance of the proposed robust filters. In addition, the proposed method works efficiently regardless of the variation of coding schemes. There is one issue remaining for future study, i.e. the exact relation between the quantization scale and parameters in the noise constraint and the potential function. Although the proposed filters are robust with the fixed parameter in the noise constraint, we believe that, with a more accurate noise model, adaptivity and robustness of the current scheme can be further improved.

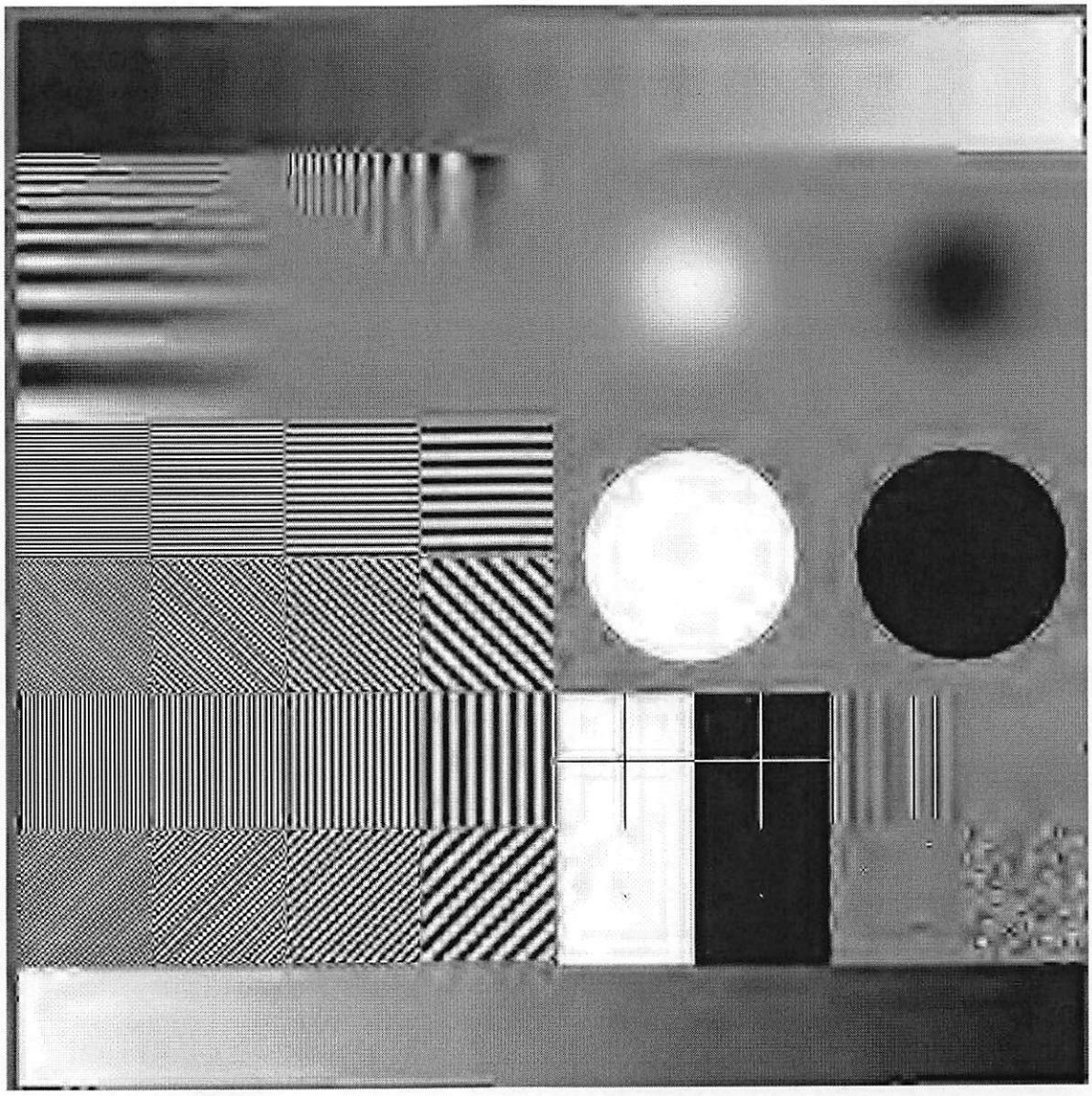


Figure 4.3: Target encoded with JPEG-2000 VM/SQ at 0.125bpp.

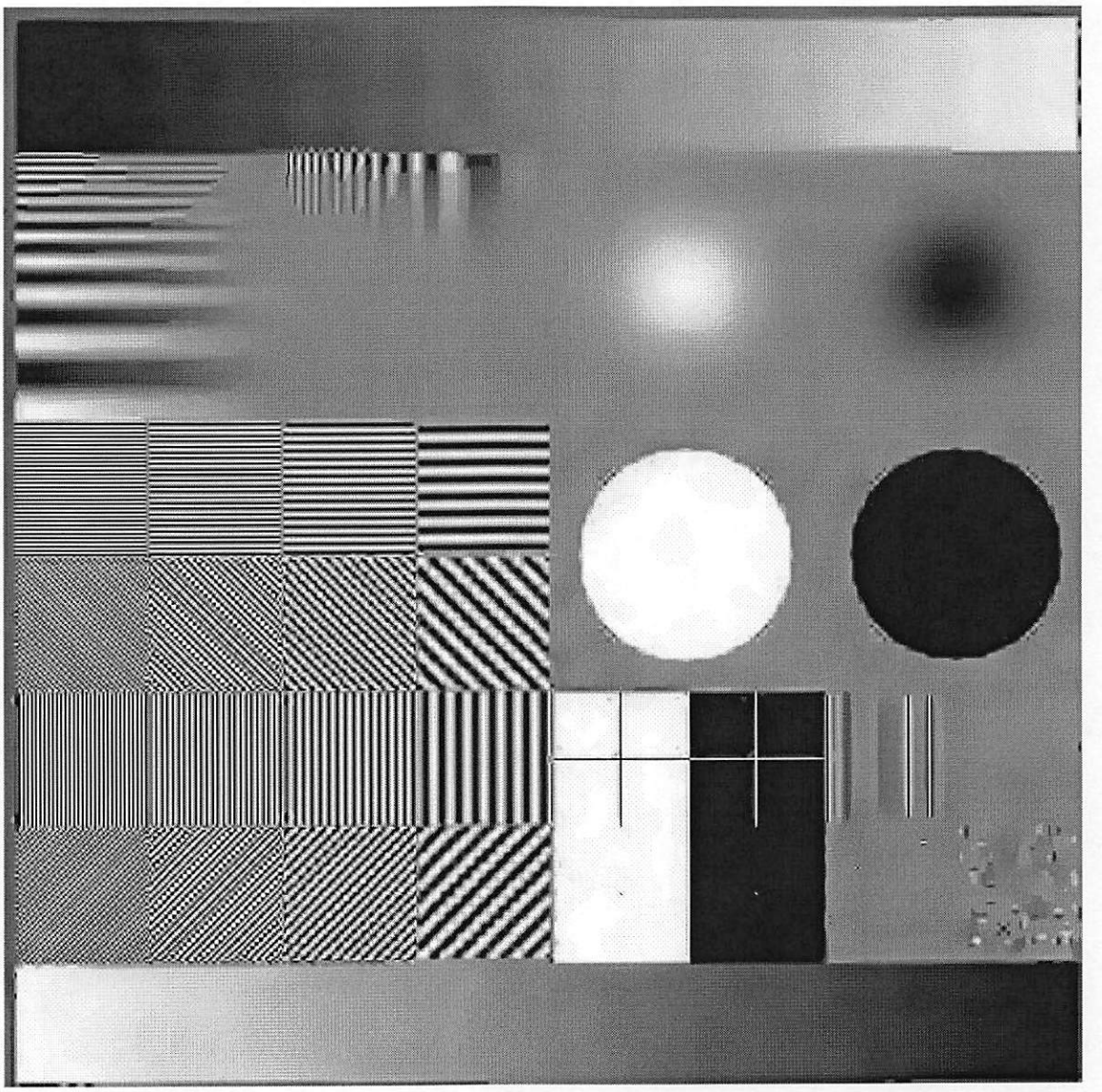


Figure 4.4: Target encoded with JPEG-2000 VM/SQ at 0.125bpp with postprocessing.

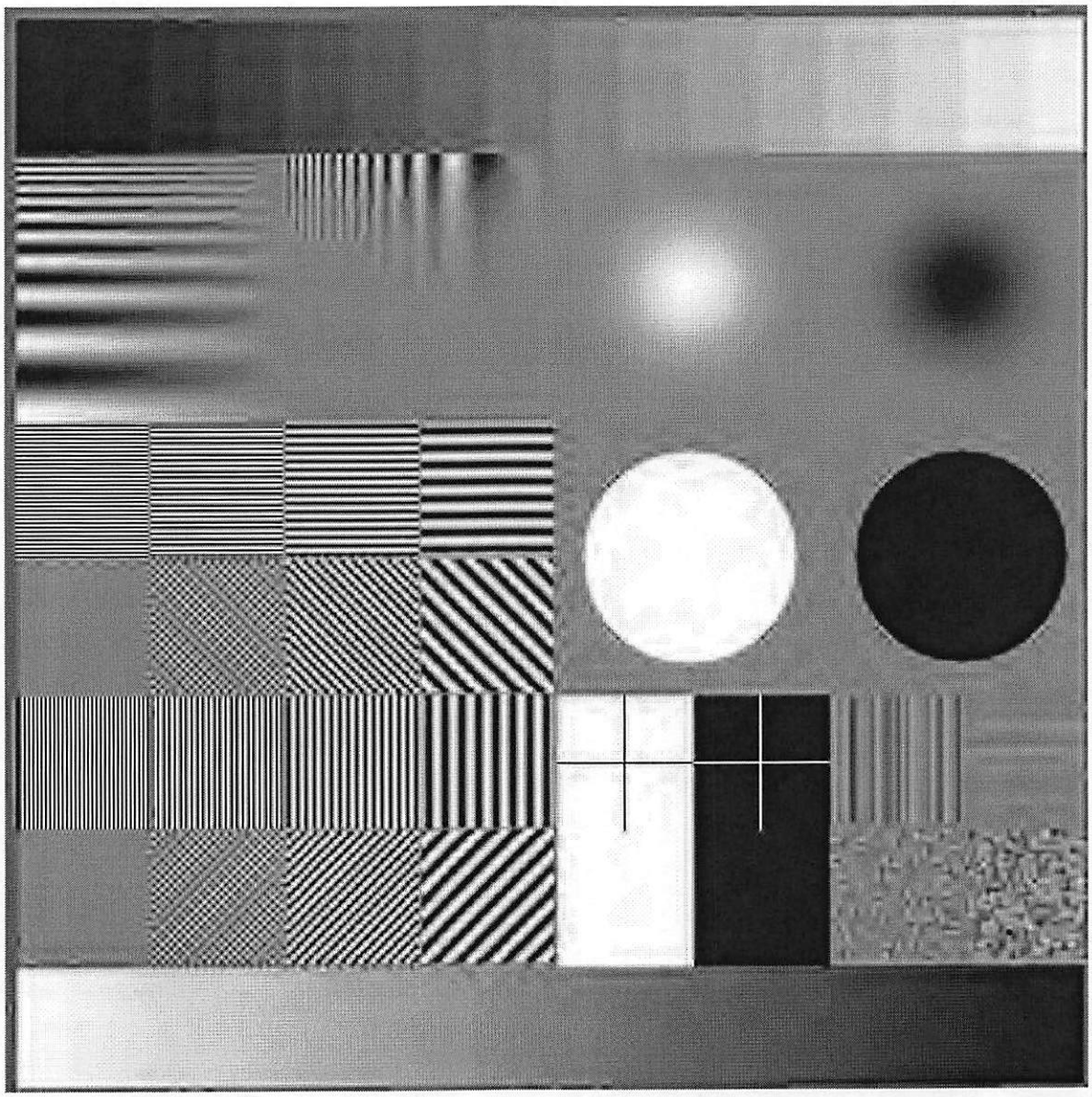


Figure 4.5: Target encoded with JPEG-2000 VM/SQ with visual weighting at 0.125bpp.

