

# Measurement of Printer Parameters for Model-Based Halftoning

by

Chen-Koung Dong

Submitted to the Department of Electrical Engineering and  
Computer Science

in partial fulfillment of the requirements for the degrees of

Bachelor of Science

and

Master of Science in Electrical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1992

©AT&T, 1992. M. . . .

MASSACHUSETTS INSTITUTE  
OF TECHNOLOGY

JUL 10 1992

LIBRARIES  
ARCHIVES

Author . . .

Department of Electrical Engineering and Computer Science

May 1, 1992

Certified by . . .

Thrasylvoulos N. Pappas

Member of Technical Staff, AT&T Bell Laboratories

Thesis Supervisor

Certified by . . .

Andrew Lippman

Associate Director, The Media Laboratory

Thesis Supervisor

Accepted by

Campbell L. Searle

Chairman, Departmental Committee on Graduate Students

# Measurement of Printer Parameters for Model-Based Halftoning

by

Chen-Koung Dong

Submitted to the Department of Electrical Engineering and Computer Science  
on April 30, 1992, in partial fulfillment of the  
requirements for the degrees of  
Bachelor of Science  
and  
Master of Science in Electrical Engineering

## Abstract

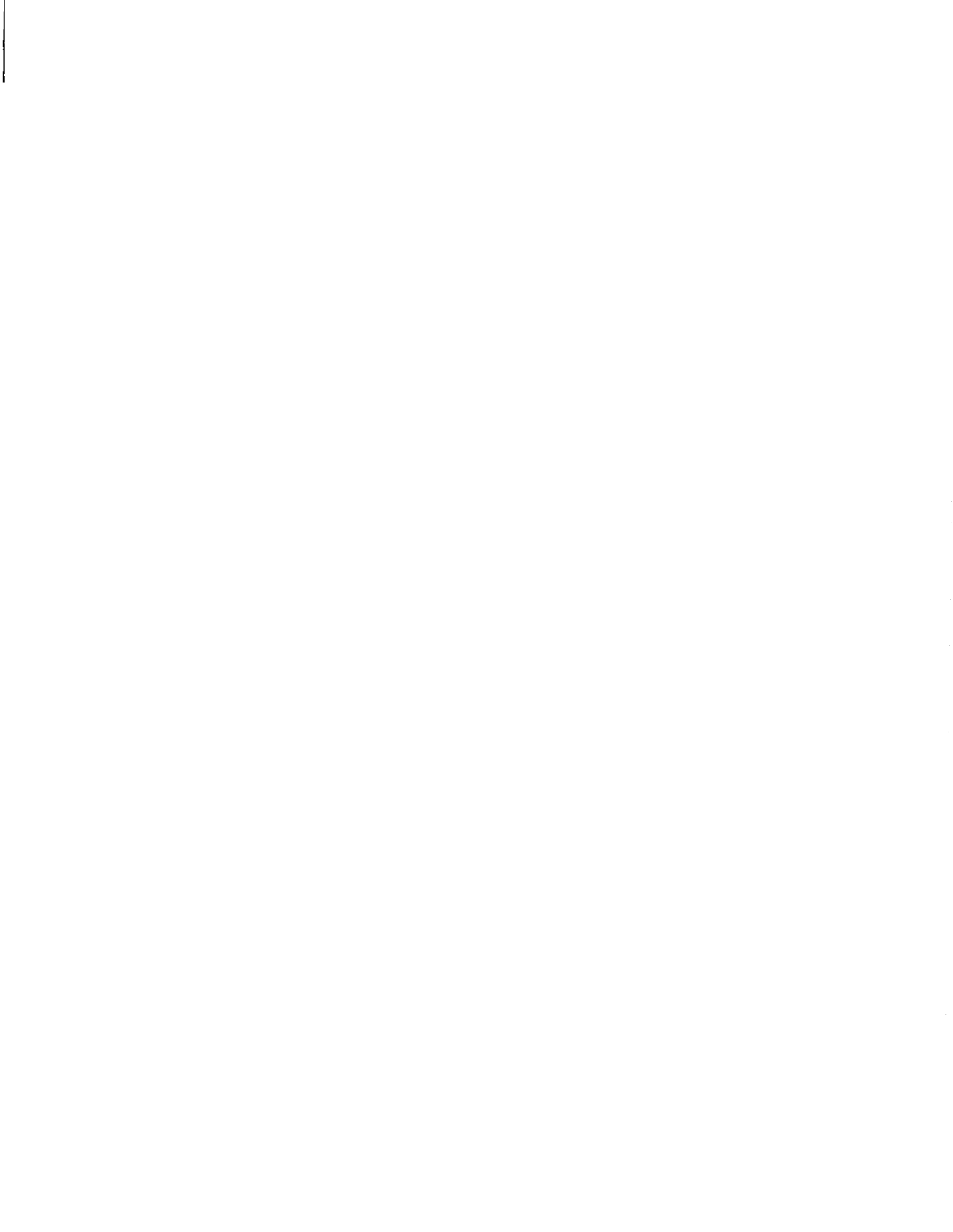
We present a new approach for estimating printer model parameters that can be applied to a wide variety of laser printers. Recently developed "model-based" digital halftoning techniques depend on accurate printer models to produce high quality images using standard laser printers (typically 300 dpi). Traditional halftoning techniques are designed to be fairly robust to printer distortions at the expense of spatial and gray-scale resolution. Since printer characteristics vary considerably, e.g. write-black vs. write-white laser printers, the model parameters must be adapted to each individual printer. Previous approaches for estimating the printer model parameters are based on a physical understanding of the printing mechanism. One such approach uses the "circular dot-overlap model," which assumes that the laser printer produces circularly shaped dots of ink. The "circular dot-overlap model" is an accurate model for many printers but cannot describe the behavior of all printers. The new approach is based on measurements of the gray level produced by various test patterns, and makes very few assumptions about the laser printer. We use a reflectance densitometer to measure the average brightness of the test patterns, and then solve a constrained optimization problem to obtain the printer model parameters. To demonstrate the effectiveness of the approach, the model parameters of two laser printers with very different characteristics were estimated. The printed models were then used with both the modified error diffusion and the least-squares model-based approach to produce printed images with the correct gray-scale rendition. We also derived an iterative version of the modified error diffusion algorithm that improves its performance.