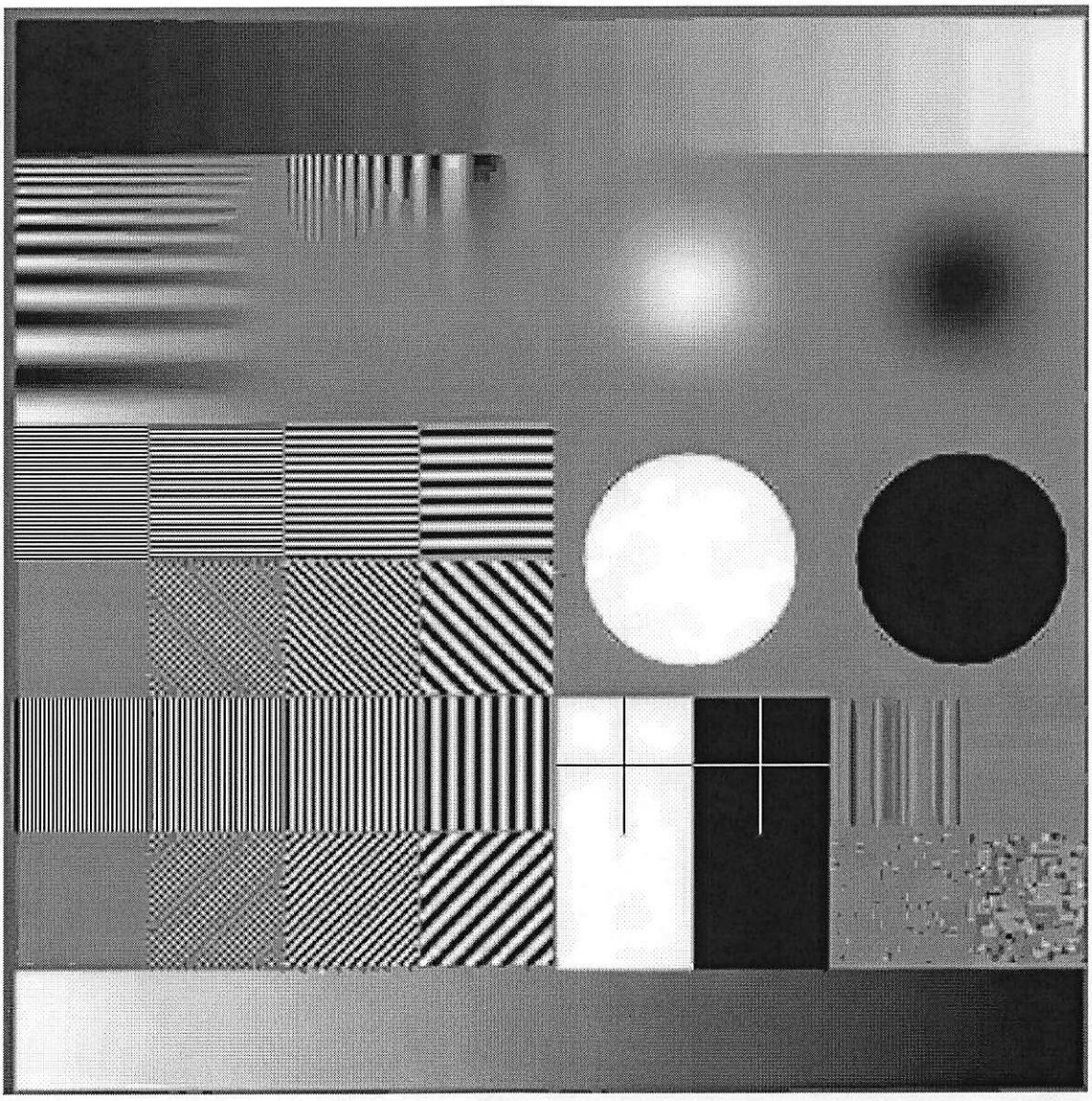


Figure 4.6: Postprocessing of Target encoded with JPEG-2000 VM/SQ with visual weighting at 0.125bpp.

Target, size 512 × 512					
			PSNR with postprocessing		
Bit rate	Codec	PSNR	1*	2	3
0.25 bpp	VM TCQ	25.42	25.55	25.58	25.59
	VM SQ	27.71	27.88	27.93	27.93
	MTWC	27.50	27.64	27.65	27.63
0.125 bpp	VM TCQ	21.09	21.16	21.19	21.20
	VM SQ	22.35	22.42	22.45	22.46
	MTWC	22.29	22.41	22.50	22.54
0.0625 bpp	VM TCQ	17.76	17.78	17.81	17.83
	VM SQ	20.01	20.05	20.10	20.13
	MTWC	20.36	20.42	20.48	20.51
Bike, size 2048 × 2560					
			PSNR with postprocessing		
Bit rate	Codec	PSNR	1*	2	3
0.25 bpp	VM TCQ	28.87	29.02	28.94	28.82
	VM SQ	29.31	29.45	29.35	29.21
	MTWC	29.49	29.72	29.67	29.55
0.125 bpp	VM TCQ	25.64	25.81	25.84	25.81
	VM SQ	25.76	25.92	25.93	25.91
	MTWC	26.06	26.23	26.26	26.24
0.0625 bpp	VM TCQ	22.95	23.07	23.12	23.13
	VM SQ	23.13	23.25	23.29	23.30
	MTWC	23.26	23.39	23.44	23.45

* number of iterations

Table 4.1: The PSNR performance comparison for the Target and Bike image with several wavelet codecs.

Cafe, size 2048 × 2560					
			PSNR with postprocessing		
Bit rate	Codec	PSNR	1	2	3
0.25 bpp	VM TCQ	23.14	23.23	23.23	23.20
	VM SQ	23.20	23.29	23.29	23.26
	TI EZW	22.97	23.07	23.09	23.08
0.125 bpp	VM TCQ	20.68	20.74	20.74	20.73
	VM SQ	20.76	20.81	20.82	20.80
	TI EZW	20.52	20.58	20.60	20.59
0.0625 bpp	VM TCQ	19.05	19.09	19.10	19.09
	VM SQ	19.10	19.14	19.14	19.13
	TI EZW	18.77	18.81	18.82	18.81
Woman, size 2048 × 2560					
			PSNR with postprocessing		
Bit rate	Codec	PSNR	1	2	3
0.25 bpp	VM TCQ	29.72	29.54	29.27	29.05
	VM SQ	29.85	29.69	29.42	29.19
	MTWC	29.64	29.55	29.37	29.20
0.125bpp	VM TCQ	27.24	27.11	26.95	26.81
	VM SQ	27.28	27.16	27.01	26.87
	MTWC	27.02	26.94	26.83	26.72
0.0625bpp	VM TCQ	25.48	25.40	25.32	25.23
	VM SQ	25.44	25.37	25.28	25.20
	MTWC	25.07	25.04	24.99	24.93

* number of iterations

Table 4.2: The PSNR performance comparison for the Cafe and Woman image with several wavelet codecs.

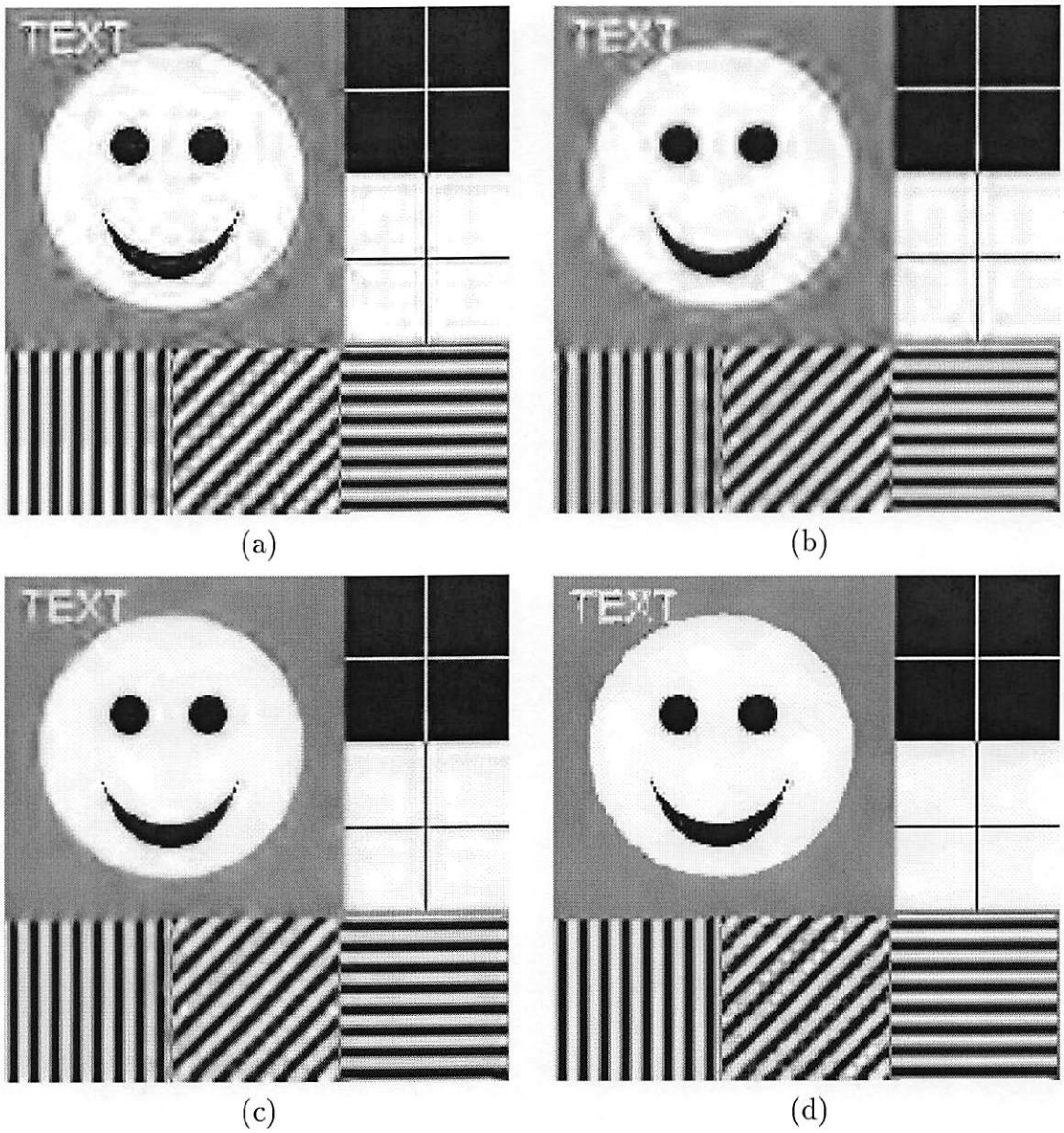


Figure 4.7: Examples of artifact reduction on a synthetic image: (a) wavelet compressed at 0.25 bpp, (b) the result by the method in [41], (c) the result by the multiscale approach in [28], and (d) the result of the proposed method with a truncated quadratic potential function (nonconvex).

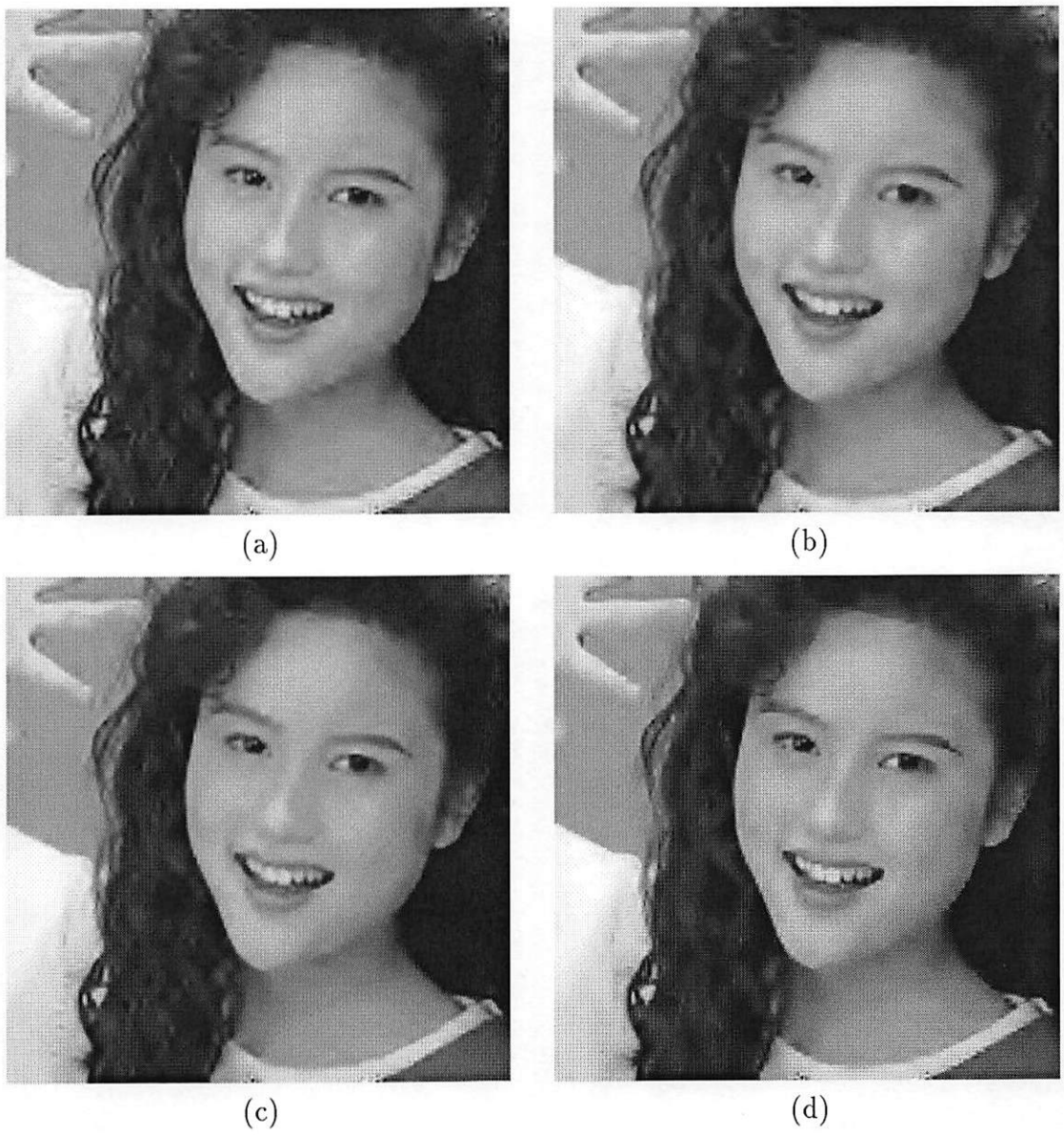


Figure 4.8: Examples of artifact reduction on a natural image: (a) wavelet compressed at 0.125 bpp, (b) the result by the method in [41], (c) the result by the multiscale approach in [28], and (d) the result of the proposed method with a truncated quadratic potential function (nonconvex).