Thesis Supervisor: Thrasyvoulos N. Pappas  
Title: Member of Technical Staff, AT&T Bell Laboratories

Thesis Supervisor: Andrew Lippman  
Title: Associate Professor

## Acknowledgments

I would like to give special thanks to Thrasos Pappas for guiding me through my thesis and for all his great insights and witty remarks. To David Neuhoff for initiating the work on estimation of printer model parameters. Special thanks to Steve Levinson for providing the optimization routines from the Harwell library. Also to my parents and my brother for being supportive. And to all my friends and co-workers for making my stay at AT&T so enjoyable.



# Contents

<b>1</b>	<b>Introduction</b>	<b>9</b>
<b>2</b>	<b>Standard Halftoning Techniques</b>	<b>14</b>
2.1	Screening . . . . .	14
2.1.1	Clustered-Dot Approach . . . . .	16
2.1.2	Dispersed-Dot Approach . . . . .	17
2.2	Error Diffusion . . . . .	17
2.3	Summary . . . . .	20
<b>3</b>	<b>Model-Based Halftoning</b>	<b>21</b>
3.1	Eye Models . . . . .	21
3.2	Printer Models . . . . .	22
3.2.1	Circular Dot-Overlap Model . . . . .	25
3.2.2	Measurement Approach . . . . .	26
3.3	Modified Error Diffusion . . . . .	27
3.4	Least-Squares Model-Based Method . . . . .	29
3.4.1	One-dimensional Least-squares . . . . .	31
<b>4</b>	<b>Measurement of Printer Model Parameters</b>	<b>32</b>
4.1	One-dimensional Measurement Model . . . . .	33
4.1.1	Formulation . . . . .	34
4.1.2	Window Size . . . . .	38
4.1.3	Selection of Equations . . . . .	39
4.2	Two-dimensional Measurement Model . . . . .	40

4.2.1	Formulation . . . . .	40
4.2.2	Window Shape and Size . . . . .	44
4.2.3	Selection of Equations . . . . .	45
4.3	Constrained Optimization . . . . .	46
4.3.1	Solution Vector . . . . .	47
<b>5</b>	<b>Application</b>	<b>49</b>
5.1	Reflectance Measurements . . . . .	49
5.1.1	Multiple Internal Reflection . . . . .	50
5.1.2	Densitometer Calibration . . . . .	52
5.2	One-Dimensional Least-Squares Model-Based Halftoning . . . . .	53
5.3	Modified Error Diffusion . . . . .	55
5.3.1	Causality Constraints . . . . .	56
5.3.2	Multi-pass Error Diffusion . . . . .	58
5.4	Least-Squares Model-Based Halftoning . . . . .	59
<b>6</b>	<b>Conclusions</b>	<b>60</b>

# List of Figures

1-1	Digital halftoning . . . . .	10
2-1	Ordered dither threshold matrices . . . . .	15
2-2	Standard error diffusion . . . . .	19
2-3	Jarvis, Judice and Ninke error filter . . . . .	19
3-1	The spatial frequency sensitivity of the eye . . . . .	23
3-2	Eye filter . . . . .	23
3-3	Definition of $\alpha$ , $\beta$ and $\gamma$ . . . . .	26
3-4	Error Diffusion With Ink-Spreading Compensation . . . . .	28
3-5	Error calculation in modified error diffusion . . . . .	29
3-6	The Visual and Printer Perceptual Models. . . . .	30
4-1	Sample periodic pattern with period "1011" . . . . .	36
4-2	Sample periodic pattern and corresponding equation . . . . .	41
4-3	Symmetry reduction . . . . .	43
4-4	Different window sizes and shapes . . . . .	45
5-1	Reflectance calibration . . . . .	51
5-2	Multiple internal reflection . . . . .	52
5-3	Mountain Effect . . . . .	55

# List of Plates

1	Dispersed vs. clustered ordered dither .....	64
2	Error diffusion .....	65
3	300 dots/inch printing .....	66
4	One-dimensional least-squares model-based halftoning .....	67
5	Modified error diffusion .....	68
6	Modified error diffusion – Circular dot-overlap model (CDOM) vs. Measurement model (MM) .....	69
7	Modified error diffusion – Cross-shaped window model .....	70
8	Least-squares model-based halftoning .....	71
9	Write-white printer model .....	72
10	Circular dot-overlap model (CDOM) vs. Measurement model (MM) .....	73



# Chapter 1

## Introduction

Digital halftoning is the process of generating a pattern of binary pixels that the eye perceives as a continuous-tone image. Digital halftoning is necessary for display of gray-scale images in media in which the direct rendition of gray tones is impossible. Examples of such media include paper and binary CRT displays. In this thesis we examine “model-based” digital halftoning techniques that have been developed recently [1, 2] and depend on accurate printer models to produce high quality images using standard laser printers (typically 300 dpi). The goal of this thesis is to develop an experimental procedure for estimating model parameters for a wide variety of printers.

Model-based halftoning can be used to improve the quality of gray-scale images transmitted by facsimile. A new approach to gray-scale facsimile is proposed in [3], in which the image is transmitted in gray-scale form using high fidelity image coders. This approach allows more efficient encoding of data and, more importantly, permits the halftoner to be tuned to the individual printer on which the document is printed.

Figure 1-1 shows how halftoning works. A halftoning algorithm generates a binary pattern of pixels which is printed and perceived by the eye. All halftoning techniques rely on the fact that the eye acts as a spatial low-pass filter. Essentially, the eye averages the brightness over an area and perceives a gray level proportional to the number of black spots (represented by 1's) and white spaces (represented by 0's) in that area. A considerable research effort has been devoted toward understanding

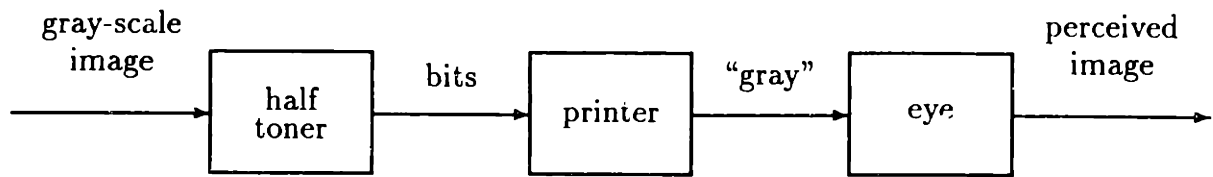


Figure 1-1: Digital halftoning

human perception in the past, and various models of the eye have been proposed. Some of the model-based techniques we examine in this thesis use an eye model that is based on estimates of the spatial frequency sensitivity of the eye by Mannos and Sakrison [4].