Chapter 5

Real-time Postprocessing for DCT-Coded Images and Video

A computationally efficient postprocessing technique to reduce compression artifacts in block-DCT coded images and video is studied in this chapter. In particular, we lay the emphasis on very low bit rate applications, say, less than 0.3bpp for images and 64kbps for video. Although deblocking for JPEG coded images has been extensively studied, the distinguishing trait of the proposed technique from previous research is the computational complexity. We propose a table-lookup technique instead of traditional filtering or optimization approaches, which greatly reduces the computational overhead yet providing competitive results.

5.1 Introduction

In a block DCT-based codec such as JPEG or H.263, compression artifacts are observed due to the process of quantization. Typically, there are two kinds of visual distortions, i.e. spatial and temporal distortions. Spatial distortions are generally caused by a coarse quantization of DCT coefficients. Examples of such impairments are blocking effects, ringing effects, blurring, and changing in the texture patterns. Temporal distortions that appear only in video coding include flickering and motion jerkiness. Among these artifacts, blocking and ringing effects are the most visible. Reduction of these artifacts can result in a significant improvement in the overall visual quality of decoded video. This work proposes a new postprocessing method that is effective in reducing severe blocking and ringing artifacts, while maintaining

a low complexity and a low memory bandwidth, which are important requirements for the implementation of a real-time video coding system.

According to research in the past, there are three key components in an efficient postprocessing algorithm. They are: classification, smoothing, and constraint. Classification is often performed since artifacts are more perceptible in the homogeneous region. Smoothing is important since most noticeable artifacts are high frequency. Finally, additional constraints are needed to avoid blurring. We follow the same principle to develop our algorithm.

5.2 Algorithm

The proposed scheme consists of three steps: classification, smoothing, and constraint. Each of them is detailed below.

5.2.1 Classification

Classification can be performed in either the transform domain or the space domain depending on the applications. The transform domain classification requires less computation. However, for some applications postprocessing is implemented as one separate unit so that DCT coefficients are not immediately available. Thus, a space domain classification technique is preferred. Both approaches are presented below. An image is classified into blocks of three categories, namely *flat*, *edge*, and *texture*, with the following procedure.

- **Transform Domain Technique.**
  - *Flat blocks*: The flat block is classified by checking the summation of absolute value of DCT ac coefficients,

$$\sum_{i=3}^{63} |ac_i| < Threshold, \tag{5.1}$$

    where the coefficients are arranged in the zigzag scan order used in JPEG and H.263. Note that the first two ac coefficients only contribute to

smooth variation within a block. Therefore they are not included in the summation.

- *Edge blocks*: We consider blocks which do not pass the test in Eq.(5.1). If at least one of their four nearest neighbors is a flat block, they are classified as edge blocks.
- *Texture blocks*: All blocks that are neither flat blocks nor edge blocks are classified as texture blocks.

- **Spatial Domain Technique.**

  A simple spatial domain classification is applied to pixels v_i with $0 \leq i \leq 7$ as shown in Fig. 5.1 to determine if it is a boundary is located in a smooth region. The classification is done with the following procedure

  $$\text{count} = \phi(v_0 - v_1) + \phi(v_1 - v_2) + \phi(v_2 - v_3) + \phi(v_4 - v_5) + \phi(v_5 - v_6) + \phi(v_6 - v_7).$$

  where $\phi(x) = 1$ if $|x| < 3$ and set to zero, otherwise. If count is equal to 6, the boundary is in flat region, otherwise, it belongs to a high activity area, and it can be further classified into edge and texture with same rule as in the transform domain technique.

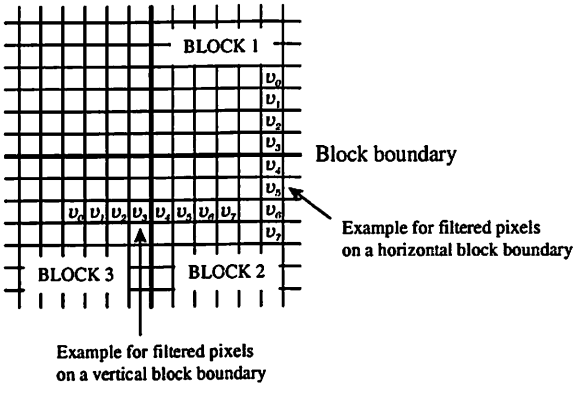


Figure 5.1: Examples of positions of filtered pixels.

5.2.2 Smoothing

The proposed scheme consists of a deblocking filter and a deringing filter for reducing two major artifacts in low bit rate DCT-based coding. Since the proposed approach

targets at real-time processing, we should avoid unnecessary computation that does not results in visual quality improvement.

- **De-blocking Filter**

  The proposed deblocking filter performs simple rules for the adaptation of local statistics and human visual system (HVS). To reduce the computational cost, the deblocking filter is performed on block boundaries only in the planar regions. This is based on the simple perceptual masking assumption that the blocking artifact is high frequency noise which is less visible in highly detail areas but very visible in the smooth region. At low bit rate, each DCT block in the flat region usually contains DC coefficients only, which results in a maximum smooth inside the blocks but a large discontinuity at block boundaries. Therefore, our deblocking filter is particularly designed to smooth this type of artifacts.

  The deblocking operation is done by updating the value of v_i, $1 \leq i \leq 6$ with

$$v_i' = v_i + \text{sign}(v3 - v4) \cdot d_i, \quad \text{if } 1 \leq i \leq 3,$$
$$v_i' = v_i - \text{sign}(v3 - v4) \cdot d_i, \quad \text{if } 4 \leq i \leq 6,$$

  where d_i is obtained by looking up the predefined table according to the offset between v_3 and v_4. An example of such a table is shown in Table 5.1. The table can be defined in more detail if $|v_3 - v_4| > 23$. However, it is observed that $|v_3 - v_4|$ is rarely over 20. Thus, for an error larger than values specified in table, we apply a linear interpolation to adjust the value of v_i.

  The design of the above table is based on the fact that the quadratic function is most effective for smoothing. The quadratic function will minimize the difference between pixels. When the offset is evenly distributed between pixels, the result will be the best. Therefore, the fast way to achieve smoothness is the use of linear interpolation in the very smooth area. Interpolation does not introduce much distortion while greatly improve visual quality. However, for real-time applications, the filtering or interpolation operation is still expensive to perform due to the multiplication or the division operation. The approach proposed above is non-iterative. The use of the look-up table only

$d = \lvert v_3 - v_4 \rvert$ $v3 > v4$	d_1	d_2	d_3	d_4	d_5	d_6
under 3	0	0	0	0	0	0
3	0	0	-1	1	0	0
4,5	0	-1	-2	2	1	0
6,7,8	-1	-2	-3	3	2	1
9,10	-1	-2	-4	4	2	1
11,12	-1	-3	-5	5	3	1
13,14,15	-2	-4	-6	6	4	2
16,17,18	-2	-4	-7	7	4	2
19,20	-2	-5	-8	8	5	2
21,22	-2	-5	-9	9	5	2
23	-3	-6	-10	10	6	3
above 23	-d/8	-d/4	-d/2	d/2	d/4	d/8

Table 5.1: A table for the deblocking filters.

requires integer-add and integer-subtract operations that speed up the computational speed tremendously. In fact, the linear filtering, diffusion filtering, or the restoration-based method will all lead to similar results in the flat areas. However, they require several iterations to smooth out a large offset.

- **De-ringing Filter**

  The robust nonlinear filtering technique in Section 4.2 is applied to the edges blocks for de-ringing. The proposed deringing filter based on the robust M-estimator can smooth ringing noise while preserving important image detail. Furthermore, no explicit edge extraction is required. This is the major difference of our method in comparison with other previous work. Theoretically, the explicit edge extraction is equivalent to the use of the truncated quadratic potential function, which totally cuts off the influence from the outliers. Even though the estimate of the truncated quadratic function is not stable around the point of the threshold, it is however still able to achieve the maximum smoothness in the flat region. This type of potential function is not proper for natural images, but good for synthetic images, such as the cartoon image with uniform color separated by edges. It is also ideal for some applications such as image segmentation.

5.2.3 Constraints

The values specified in the table already make the distortion bounded, so that there is no need to specify constraints for smooth blocks. For constraints used in the de-ringing filter, it is set as the value of the smallest quantization step size.

5.3 Experimental Results

5.3.1 Postprocessing of JPEG Compressed Images

In order to show the robustness of the proposed technique, we applied the proposed technique to a number of *de facto* standard test images compressed by JPEG. For comparative study, three postprocessing algorithms: POCS [60], multiscale MAP(MMAP) [28] and the proposed method (3×3 window, Huber potential function for deringing filter) are performed and the PSNR result as well as the speed are summarized in Table 5.3. We see from the results that MMAP has the best PSNR performance, but the processing time is much longer than the other two methods due to the high complexity. The proposed method provides a slightly worse objective improvement, but with much less computing time. The POCS method only restores the boundary pixels so that the subjective quality improvement is the worst among the three. Its complexity is also high since it requires several iterations to converge, where each iteration has to perform one DCT/IDCT pair. Furthermore, both MMAP and POCS require to buffer the whole frame in the iterative process. Our new scheme only utilizes information in neighboring blocks, so that it saves memory bandwidth. It is clear that our new scheme is more suitable for on-line real-time process than the other two methods.

It is worthwhile to point out that PSNR is not the best measure to evaluate the postprocessing algorithm, since it does not completely correlate with human perception. We show the subjective quality of different methods in Figure 5.2– 5.5. It was found that both MMAP and the proposed method achieve good subjective quality.

5.3.2 Postprocessing of H.263 Compressed Videos

The proposed nonlinear filter can be used as a pure postfilter or incorporated with motion estimation as a loop filter. In this work, we apply it as a pure post-filter and compare the result with the post-filter defined in H.263. Throughout the test, the UBC TMN8 codec is used. The experiment is performed on Class A, Class B, and Class C sequences. Each sequence with 300 frames was used for simulation. The sequences are compressed at a given bit-rate, image-size, and frame rate with the IPPPP... structure. That is, only the first frame is intra-coded while others are all inter-coded. The average PSNR result of applying the deblocking filter alone, and with both filters applied are shown in Table 5.4. The result also compared with TMN8 decoder and decoder with post-filtering.

Images shown in Figs. 5.6 and 5.7 illustrate the subjective improvement of the proposed approach. From the subjective and objective data presented, it is clear that the proposed deblocking approach is more efficient than the H.263 postfilter. With only the use of the deblocking filter, the postfilter is able to achieve real-time playback at the targeted frame rate. The real-time playback speed for a software decoder running on a PC with Pentium 200 and 64M memory is shown in Table 5.5. The ringing noise is in high activity area and may not be perceivable at some place. Since deringing is more expensive, we provide 3 options from the weak to the strong to determine the amount of de-ringing. By weak, moderate and strong, we mean 2, 3 and 4 pixels around boundaries are filtered, respectively. From the table, we see that, with the deblocking only, it is able to playback at the desired rate. Since deringing is a little more expensive compared with the deblocking filter, the maximum frame rate for decoding with both deblocking and deringing with such PC is around 4-6 frames/sec. For memory, since the deblocking and deringing filters does not use any information outside a small local window, the processing can be done without buffering large images. This is important for many applications that have the memory constraint.

5.3.3 Loop-filtering of H.263 Compressed Videos

The proposed de-blocking filter can also be used as a loop filter. Since the loop filter is placed in the motion compensation loop, the encoder and the decoder should apply

the same de-blocking filter. Results of TMN8, TMN8 with its default loop-filter and TMN8 with the proposed loop-filter are shown in Table 5.6. Results of the proposed method outperforms TMN8 loop-filter up to 0.3dB. This is due to the fact that TMN8 only considers 4 pixels around the block boundary, while a larger extent was modified in our approach to achieve better smoothness. The proposed loop-filter in conjunction with the advanced prediction mode can boost up the performance as high as 2.3 dB. The improvement is substantial when the original video contains more homogenous areas (e.g. the Container sequence). The complexity and the memory required for the proposed loop-filter and the TMN8 loop filter are listed in Table 5.2. The actual computations required for the proposed loop filter is varying for different sequences, since it is not applied to every boundary pixels. The comparison of the loop-filter and the post-filter can be summarized as below.