Another important factor affecting the performance of a halftoning technique is the behavior of the display device. Most halftoning techniques assume that the displayed binary pattern consists of identically shaped dots of two colors, usually on a rectangular grid. This assumption does not hold for most printing devices which introduce significant distortions. Such distortions make most halftoning techniques unsuitable for printers. Printer characteristics are known to vary considerably from printer to printer. For example write-black laser printers have very different characteristics than write-white laser printers. Traditional halftoning techniques are designed to be fairly robust to printer distortions. As a result, they compromise both spatial and gray-scale resolution. Model-based halftoning techniques [1, 2, 5], on the other hand, exploit the characteristics of each particular printer to maximize the quality of the printed images. Thus, they depend on an accurate printer model, whose parameters must be adapted to each individual printer.

The printer model describes the “gray” level, that is the reflectance, of the printed pattern. The parameters of the printer model can be derived from a physical understanding of the printing mechanism or from direct measurements of the reflectance of various printed patterns. One approach for estimating the printer model parameters uses the “circular dot-overlap model,” which assumes that the laser printer produces circularly shaped dots of ink [1]. The gray level at each pixel of the printed image

is calculated to be the area of the pixel that is covered by ink. Thus, by estimating the radius of these circularly shaped ink dots, the gray values at each location can be calculated.

The “circular dot-overlap model” is an accurate model for many printers but cannot describe the behavior of all printers. The black dots produced by actual printers are not perfectly round, they are not perfectly black, and their size and shape may depend on the presence of adjacent dots. For example, as was pointed out in [1], a white line surrounded by several black lines is brighter than when surrounded by two single lines. For some printers the circular dot approximation may not be valid at all. In this thesis we develop an approach for estimating printer model parameters that is based on measurements of the gray level produced by various test patterns, and makes very few assumptions about the laser printer. The “measurement approach” is intended to be general so that it can be applied to any write-white or write-black printer.

The first step in developing the measurement approach is to select a set of test patterns to be generated. The reflectance of the printed test patterns is then measured using a reflection densitometer. The densitometer measures the average reflectance of each pattern over an area whose diameter is roughly 4 mm. As was suggested in [6], the measured reflectance can be related to the printer model parameters by a set of linear equations. This system of equations cannot be solved directly because of discrepancies in the measurements and, more importantly, because of possible rank deficiency of the associated matrix. Instead, we formulate a constrained optimization problem which incorporates various constraints on the model parameters. Finally, the constrained optimization problem is solved by standard iterative techniques to provide the printer model parameters.

To demonstrate the effectiveness of the approach, we estimated the model parameters of two laser printers with very different characteristics. The printed models were then used with two model-based halftoning techniques: the modified error diffusion algorithm and the least-squares model-based (LSMB) halftoning algorithm. The most important criterion for the effectiveness of the measurement approach is the accuracy

of the gray-scale rendition of the printed images.

The modified error diffusion algorithm [1] uses the printer model to predict the gray level of the printed pixels and thus account for printer distortions. This algorithm provides high quality reproductions with reasonable complexity. However, since the error diffusion algorithm is causal and the printer models are noncausal, certain problems may arise. For the circular dot-overlap printer model this results in a small (barely noticeable) bias in the darkness of the printed images. For the measured printer models this may result in instabilities in the algorithm. We propose two solutions to this problem. One is to impose further constraints on the printer parameters. The other is to use an iterative (multi-pass) version of the modified error diffusion algorithm. The multi-pass error diffusion algorithm eliminates the instabilities, produces the correct gray-scale, and even eliminates some low-frequency artifacts.

One drawback of the error diffusion algorithm is that it does not make use of an explicit eye model. This results in some well known artifacts and asymmetries [1]. In contrast to the modified error diffusion algorithm, the least-squares model-based halftoning technique [2] exploits both a printer model and a model of visual perception. It produces an “optimal” halftoned reproduction, by finding the binary image that minimizes the squared error between the output of the cascade of printer and visual models in response to the binary image and the output of the visual model in response to the original gray-scale image. The solution to the two-dimensional least-squares problem is only approximate and is obtained by iterative techniques. The LSMB method produces images that are sharper than those produced by the modified error diffusion algorithm. The least-squares approach eliminates the problems associated with error diffusion. More importantly, the gray scale rendition of the halftoned images is good.

The experimental results demonstrate that when the measured model parameters are used with the two model-based techniques, they produce the correct gray scale and maintain the overall performance of these techniques. For one of the laser printers we used, the circular dot-overlap model and the measurement approach produce results

that are equally good. The circular dot-overlap model does not apply to the second laser printer. Thus, even though the measurement approach is considerably more complicated than the circular dot-overlap model, it can be used to establish the limitations of simpler and easier to calibrate models.

The remainder of this thesis is organized as follows. Chapter 2 discusses the standard halftoning techniques. In Chapter 3, the model-based approaches to halftoning are reviewed. Our results are presented in Chapters 4 and 5. Chapter 4 describes the measurement approach for estimating printer model parameters. The application of the measurement approach to specific laser printers and halftoning techniques is presented in Chapter 5. The conclusions are summarized in Chapter 6.

One special note concerning the plates that are included at the end of the report should be mentioned here. These plates are original laser printer outputs and should not be xeroxed. They illustrate the characteristics of a particular laser printer and demonstrate the performance of the various halftoning techniques. Xeroxed copies of these plates may severely distort the characteristics of the printed images.

# Chapter 2

## Standard Halftoning Techniques

Digital halftoning is the process of converting a continuous-tone gray scale image into a binary image. This process is necessary in the reproduction of images in media where pixels can take only two possible values, usually black and white. The illusion of continuous tone images in such media is created by the arrangement of pixels of the two colors in specified formations. In the past, digital halftoning algorithms have been studied extensively (see [7]). The most popular techniques currently used are *screening* and *error diffusion*.

In our study of halftoning techniques, it will be assumed that the gray-scale images have been sampled so there is one pixel for each binary value generated. When the samples of a given image are fewer than the number of dots to be generated, interpolation is necessary. In this thesis, the test images were upsampled by a factor of 2 using an interpolation scheme consisting of an *expander* [8, pp. 105–109] and an equiripple low-pass FIR filter [8, pp. 462–480]. For the specification of the binary values of the halftoned image, we will use a “1” to denote a black pixel and a “0” to denote a white pixel.

### 2.1 Screening

Traditional halftoning relies on a simple but effective technique known as screening. In screening, the image is compared, pixel by pixel, to an array of image-independent

.576	.635	.608	.514	.424	.365	.392	.486
.847	.878	.910	.698	.153	.122	.090	.302
.820	.969	.941	.667	.180	.031	.059	.333
.725	.788	.757	.545	.275	.212	.243	.455
.424	.365	.392	.486	.576	.635	.608	.514
.153	.122	.090	.302	.847	.878	.910	.698
.180	.031	.059	.333	.820	.969	.941	.667
.275	.212	.243	.455	.725	.788	.757	.545

(a) "Classic-4" ( $8 \times 8$ )

.513	.272	.724	.483	.543	.302	.694	.453
.151	.755	.091	.966	.181	.785	.121	.936
.634	.392	.574	.332	.664	.423	.604	.362
.060	.875	.211	.815	.030	.906	.241	.845
.543	.302	.694	.453	.513	.272	.724	.483
.181	.785	.121	.936	.151	.755	.091	.956
.664	.423	.604	.362	.634	.392	.574	.332
.030	.906	.241	.845	.060	.875	.211	.815

(b) "Bayer-5" ( $8 \times 8$ )

Figure 2-1: Ordered dither threshold matrices

thresholds to produce the binary array that specifies the halftone image. The binary pixel is black whenever the gray level of the image is greater than the corresponding threshold and white otherwise.

There are many ways to generate the array of thresholds. When the thresholds are generated randomly, the technique is called *random dither*. Random dither is seldom used in practice because the patterns it produces have visible and objectionable low-frequency artifacts [7]. When the thresholds are periodic, the technique is called *ordered dither*. The threshold array is specified by one period which is typically a rectangular array, e.g. an  $8 \times 8$  matrix. Figure 2-1 shows the threshold matrices we used in this thesis. The matrix in Figure 2-1(a) specifies a *clustered* ordered dither technique. The thresholds are arranged so that they produce clusters of black bits. The matrix in Figure 2-1(b) specifies a *dispersed* ordered dither technique. The thresholds are arranged so that they produce dispersed black and white bits.