- The major advantage of applying the loop-filter rather than the post-filter is that it requires less computation and memory at the decoder side. This may be critical for some applications which have the memory constraint.
- The loop filter tends to result in a slightly better coding gain since the reference frame has been smoothed before motion estimation and the motion estimation accuracy is improved.
- The major drawback of the loop filter is that it has to be standardized. The computational complexity at the encoder side slightly increases.

Algorithm Reference	Memory Usage	Complexity
TMN8 loop filter	No overhead	$21N$
Proposed de-blocking	No overhead	$11N + 9M$

Table 5.2: Comparison of complexity and memory requirements between the TMN8 loop-filter and the proposed de-blocking filter as a loop filter, where N is the total number of boundary pixels and M is the number of boundary pixels in smooth regions.

5.4 Conclusion

A statistical model that characterizes both compression degradation and image smoothness is examined to remove the compression artifact. The model provides a unified framework for artifact reduction and image enhancement. To overcome the complexity of constrained optimization, efficient nonlinear filters are proposed to obtain an approximate minimum quickly. For debloking, a table look-up scheme is applied and the blocking artifact can be reduced in real time. For deringing, the nonlinear filter can reduce ringing noise without blurring the image details. The new postprocessing algorithm can greatly enhance the visual quality of decoded video with a low computational cost. The real-time postprocessing is made possible due to the low complexity and the low memory bandwidth requirements of the proposed algorithm.

Image (size)	bpp	PSNR (dB)	PSNR Improvement			Time** (sec)		
			POCS*	MMAP	Proposed	POCS*	MMAP	Proposed
Goldhill (720 × 576)	0.20	29.49	0.50	0.67	0.37	6.4	87.3	3.5
Hotel (720 × 576)	0.27	29.09	0.46	0.80	0.38	7.8	83.3	3.4
Woman (512 × 640)	0.21	29.35	0.34	0.67	0.47	6.0	68.8	2.4
Cafe (512 × 640)	0.54	22.96	0.15	0.49	0.14	9.7	67.2	2.3
Lena (512 × 512)	0.25	30.70	0.68	1.01	0.62	4.6	54.1	1.7
Cameraman (256 × 256)	0.38	25.17	0.18	0.37	0.15	2.4	10.7	0.5

* de-blocking only

* All tests are based on Pentium 200 with 64M memory.

Table 5.3: Comparison of PSNR and speed among different postprocessing algorithms

Sequence	Fixed Q*	Average PSNR(Y)			
		TMN8	TMN8 deblocking	proposed	
				deblocking	deblocking & deringing
10kbps,QCIF,7.5Hz					
Hall monitor	17	31.649	31.681	31.733	31.743
Mother & dau.	15	32.782	32.861	32.853	32.867
Container	17	30.904	30.925	30.957	30.797
24kbps,QCIF,10Hz					
Hall monitor	9	31.686	31.716	31.768	31.767
Mother & dau.	8	32.782	32.850	32.523	32.862
Container	10	30.917	30.938	30.971	30.863
Silent	14	30.76	30.78	30.841	30.848
48kbps,QCIF,10Hz					
Foreman	13	31.092	31.112	31.126	31.197
Coast guard	14	29.282	29.288	29.392	29.067

* Q for the first frame is 13

Table 5.4: The PSNR results of Class-A, Class-B, and Class-C sequences coded at different bit rates

Sequence	Fixed Q	TMN8 (fps)	proposed			
			deblock only (fps)	deblock & w. dering (fps)	deblock & m. dering (fps)	deblocking & s. dering (fps)
10kbps,QCIF,7.5Hz						
Hall monitor	17	7.71	7.69	6	5.21	4.52
Mother & dau.	15	7.71	7.69	6.3	5.33	4.62
Container	17	7.71	7.69	6.4	5.44	4.73
24kbps,QCIF,10Hz						
Hall monitor	9	10.18	10.16	6.29	5.31	4.6
Mother & dau.	8	10.18	10.16	6.32	5.34	4.63
Container	10	10.18	10.16	6.41	5.41	4.7
Silent	14	10.18	10.16	4.58	3.81	3.29
48kbps,QCIF,10Hz						
Foreman	13	10.18	10.15	5	4.28	3.69
Coast guard	14	10.18	10.15	4.31	3.59	3.08

Table 5.5: Real-time playback speed (including decoding and display) in terms of frames per sec (fps)

Sequence	Fixed QP*	Average PSNR(Y)		
		TMN8	TMN8 loop filter	proposed loop filter
10kbps,QCIF,7.5Hz				
Hall monitor	18	29.172	29.897	29.898
Mother & dau.	15	31.749	32.272	32.331
Container	17	27.129	29.164	29.426
24kbps,QCIF,10Hz				
Hall monitor	9	32.658	34.000	34.044
Mother & dau.	8	34.211	35.151	35.226
Container	10	30.682	32.141	32.475

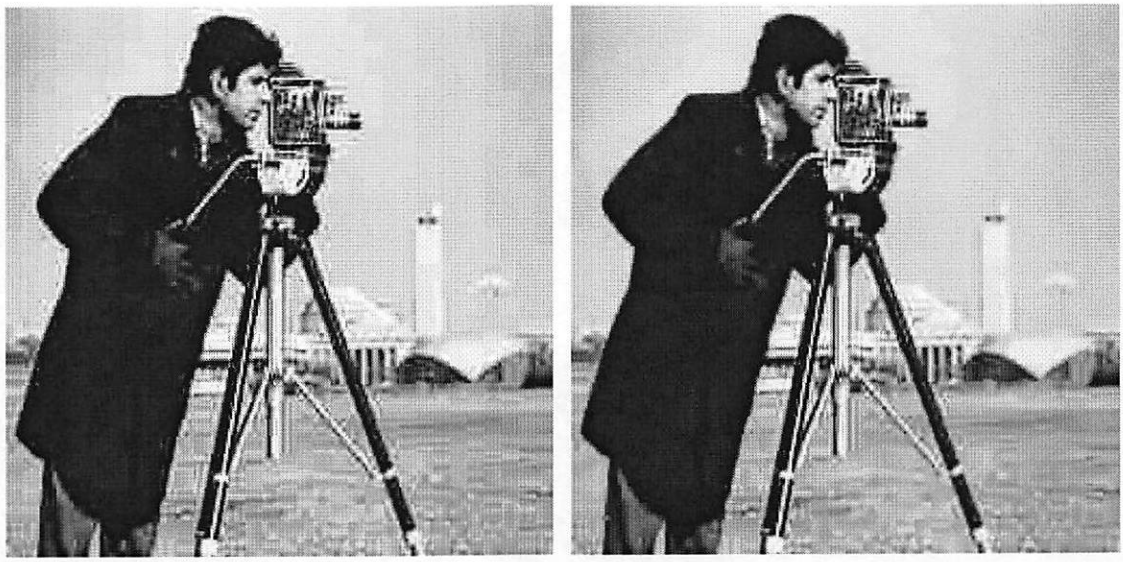
* with advanced prediction mode(Annex F)

Table 5.6: The PSNR performance comparison of the TMN8 loop-filter and the proposed de-blocking filter used as a loop filter.

Figure 5.2: Lena encoded by JPEG at 0.25bpp.

Figure 5.3: JPEG-coded Lena at 0.25bpp postprocessed by the proposed technique.

Figure 5.4: JPEG-coded Lena at 0.25bpp postprocessed by the MMAP technique.



(a) (b)

Figure 5.5: Example of artifact reduction: (a) JPEG-coded Cameraman at 0.37 bpp, and (b) postprocessed by the proposed technique.

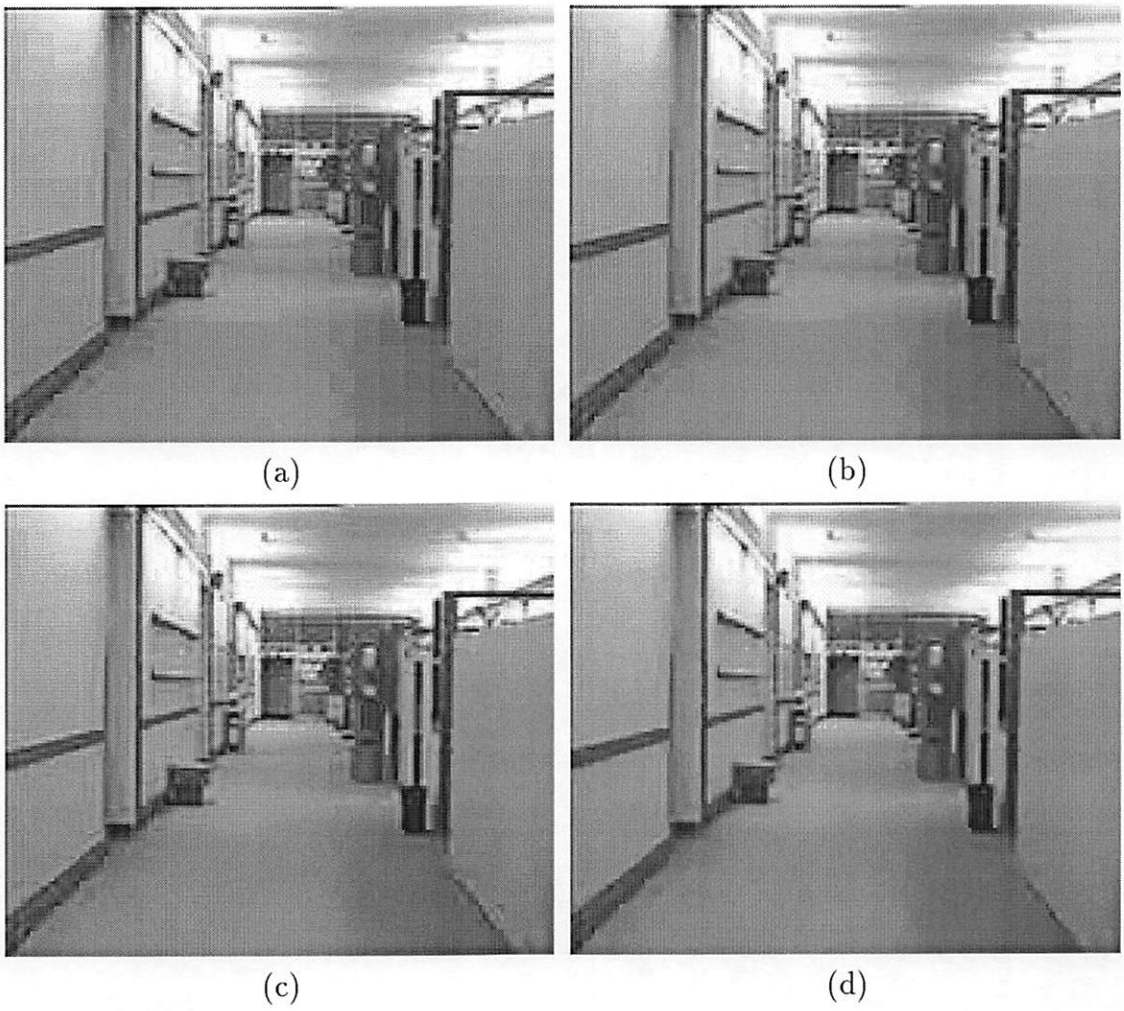


Figure 5.6: Example of artifact reduction on No. 1 frame of "Hall Monitor": (a) decoded at 10 kps without postprocessing, (b) the result with TMN8 deblocking, (c) the result with the proposed deblocking, and (d) the result with the proposed deblocking and deringing.