The popularity of screening techniques is attributed to their simplicity and ease of implementation. In the following two subsections we evaluate the performance of ordered dither techniques. The resolution of the images we used to test the halftoning methods in this thesis is  $460 \times 460$  pixels. We upsampled by a factor of 2 to obtain  $920 \times 920$  pixels that result in  $3.07 \times 3.07$  inch printed images. The plates at the end of the thesis were all printed on a 300 dpi HP LASERJET II printer, except Plate 9, which was printed on a DATA PRODUCTS LZR 2665 printer.

### 2.1.1 Clustered-Dot Approach

Clustered ordered dither is designed to simulate traditional analog halftoning techniques used in printing. Its main advantage is that it produces images that are very robust to dot overlap and other printer distortions. When the ink dots are printed in clusters or *macrodots*, most of the black dots overlap other black dots rather than white spaces. Thus the effect of dot overlap is minimized, and the accuracy of gray-scale rendition of the printed image is more or less maintained.

The macrodots are formed by choosing the elements of the threshold matrix so that they decrease towards a fixed point. In the case of the “Classic-4” threshold matrix, two macrodots will be formed, one in the lower left quadrant and one in the upper right. The spacing of the dots is fixed. The darker the image the bigger the macrodots that are generated. This is precisely how traditional analog halftoning works.

Plate 1 demonstrates the robustness of clustered-dot techniques. Plate 1(a) shows two test images halftoned with the “Classic-4” screen and printed on a write-black printer with a fair amount of dot overlap. The images in Plate 1(b) were halftoned with the “Bayer-5” screen, which disperses the dots, and results in printed images that are too dark. Thus, clustered screens like “Classic-4” are generally preferred for printing in the presence of dot overlap.

Although the rendition of gray levels is fairly reasonably maintained in the clustered dot approach, there are serious drawbacks. For one, the spatial resolution of the printed images is compromised. It can be seen in Plate 1 that “Bayer-5” pro-



duces images that are sharper, even though their gray-scale is distorted. Also, the macrodots of the clustered technique are more visible and unpleasant to the eye than the patterns of the dispersed technique. We usually refer to such patterns as low-frequency artifacts. As the period of the screen increases, the number of gray levels we can generate increases but the macrodots become more visible. Thus there is a tradeoff between the gray-scale resolution and the severity of the low-frequency artifacts. The low-frequency artifacts become more severe when we use printers with lower resolutions.

Finally, even though the clustered-dot approach minimizes the effects of ink spreading, the rendition of gray scale is still not perfect. The resulting images do not accurately reflect the original gray levels. The effects of printer distortion are still apparent for some printers. This can be corrected using direct measurement of printed images [7, p. 36], or a printer model [9].

### **2.1.2 Dispersed-Dot Approach**

In dispersed-dot techniques, the threshold matrix is designed to maximize the spreading of the dots in the printed image. This increases the spatial resolution of the printed images and minimizes the low-frequency artifacts. The disadvantage of dispersed-dot techniques is that they suffer major degradations in terms of gray-scale rendition. The images produced tend to be too dark or too light, depending on the type of printer used. As we saw in Plate 1(b), “Bayer-5” produces sharper printed images with less objectionable low-frequency artifacts than “Classic-4,” but also results in far greater distortions of the gray scale.

## **2.2 Error Diffusion**

A halftoning technique that produces sharper images than the screening techniques is error diffusion [10]. In error diffusion, the image pixels are compared to thresholds, like in screening, but in this case the threshold for each image pixel is dependent upon “prior” image pixels (usually above and to the left). Alternatively, each image

pixel is compared to a fixed threshold, after a correction factor is added to its original gray level to account for past “errors.” Thus, error diffusion uses feedback to correct for errors it has committed in the past. We will outline the error diffusion algorithm below. A block diagram of the algorithm is shown in Figure 2-2.