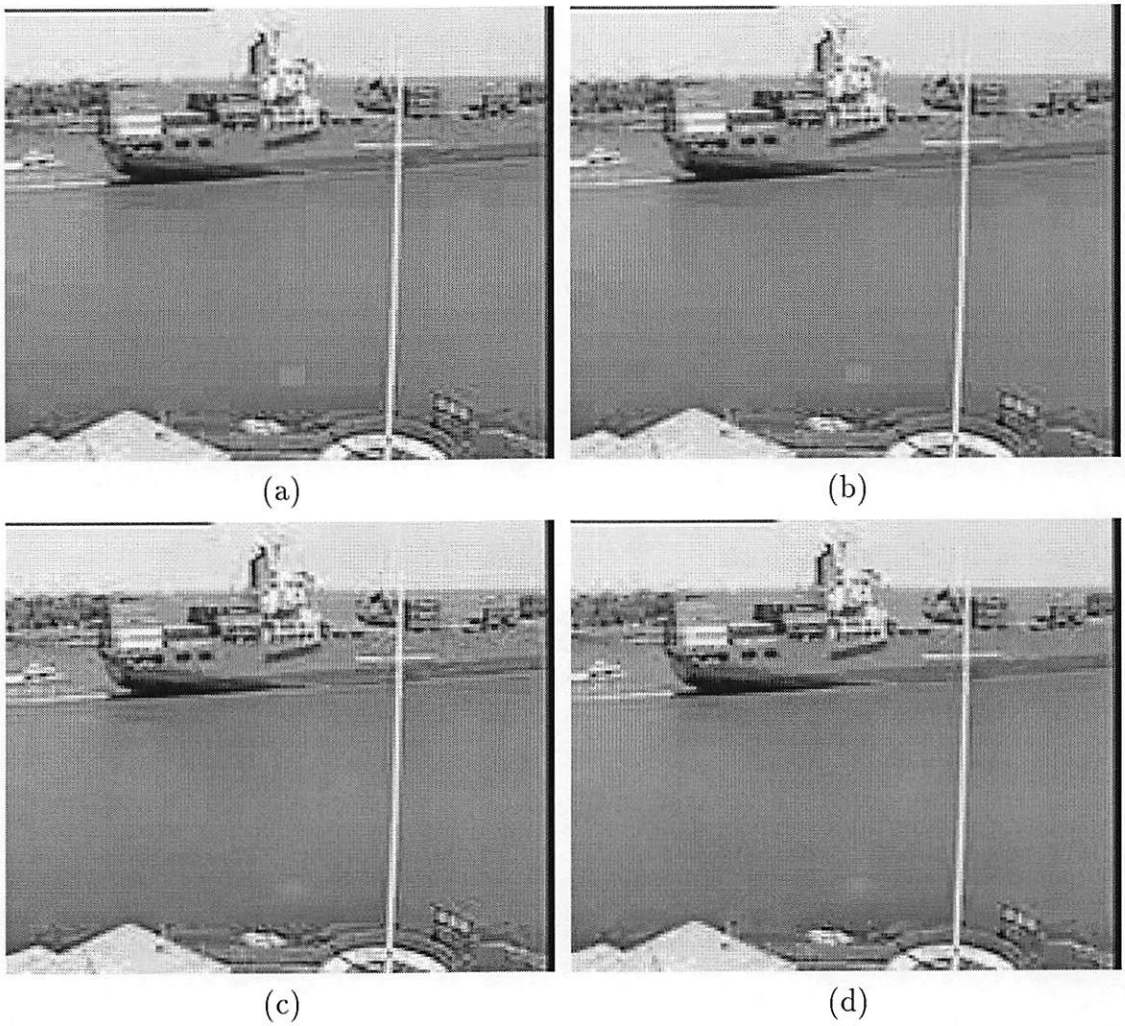


Figure 5.7: Example of artifact reduction on No. 1 frame of "Container ship": (a) decoded at 10 kps without postprocessing, (b) the result with TMN8 deblocking, (c) the result with the proposed deblocking, and (d) the result with the proposed deblocking and deringing.

Chapter 6

Inverse Halftoning via Nonlinear Filtering

Halftone images are binary images that provide a rendition of gray-level images. Inverse halftoning addresses the problem of recovering a continuous-tone image from a halftone image. This chapter introduces a new inverse halftoning algorithm based on the nonlinear filtering technique proposed in Chapter 3. The distinct features of our approach are: (a) efficiently smoothing halftone patterns in large homogeneous areas, (b) additional edge enhancement capability to recover the sharp edge quality, (c) an excellent PSNR performance with only local integer operations and requiring a small memory buffer.

6.1 Introduction

Inverse halftoning has applications in halftone manipulation, conversion, and compression when a halftone is the only available version of an image. In this chapter, our nonlinear filtering technique is used to solve the inverse halftoning problem. Although the proposed framework can work with any halftone process, we only focus on the most popular technique, i.e. error diffusion, in this research.

In error diffusion halftones, the error between the continuous-tone image and the binary halftone at each pixel is diffused over a causal neighborhood. Most of the quantization noise power falls at high spatial frequency. A good inverse halftoning scheme should remove as much noise as possible, while preserving the important image details. Several inverse halftoning methods have been proposed in the literature, including linear adaptive filtering, nonlinear filtering, wavelets,

MAP estimation and POCS. These techniques can be well explained with the block diagram given in Figure 6.1.

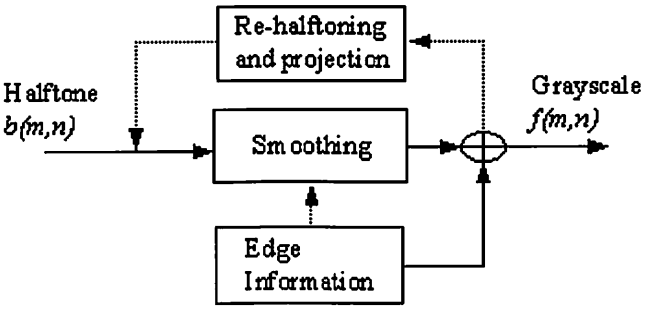


Figure 6.1: The block diagram of a generic inverse halftoning algorithm.

A generic inverse halftoning process includes two key modules: the halftone noise smoothing and image details recovering. Smoothing can be performed by using simple linear filtering techniques, e.g. the halfband lowpass filtering [54], the Gaussian lowpass filtering and singular value decomposition (SVD) [7, 13], or model-based nonlinear smoothing, e.g. the nonlinear filtering proposed in this work and MAP estimation [50]. Some work adopts edge information to improve smoothing results such as spatial varying FIR filtering [23] and wavelet de-nosing [57]). All the above methods can reduce halftone noise efficiently. However, the selection of bandwidth parameters (or, equivalently, thresholds in wavelet and model-based techniques) plays a critical role in the overall performance of a given algorithm. Besides, halftone noise contributes its energy mostly at high frequency components, which can potentially corrupt or reduce the quality of image edges. This is the fundamental limitation of inverse halftoning algorithms that rely on the smoothing strategy only. Thus, an edge recovery procedure is required to obtain a high quality inverse halftoned image. To recover edges, a bandpass filter with thresholding was used to extract and enhance edges in [7] and a Gaussian filter applied to a highpass band was studied in [57] . In the restoration-based approach [50, 2], there is no explicit edge extraction step while the edge information is recovered by projection to the original halftone image with the assumption that halftone kernel is known.

In this work, we present a single-pass blind inverse halftoning algorithm with a low computational complexity and a low memory buffer. It is called blind since we

do not assume that the halftone kernel is available. Our scheme performs nonlinear filtering in conjunction with edge enhancement to improve the quality of an inverse halftoned image. The proposed approach is universal in the sense that no prior knowledge about the forward halftoning process is assumed. The remarkable feature of this simple algorithm is that it outperforms the best results obtained by other blind inverse halftoning methods.

6.2 Two Problems

Previous work on blind inverse halftoning was performed without knowing the diffusion kernel. It is nevertheless assumed that a "good" halftoning scheme has been applied so that the resulting halftone image will not have localized energy in smooth areas, and edges are well represented. This is however not true in general. In this section, we address two major problems that have not yet been pointed out in previous research in blind inverse halftoning.

6.2.1 Halftone Artifacts

One of the most observable artifacts associated with error diffusion is the periodically repeating pattern appearing in areas where the image is a constant or slow varying. This is called the limit cycle behavior (or idle tone), which has been studied in [10]. The limited cycle phenomenon is resulted from a constant input, especially, when the input value is a constant that slightly deviated from $\frac{1}{2}$, $\frac{1}{3}$, and $\frac{1}{4}$. The stability of these periodic patterns is determined by the choice of error diffusion weights and the support of the error diffusion kernel.

An example of limited cycle patterns is shown in Fig. 6.2. In this example, the input is set to a constant $\frac{130}{256}$. It will result in a low frequency noise and make a small perturbation to the checkerboard pattern (Note that most error diffusion kernels generate a checkerboard pattern at mid-gray). In Fig. 6.2 (d), the Floyd-Steinberg kernel is used. The dominating pattern is the checkerboard. The worm pattern appearing in Fig. 6.2 (e) is generated by equally weighted 4 coefficients in the diffusion kernel. Fig. 6.2 (f) is the result of the Jarvis kernel [21]. The Jarvis kernel has a larger support than the other two kernels. Increasing the support

generally can reduce limited cycle patterns. The limited cycle behavior introduces low frequency noise or structure noise (e.g. worm patterns), which is difficult to remove by using the linear lowpass filter or even the wavelet denoising technique. A good blind inverse halftoning algorithm should be able to deal with the limited cycle behavior.

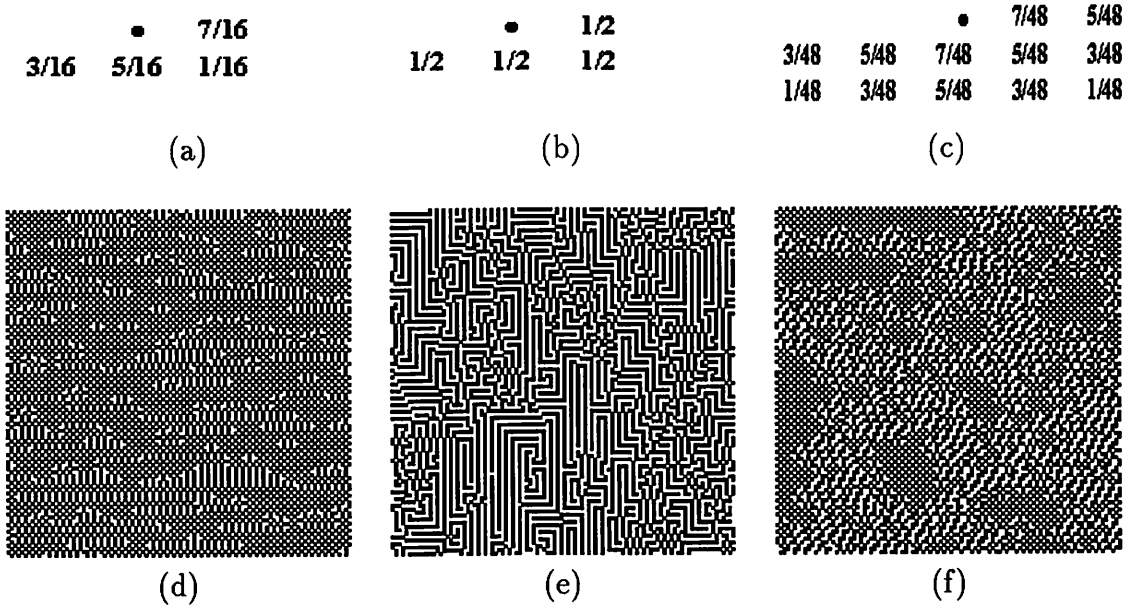
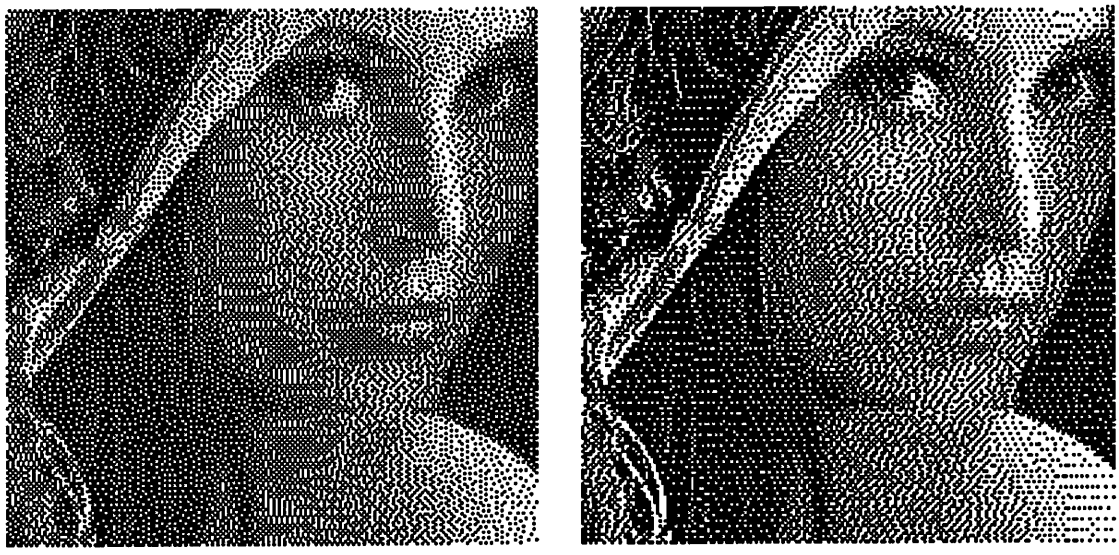


Figure 6.2: Examples of the limited cycle behavior with a constant input set to 130. (a)-(c)are error diffusion kernels used in (d)-(f), respectively, (d) is dominated by checkerboard patterns, (e) is dominated by worm patterns, and (f) exists no obvious dominate patterns.

6.2.2 Edge Sharpness

The sharpness of an edge is determined by halftoning schemes, and halftone kernels. In general, an error-diffused image has more edge information than the one using clustered-dot or disperse-dot methods. The error diffusion kernels also affect the quality of edges. Larger diffusion kernels tend to have sharper edges. An example to illustrate the influence of the error diffusion kernel is shown in Fig. 6.3. Clearly, the 12-coefficient Jarvis kernel [21] produces sharper edges than the 4-coefficient Floyd-Steinberg kernel at the cost of more noise in the slow varying regions. If significant edge information is lost during the forward halftoning process, the reconstructed

inverse halftone will have rather smooth edges. In blind inverse halftoning, we cannot assume the edge is always perfectly rendered. The loss of edge resolution cannot be solved by edge preserving smoothing algorithms. Additional techniques, such as edge enhancement, must be perform to recover sharp edges.



(a) (b)

Figure 6.3: Examples of edge quality influenced by the diffusion kernel: (a) the Floyd-Steinberg kernel and (b) the Jarvis kernel.

6.3 Proposed Inverse Halftoning Algorithm

The proposed inverse halftoning algorithm consists of two components, noise smoothing and edge enhancement, to overcome the problems addressed in previous section. A block diagram of the proposed algorithm is shown in Figure 6.4. The lowpass filter is served as a pre-processor to provide a rough estimate of the continuous-tone image and refined by a nonlinear filter that can efficiently suppress low frequency noise resulted from limited cycle patterns in smooth regions. The loss of edge resolution is compensated by an edge enhancement technique using a bandpass filter and morphological filtering.

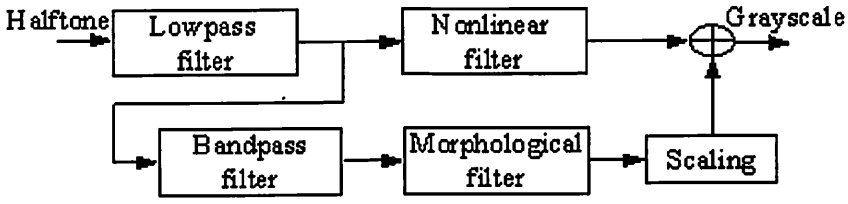


Figure 6.4: Block diagram of the proposed halftoning algorithm.

6.3.1 Design of Smoothing Filters

As discussed in the early part of the thesis, robust nonlinear filters can suppress noise in continuous-tone images efficiently. It is however a location filter that is difficult to apply to the inverse halftoning problem directly. To solve this problem, we pre-process the halftone image with lowpass filtering, and then feed the output into the robust nonlinear filter for refinement.

The lowpass filter plays a critical role for the final performance. To achieve better computational efficiency, the 2D lowpass filter should have the following desirable properties:

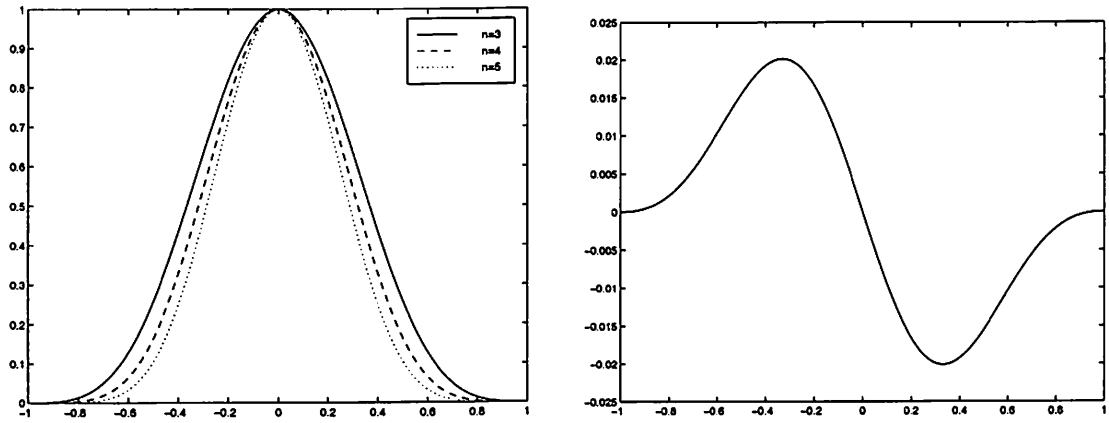
- an FIR filter with a small region of support;
- a separable filter for the ease of implementation;
- a controllable cutoff frequency that can be determined by a single parameter with a fixed size of support.

A class of 1D filters satisfying the above criteria is used in our algorithm. They are designed based on a spline function θ,

$$\theta(x,n) = (\frac{\sin \pi x}{\pi x})^n, \quad n = 1,2,\ldots,N. \tag{6.1}$$

For a given filter length, coefficients of the filter can be determined by uniformly sampling the above function between [-1,1], and the cutoff frequency can be adjusted by the value of n. A larger value of n leads to a higher cutoff frequency and, consequently, a less smoothing capability. Coefficients obtained by the above function should be normalized to achieve unity gain at DC. The shape of this function with different n is depicted in Figure 6.5(a). The above lowpass filter is superior to the

Gaussian filter with respect to error diffusion halftones. The output of this filter is a rough estimate of the original continuous-tone image and will be sent to a nonlinear filter and a bandpass filter for further process.



(a) (b)

Figure 6.5: (a) A family of spline functions used in the lowpass filter design and (b) the derivative of a cubic spline.

The nonlinear filter is the same as the de-ringing filter described in Section 4.2. To reduce halftone noise, Huber and truncated quadratic potential functions are better choices for the cost function, since they are efficient in smoothing and can be implemented using only integer operations. For images containing more homogeneous regions, the truncated quadratic function works the best.

There are three parameters in a nonlinear filter. One is the filter size. Since the diffusion kernel has smaller extent, a nonlinear filter of window size 3×3 is usually sufficient for most images. If the underlying image contains a large amount of smooth regions, a larger window size can be applied to reduce clustered noise. The second parameter is the threshold T used in potential functions. For the truncated quadratic potential function, T can be chosen between 5 to 20. A larger value of T tends to produce a smoother result. For the Huber function, T ranges from 1 to 10. For $T = 1$, the nonlinear filter is equivalent to the median filter. For better quality, T can vary adaptively based on edge information. The third parameter is the threshold for clipping, which is the value of $Th1$ in (4.2). By empirically testing the average distortion of inverse halftoned results for a pre-selected set of images,

the value of T can be chosen to be one half or one quarter of the average value. The purpose of the clipping operation is to avoid additional distortion.

The cascade of lowpass and nonlinear filters can efficiently smooth the low-frequency clustered halftone noise in large smooth regions. Examples to demonstrate this phenomenon are shown in Figs. 6.6 (a)-(d). The original image has two constant gray levels ($\frac{127}{256}$ and $\frac{200}{256}$). The resulting halftone is shown in (a). The result of the proposed algorithm without edge enhancement is shown in (b). In this case, a quadratic function with $T = 10$ is used and the window size is 9×9. Results obtained by other blind inverse halftoning algorithms [23, 7] are shown in (c)-(d) for comparison. Clearly, the proposed algorithm is more efficient in reducing noise in large smooth regions. Note that the proposed technique can preserve the quality of the sharp edge without further enhancement for this case. This is due to the fact that the original image has a crispy edge. Since there is no intermediate level, the resolution of edge can be easily preserved. In addition, this image contains two constant gray levels separated by the sharp transition. It fits the truncated quadratic model the best so that the sharp edge can be well recovered by the proposed nonlinear filter with only a minor distortion.

6.3.2 Edge Enhancement

In the halftoning process, the gray-scale resolution (i.e. intensity) is represented by the density of randomly distributed binary patterns. In some edge regions, such as those around an unsharp edge, the intensity is changing gradually within small extent. For this case, intermediate intensities at the edge transition area cannot be well represented by halftone patterns. The resulting edge after inverse halftoning is smoother than the original one due to the loss of grayscale resolution. This loss of resolution cannot be recovered even with the proposed nonlinear filter. Note that our nonlinear filters use the first-order model for discontinuity that does not fit smoother edges. The use of a higher order model for edge approximation may solve this problem at the price of a higher complexity.

The edge enhancement approach presented in this section is based on heuristic arguments. It however provides a good solution to the loss of grayscale resolution problem at a low computational cost. With this approach, useful edge information is

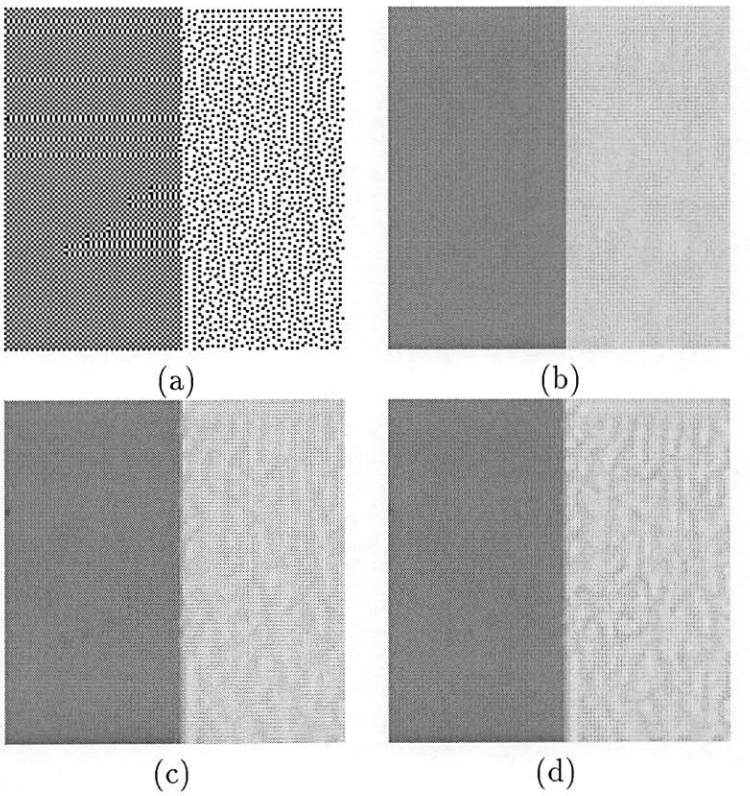


Figure 6.6: Examples of smoothing a halftone image: (a) the halftone of an image with two constant gray levels 130/256 and 200/256; (b) the result by applying the proposed smoothing filters; (c) the result by applying the adaptive filtering technique [23]; (d) the result by applying the nonlinear smoothing technique [7].

obtained by applying a bandpass filter to a lowpass-filtered halftoning image. There are two ways to realize the bandpass filter. One is applying two lowpass filters with different cutoff frequencies and calculating the difference between the two filtered results. The other method is taking derivative of (6.1) and sampling it with a desired filter length. The derivative of (6.1) is depicted in Fig. 6.5 (b).

In this work, we adopt the first approach, since a filter of shorter length can be used. After bandpass filtering, a morphological operation is applied to refine the bandpass filtered result. The morphological filter includes two steps. First, thresholding is applied to the output of the bandpass filter to generate a binary edge map E. Then, a binary median filter is applied to E for removing noise in this map. This edge information is used to refine the original bandpass filter output by setting values of non-edge pixels to zero. Finally, the refined result B is added back

to the nonlinear filter output S to obtain the inverse halftoning image I. In terms of mathematics, we have

$$I(m,n) = S(m,n) + \lambda B(m,n), \qquad 1 \leq m,n \leq N, \tag{6.2}$$

where λ is a scaling factor to control the amount of edge enhancement. This enhancement technique has a very low complexity, since only windows of a small size and integer operations are used.

6.4 Experimental Results

We performed our experiments on five images, *Lena, Peppers, Woman, Cafe* and *Bike*. These images are error diffused using Floyd-Steinberg kernel. We compared our results with all previous blind inverse halftoning algorithms as follows.

6.4.1 Comparison of Lowpass Filters

The lowpass filter that performs the initial estimate of a continuous-tone image plays a critical role in inverse halftoning results. Equation (6.1) provides a class of lowpass filters. The cutoff frequency of this lowpass filter should be high enough to preserve spatial details while providing a sufficient grayscale resolution for the robust nonlinear filter. To achieve a good trade-off between the performance and the complexity, a 7×7 separable FIR filter with coefficients listed in Table 6.1 is used in the implementation. These coefficients are obtained by setting $n = 3$ in (6.1) and normalized to a unity gain. We compare the performance of this filter with other filters used in other inverse halftoning algorithms, including the 9×9 Gaussian filter [13], the halfband lowpass filter [54], and the wavelet [57]. The PSNR values for several test images after the lowpass filter are shown in Table 6.2. The results indicate that the proposed filter is superior to others.

n	-3	-2	-1	0	1	2	3
H	0.0089	0.0852	0.2409	0.3300	0.2409	0.0852	0.0089

Table 6.1: Coefficients of 7-tape lowpass filter used in our experiment.

Algorithm		Gaussian [13]	Halftband [54]	Wavelet [57]	Proposed
PSNR (dB)	Lena	28.64	29.60	30.38	30.79
	Peppers	27.59	27.76	28.56	28.65

Table 6.2: Performance comparison of different lowpass filters.

6.4.2 Blind Inverse Halftoning

Our technique does not assume the availability of the error diffusion kernel. The PSNR results of our proposed technique and previous blind inverse halftoning algorithms, including statistical smoothing [54], wavelet [57], adaptive filtering [23], and nonlinear noise removal [7], are listed in Table 6.3. Non-blind inverse halftoning algorithms assumes that the halftone kernel is known and performs projection to improve the quality. Results of these techniques are also provided in Table 6.4 for comparison. We can see clearly from Table 6.3 that our technique achieves the best result among all blind inverse halftoning algorithms. Our method is also robust across all test images. It is interesting that our blind inverse halftoning results are comparable to or even outperform those obtained from iterative projection. It is actually not surprising that the kernel information does not boost up the inverse halftoning performance since the original halftone image is not a strict constraint in the space domain. This is especially obvious when intensities are around mid-gray, projection to this set may not move the estimate toward the real solution, but a legal solution.

If the purpose of inverse halftoning is to obtain the original continuos-tone image, then subjective quality of inverse halftone is also important. In Fig. 6.7, the center portion of *Lena* and *Peppers* images is enlarged and compared to that obtained from a wavelet-based blind inverse halftoning approach, which is known to produce the best visual quality among previous work. As mentioned before, the size of nonlinear filters effects the smoothing capability in large homogeneous areas. Examples obtained from 3×3, and 7×7 nonlinear filters are shown in Fig. 6.8. Clearly, nonlinear filters with a larger extent strongly smooth the homogeneous region while important edges are still well preserved. In general, a larger filter size produces more pleasant visual quality, although the PSNR value may sometimes be worse than those obtained from a smaller window size.

Algorithm, Reference	PSNR				
	Lena	Peppers	Woman	Bike	Cafe
Statistical [54]	31.00	29.30	-	-	-
Wavelet [57]	31.50	30.43	-	-	-
Adaptive [23]	31.33	29.03	29.56	22.64	22.03
Nonlinear [7]	31.26	28.96	29.21	22.26	21.39
Proposed	31.62	30.82	29.60	22.97	22.41

Table 6.3: Performance comparison of different blind inverse halftoning algorithms.

Algorithm Reference	PSNR	
	Lena	Peppers
Statistical [54]	32.00	30.30
Wavelet [57]	31.67	30.69
POCS [13]	30.4	-

Table 6.4: Performance comparison of inverse halftoning algorithms that assume the halftone kernel is known.

6.4.3 Complexity and Memory

The complexity of the proposed technique is slightly higher than the linear filtering approach [23, 7], but much less than the wavelet based method [57], since reconstruction from wavelet maxima requires iterative projection.

To give an example, if we use a 7×7 window for all three lowpass filters, 5×5 for the binary median filter, and 3×3 for the nonlinear filter, the number of operations required per pixel is less than 500. The execution of the proposed algorithm proceeds in the raster scan fashion. Several rows are required for the filtering process. In this example, at least 11 rows have to be buffered.

The complexity and memory requirements for several inverse halftoning schemes are compared in Table 6.5. The memory usage is the minimum number of data required for buffering to process one row of an image. The computational complexity is estimated from algorithms given in the corresponding citation. "Low" means a number fewer than 500 operations per pixels, "Median" denotes 500-2000 operations per pixel, and "High" means more than 2000 operation required. Clearly, scan-based techniques such as our algorithm have a memory requirement that is proportional to image width N, while the frame-based methods have to store the entire frame and the complexity is proportional to N^2 for an image of size $N \times N$. Furthermore,

most frame-based techniques are iterative so that we have to buffer several copies of the image. The wavelet approach requires the largest amount of memory buffer due to the use of over-complete wavelet decomposition, and all bands created in each decomposition have to be buffered.

Algorithm Reference	Memory Usage	Complexity
POCS [13]	$8N^2$	High
MAP [13]	$8N^2$	High
Wavelet [57]	$9N^2$	Med.
Adaptive[23]	$7N$	Low
Nonlinear[7]	$11N$	Low
Proposed	$11N$	Low

Table 6.5: Performance comparison of inverse halftoning schemes in terms of complexity and memory.

6.4.4 Other Diffusion Kernels

We also test our algorithm on halftone images generated by using the Jarvis kernel [21], and multiscale halftoning algorithm [22]. Results from the proposed method and other two blind inverse halftoning algorithms are compared in Table 6.6. Throughout these tests, a 3×3 nonlinear filter with the *Huber* potential function is used in our algorithm. Threshold T in the potential function is set to 2 and constraint $Th1$ is set to 15. From Table 6.6, it is clear that the proposed algorithm outperforms other two methods for all images. The inverse halftoned images are provided in Figs. 6.9 - 6.10 for subjective evaluation.

As discussed before, the Jarvis kernel produces sharper edges at the price of more noise in homogeneous regions. We would like to point out that the Jarvis kernel produces higher contrast edges. Lowpass filtering of these edges generates an enhanced version of the original ones. In our approach, the *Huber* function is used as a cost function. It allows a smoother transition of large discontinuities, results in an edge that is more close to the original one and produces better results in homogeneous areas since it is a quadratic cost function when the residual below threshold T. Multiscale error diffusion also has a higher gain at edges than the

Floyd-Steinberg kernel. As a result, edges in the inverse halftone images from these two methods are overly sharpen.

We have following conclusions from this experiment.

- The proposed method is more robust w.r.t. different error diffusion kernels. Other inverse halftoning methods with pre-designed passbands may fail when halftoning algorithms perform visual enhancement such as the use of the Jarvis kernel.

- Compared to most other methods which are difficult to optimize for different halftone images, our approach can simply replace the cost function in the nonlinear filter to obtain better results. This is another advantage of the model-based filtering technique.

Haltone Algorithm		PSNR		
		Proposed	Adaptive[23]	Nonlinear[7]
Jarvis	Lena	25.76	24.65	25.29
	Peppers	25.42	23.91	24.35
Multiscale	Lena	28.29	26.73	28.16
	Peppers	29.07	27.99	28.71

Table 6.6: Performance of blind inverse halftoning algorithms applied to halftone images generated by different error diffusion algorithms.

6.5 Conclusion

In this chapter, we present an efficient blind inverse halftoning algorithm which yields good objective and subjective quality at a low computational cost and memory usage. Since the algorithm uses mostly integer arithmetic and local operations, it is suitable for efficient software and hardware implementation.

(a) Proposed

(b) Wavelet

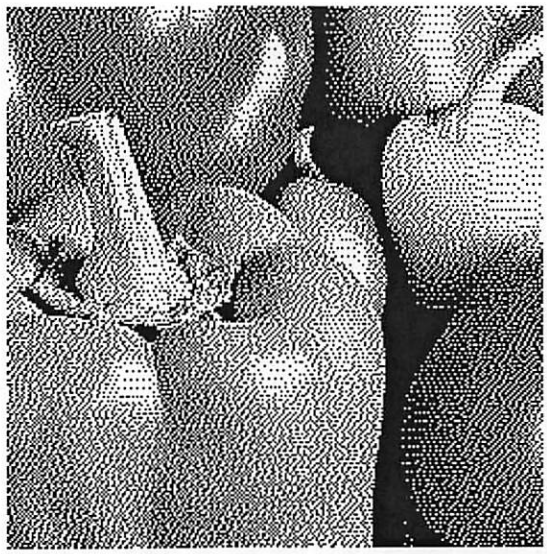
(c) Proposed

(d) Wavelet

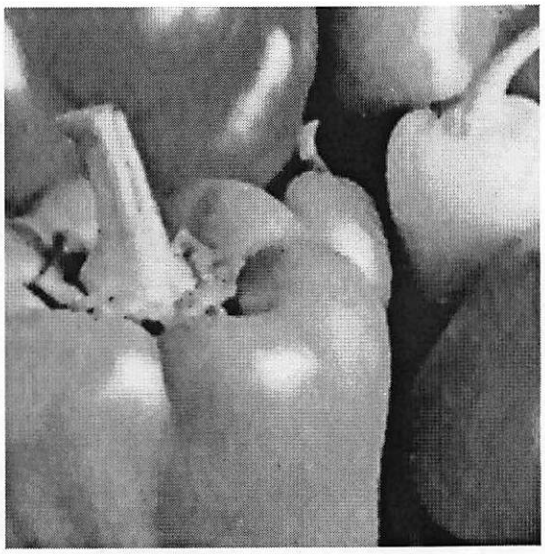
Figure 6.7: Inverse halftoning results of the proposed technique and results from the wavelet approach in [57].

(a) 3×3 (b) 7×7

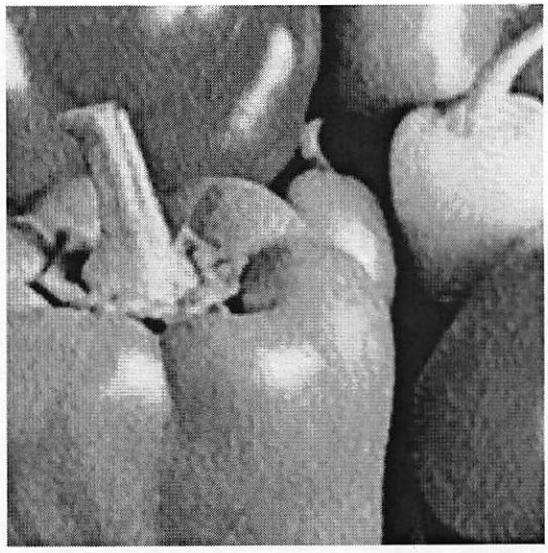
Figure 6.8: Results of nonlinear filtering with a different window size.



(a)Halftone

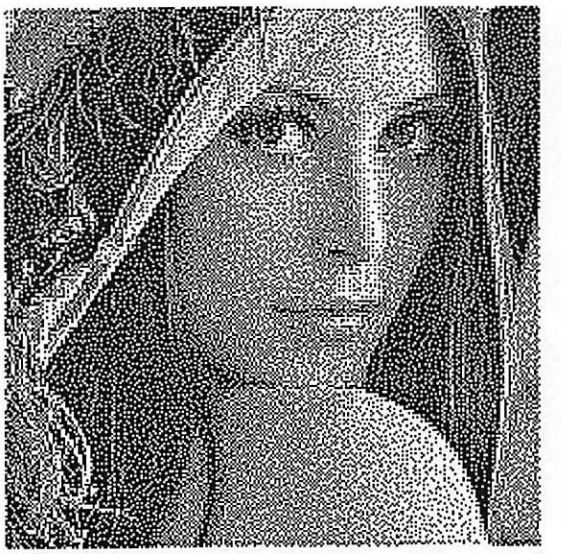


(b)Proposed



(c)Algorithm in [7]

Figure 6.9: The inverse halftoning of the diffused image with the Jarvis kernel.



(a)Halftone

(b)Proposed

(c)Algorithm in [7]

Figure 6.10: The inverse halftoning of a multiscale error diffused image.

Chapter 7

Conclusion and Future Work

7.1 Summary of the Research

In this thesis, we studied two ill-posed image recovery problems: compression artifact reduction and inverse halftoning. Previous work to solve these problems is usually based on iterative optimization techniques, which are usually too slow to be useful for practical applications. The primary goal of this research is to develop cost-effective solutions to these two problems.

7.1.1 Compression Artifact Reduction

A low complexity postprocessing algorithm based on robust estimation and nonlinear filtering was proposed. Unlike the previous post-filters which are often designed in an ad hoc fashion, we formulated the artifact reduction as a constrained optimization problem and solved it by using the nonlinear filtering technique. The proposed nonlinear filters can obtain sub-optimal estimates that are very close to the optimal one while substantially reducing the computational complexity.

The proposed method is generic and applicable to different compression algorithms, because robust estimation along with the MRF model provides a powerful and flexible framework for image smoothing. By properly selecting the potential function, the user is able to control the degree of smoothing while preserving or even enhancing the sharpness of edges. The noise constraint is added to ensure that the distortion introduced by the nonlinear filter is bounded and prevents features from being removed due to smoothing.

In the case of the postprocessing of compressed video, the constraint on the computational resource is even more critical. A fast table-lookup technique is adopted to replace the deblockinng filters for further complexity reduction.

The proposed postprocessing algorithm brings a significant improvement to the quality of low-bit-rate compressed images and video. It reduces noticeable artifacts existing in standard decompression techniques. Furthermore, real-time postprocessing is made possible due to the low complexity and the low memory requirement of the proposed algorithm.

7.1.2 Inverse Halftoning

A blind inverse halftoning algorithm of low complexity was presented in Chapter 6. The cascade of lowpass and nonlinear filters in conjunction with the edge enhancement technique produces high quality inverse halftoned images at a low computational cost.

Results of the proposed technique are comparable to that obtained by using sophisticated restoration-based techniques that require the knowledge of diffusion kernel. Besides, our method is more robust to different error diffusion kernels than other blind inverse halftoning algorithms since our nonlinear filtering is a model-based approach. That is, the smoothing is based on a generic image model for an ideal image. The advantage of using our technique is even more substantial when an image contains constant or slow varying gray levels in large areas separated by sharp edges.

7.2 Future Work

Since our nonlinear filtering solutions for these two inverse problems are sub-optimal approximations of results obtained via constrained optimization, we do not claim the optimality of the proposed methods. They however provide good solutions to these difficult inverse problems and meet practical constraints on the complexity and the memory requirement. There are several research topics worth further study in the future and presented below.

- **Further improvement of proposed postprocessing filters**

  1. The speed of the deringing filter will be improved. This can be achieved by reducing the computational redundancy when calculating minima of adjacent windows.
  2. The proposed filters have a close relationship with filters based on the order statistics, such as tri-mean filters, trimmed-mean filters and hybrid mean/median filters. The exact relationship deserves more investigation.
  3. The parameters in the noise constraint and the potential function can be adjusted adaptively according to local statistics and the estimated noise level.

- **Postprocessing by other nonlinear filters**
  We will investigate other types of nonlinear filters which may achieve a good performance in the postprocessing of compressed media. One interesting example is the cluster filter based on the maximum likelihood estimation criterion.

- **New applications of current postprocessing technique**
  We will look for new applications of the developed robust nonlinear filtering technique. A direct extension will be the applications requiring interpolation, such as resolution conversion, frame interpolation, and lossy block recovery etc. The robust technique can preserve discontinuities better then the traditional linear interpolation technique.

- **Postprocessing of other compressed media**
  Postprocessing of other compressed media is also an interesting topic to be explored. For example, the frame error in the low-bit-rate compressed audio could be reduced by some postprocessing technique.

- **Inverse halftoning for clustered-dot halftones**
  The extension of the proposed technique to clustered-dot halftone images may not be straightforward, since the screen sizes of clustered-dot halftones vary a lot. The halftone image that has a larger screen is generally more difficult to

be restored. Since clustered-dot halftoning is widely used in color halftoning for printing and display applications, it is worth further study.

- **Inverse halftoning for color halftone images**

  The proposed inverse halftoning method can potentially be extended to the solution of multi-level or color error-diffused halftones. However, there are concerns in color inverse halftoning. First, the information carried in a color halftone image is only an index to a predefined color palette. The correct inverse halftoning result can be obtained only if both halftone and inverse halftone systems using the same palette. Note that in binary halftone, we have common assumption that the palette contains black and white. Second, colors in the palette are not necessarily uniformly quantized from three color components and, in most cases, it is obtained by using perceptually optimized non-uniform quantizers. Therefore, simply applying the same inverse halftoning algorithm to each color space may not lead to a good result. The color inverse halftoning is a challenging problem which needs to be further explored.

- **Inverse halftoning for efficient lossy and lossless compression of a halftone image**

  Compared to compression of continuous-tone images, the binary image compression has less been addressed. Previous work mostly adopted the context-based coder and the dictionary-based approach. However, this is not efficient for error-diffused halftone images. The emerging JIBG2 standard has considered this problem and supported a new mode so that halftone images can be compressed at higher bit-depth. This allows a binary halftone to be compressed by other start-of-the-art lossy image compression techniques. However, this cannot be done without inverse halftoning and re-halftoning. What is the most efficient inverse halftoning for lossy compression remains open.

Reference List

[1] M. Analoui and J. Allebach, "New results on reconstruction of continuous-tone from halftone," in *Proc. IEEE ICASSP*, pp. 313–316, 1992.

[2] G. Bozkurt and A. E. Çetin, "A set theoretic inverse halftoning technique," in *in Proc. SPIE*, vol. 3653, Jan. 1999.

[3] R. Castagno and J. A. Villaroel, "A spline-based adaptive filter for the removal of block artifacts in image sequences coded at very low bitrate," in *Proc. IEEE International Conference on Image Processing*, Sept. 1996.

[4] S. S. O. Choy, Y.-H. Chan, and W.-C. Siu, "Reduction of block-transform image coding artifacts by using local statistics of transform coefficients," in *IEEE Signal Processing Letters*, pp. 5–7, Jan. 1997.

[5] R. R. Coifman and D. Donoho, "Translation-invariant de-noising," in *Wavelet Digest*, 1995.

[6] S. Comes and B. Macq, "Human visual quality criteria," in *SPIE: Visual Communication and Image Processing '90*, Oct. 1990.

[7] N. D.-Venkata, T. D. Kite, M. Venkataraman, and B. L. Evans, "Fast blind inverse halftoning," in *Proc. IEEE International Conference on Image Processing*, 1998.

[8] D. L. Donoho, "Denoising by soft-thresholding," *IEEE Trans. on Information Theory*, Vol. 41, pp. 613–627, 1995.

[9] DRAFT ITU-T Recommendation H.263, *Video coding for low bitrate communication*, Aug. 1997.

[10] Z. Fan and R. Eschbach, "Limited cycle behavior of error diffusion," in *Proc. IEEE International Conference on Image Processing*, vol. 2, pp. 1041–1045, 1994.

[11] R. Floyd and L. Steinberg, "An adaptive algorithm for spatial grayscale," in *in SID Int. Symp. Digest Technical Papers*, 1975.

[12] R. A. Gopinath, M. Lang, H. Guo, and J. E. Odegard, "Wavelet-based post-processing of low bit rate transform coded images," in *Proc. IEEE ICASSP*, pp. 913–917, 1994.

[13] S. Hein and A. Zakhor, "Halftone to continuous-tone conversion of error-diffusion coded images," *IEEE Trans. on Image Processing*, Vol. 4, No. 2, pp. 208–215, 1995.

[14] S.-W. Hong, Y.-H. Chan, and W.-C. Siu, "A practical real-time post-processing technique for blocking effect elimination," in *Proc. IEEE International Conference on Image Processing*, pp. II.21–II.24, Sept. 1996.

[15] ISO/IEC 11172, *Coding of moving pictures and associated audio - for storage at up to about* 1.5 Mbits/s, Nov. 1992.

[16] ISO/IEC 13818, *Generic coding of moving pictures and associated audio information*, Nov. 1994.

[17] ISO/IEC 14496-2, *Information technology - Coding of audio-visual objects: visual, committee draft*, Mar. 1998.

[18] ISO/IEC JTC1 CD 10918-1, *Digital Compression and Coding of Continuous-Tone Still Image, Part 1 Requirements and Guidelines*, Feb. 1991.

[19] ISO/IEC JTC1/SC29/WG1 N875, *JPEG 2000 Verification Model Version 1.0*, Jun. 1998.

[20] ITU-T Recommendation H.261, *Video codec for audiovisual services at $p \times 64$ kbit/s*, Mar. 1993.

[21] J. Jarvis, C. Judice, and W. Ninke, "A survey of techniques for displaying of continuous-tone pictures on bilevel displays," *Computer Graphics Image Processing*, Vol. 5, pp. 13–40, 1976.

[22] I. Katsavounidis and C. C. J. Kuo, "A multiscale error diffusion technique for digital halftoning," *IEEE Trans. on Image Processing*, Vol. 6, pp. 483–490, Mar. 1997.

[23] T. D. Kite, N. D.-Venkata, B. L. Evans, and A. C. Bovik, "A high quality, fast inverse halftoning algorithm for error diffusion halftones," in *Proc. IEEE International Conference on Image Processing*, 1998.

[24] A. Kundu, "Enhancement of JPEG coded images by adaptive spatial filtering," in *Proc. IEEE International Conference on Image Processing*, pp. 187–190, Oct. 1995.

[25] C. J. Kuo and R. J. Hsieh, "Adaptive postprocessing for block encoded images," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 5, pp. 298–304, Aug. 1995.

[26] K. Y. Kwak and R. Haddad, "Projection-based eigenvector decomposition for reduction of blocking artifacts of DCT coded image," in *Proc. IEEE International Conference on Image Processing*, pp. II.527–II.530, Oct. 1995.

[27] G. Lakhani, "Improved equations for JPEG's blocking artifacts reduction approach," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 7, pp. 930–934, 1997.

[28] J. Li and C.-C. J. Kuo, "Coding artifact removal with multiscale postprocessing," in *Proc. IEEE International Conference on Image Processing*, Oct. 1997.

[29] S. Z. Li, "Discontinuous MRF prior and robust statistics: a comparative study," *Image and Vision Computing*, Vol. 13, pp. 227–233, Apr. 1995.

[30] T. S. Liu and N. Jayant, "Adaptive post-processing algorithms for low bit rate video signals," in *Proc. IEEE ICASSP*, pp. 401–804, 1994.

[31] J. Lu, "Image deblocking via multiscale edge processing," in *Proc. SPIE*, Aug. 1996.

[32] J. Luo, C.-W. Chen, and K. J. Parker, "Image enhancement for low bit rate wavelet-based compression," in *International Symposium on Circuits and Systems*, June 1997.

[33] J. Luo, C. W. Chen, K. J. Parker, and T. Huang, "Artifacts reduction in low bit rate DCT-based image compression," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 5, No. 9, pp. 1363–1368, 1996.

[34] W. E. Lynch, A. R. Reibman, and B. Liu, "Postprocessing transform coded images using edges," in *Proc. IEEE ICASSP*, pp. 2323–2326, May 1995.

[35] B. Macq, M. Mattavelli, O. V. Claster, E. van der Plancke, S. Comes, and W. Li, "Image visual quality restoration by cancellation of the unmasked noise," in *Proc. IEEE ICASSP*, pp. V.53–V.56, 1994.

[36] S. Mallat and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Trans. on Pattern Anal. Machine Intell.*, Vol. 14, pp. 710–732, July 1992.

[37] T. Mitsa and K. J. Parker, "Digital halftoning technique using a blue noise masking," *J. Opt. Soc. Amer*, Vol. 9, pp. 1920–1929, 1992.

[38] Y. Nakajima, H. Hori, and T. Kanoh, "A pel adaptive reduction of coding artifacts for MPEG video signals," in *Proc. IEEE ICASSP*, pp. 928–932, 1994.

[39] M. Nikolova, "Regularization functions and estimators," in *Proc. IEEE International Conference on Image Processing*, Sept. 1996.

[40] B. Niss, *Prediction of AC coefficients from the DC values*, ISO/IEC JTC1/SC2/WG8 N745, May 1988.

[41] T. P. O'Rourke and R. Stevenson, "Improved image decompression for reduced transform coding artifacts," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 5, No. 6, pp. 490–499, 1995.

[42] T. Ozcelik, J. C. Brailean, and A. K. Katsaggelos, "Image and video compression algorithms based on recovery techniques using mean field annealing," *Proceedings of the IEEE*, Vol. 83, No. 2, pp. 304–315, 1995.

[43] T. N. Pappas, "Model-based techniques for digital halftoning," in *Proc. IEEE International Conference on Image Processing*, pp. 26–30, 1994.

[44] B. Ramamurthi and A. Gersho, "Nonlinear space variant postprocessing of block coded images," *IEEE Trans. on Acoustic, Speech, and Signal Processing*, Vol. 34, pp. 1258–1267, Oct. 1986.

[45] H. C. Reeves and J. S. Lim, "Reduction of blocking effects in image coding," *Optical Engineering*, Vol. 23, pp. 34–37, Jan/Feb 1984.

[46] R. Rosenholtz and A. Zakhor, "Iterative procedures for reduction of blocking effects in transform coding," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 1, pp. 81–94, Oct. 1992.

[47] K. Sauer, "Enhancement of low bit-rate coded images using edge detection and estimation," *Computer, Vision, Graphics, and Image Processing*, Vol. 53, pp. 52–62, Jan. 1991.

[48] R. Stevenson, "Reduction of coding artifacts in transform image coding," in *Proc. IEEE ICASSP*, Apr. 1993.

[49] R. Stevenson, "Reduction of coding artifacts in low-bit-rate video coding," in *Proceedings of the 1995 Midwest Symposium on Circuits and Systems*, Aug. 1995.

[50] R. L. Stevenson, "Inverse halftoning via MAP estimation," *IEEE Trans. on Image Processing*, Vol. 6, pp. 574–583, Apr. 1997.

[51] R. Ulichney, *Digital Halftoning*, Cambridge, MA:MIT Press, 1987.

[52] R. Ulichney, "Dithering with blue noise," in *Proc. IEEE International Conference on Image Processing*, vol. 76, pp. 56–79, Jan. 1988.

[53] R. Ulichney, "The void-and-cluster method for dither array generation," in *in Proc. SPIE*, vol. 1913, pp. 332–343, Feb. 1993.

[54] P. W. Wong, "Inverse halftoning and kernel estimation for error diffusion," *IEEE Trans. on Image Processing*, Vol. 4, pp. 486–498, Apr. 1995.

[55] P. W. Wong, "Entropy constrained halftoning using multiscale tree coding," *IEEE Trans. on Image Processing*, Vol. 6, pp. 1483–1579, Nov. 1997.

[56] S.-W. Wu and A. Gersho, "Improved decoder for transform coding with application to the JPEG baseline system," *IEEE Trans. on Communications*, Vol. 40, pp. 251–254, Feb. 1992.

[57] Z. Xiong, M. T. Orchard, and K. Ramchandran, "Inverse Halftoning using Wavelets," in *Proc. IEEE International Conference on Image Processing*, pp. 569–572, 1996.

[58] Z. Xiong, M. T. Orchard, and Y. Q. Zhang, "A simple de-blocking algorithm for JPEG compressed images using overcomplete wavelet representations," in *in Proc. ISCAS*, May 1997.

[59] L. Yan, "A nonlinear algorithm for enhancing low bit-rate coded motion video sequence," in *Proc. IEEE ICASSP*, pp. 923–927, 1994.

[60] Y. Yang, N. Galatsanos, and A. K. Katsaggelos, "Regularized reconstruction to reduce blocking discrete cosine transform compressed images," *IEEE Trans. on Circuits and Systems for Video Technology*, Vol. 3, No. 6, pp. 421–432, 1993.

[61] Y. Yang, N. Galatsanos, and A. K. Katsaggelos, "Projection-based spatially adaptive reconstruction of block- transform compressed images," *IEEE Trans. on Image Processing*, Vol. 4, No. 7, pp. 896–908, 1995.

[62] A. Zakhor and S. L. F. Eskafi, "A new class of B/W halftoning algorithms," *IEEE Trans. on Image Processing*, Vol. 2, pp. 499–509, Oct. 1993.