We first introduce the following notation in our description of error diffusion. Let  $[y_{i,j}]$  be a gray-scale image (after interpolation), where  $y_{i,j}$  denotes the pixel located at the  $i$ -th column and the  $j$ -th row of a Cartesian grid. Without loss of generality, we assume that the image is scanned left to right top to bottom. The binary image  $[b_{i,j}]$  produced by error diffusion is obtained by the following set of equations [1]

$$v_{i,j} = y_{i,j} - \sum_{m,n} h_{m,n} e_{i-m,j-n} \quad (2.1)$$

$$b_{i,j} = \begin{cases} 1, & \text{if } v_{i,j} > t \\ 0, & \text{otherwise} \end{cases} \quad (2.2)$$

$$e_{i,j} = b_{i,j} - v_{i,j} \quad (2.3)$$

Here  $v_{i,j}$  is the “corrected” value of the gray-scale image. The error  $e_{i,j}$  at any “instant”  $(i,j)$  is defined as the difference between the “corrected” gray-scale image and the binary image. The “past” errors are low-pass filtered and subtracted from the current image value  $y_{i,j}$  before it is thresholded to obtain the binary value  $b_{i,j}$ , where  $[h_{i,j}]$  is the impulse response of the low-pass filter. Thus, errors are “diffused” over the image.

The threshold  $t$  is typically fixed at 0.5, the middle of the gray-scale range. The low-pass filter  $[h_{i,j}]$  has non-symmetric half-plane support, so that only the past and already computed errors are filtered. This is the two-dimensional equivalent of “causality,” and enables the algorithm to make instantaneous decisions at each point. Thus, error diffusion requires only one pass through the data. The filter coefficients are positive and their sum is equal to one. This guarantees stability.

Various error diffusion filters have been suggested in the literature (see [7]). In our experiments we will use the filter proposed by Jarvis, Judice and Ninke [11], shown in Figure 2-3.

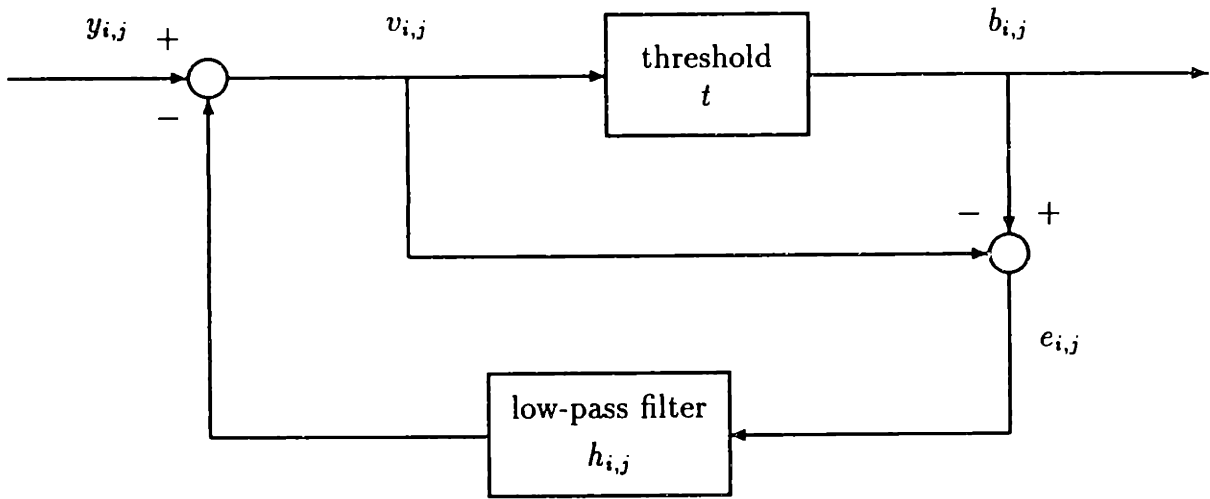


Figure 2-2: Standard error diffusion

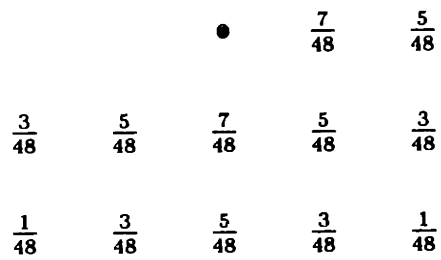


Figure 2-3: Jarvis, Judice and Ninke error filter

Error diffusion allows more flexibility in the placement of dots than screening techniques. This results in sharper images. However, as with all dispersed-dot schemes, error diffusion is very sensitive to printer distortions such as dot overlap. In the presence of dot overlap, error diffusion produces very dark images in write-black printers, as can be seen in Plate 2(a). This has limited its application to cases where no dot overlap occurs, such as binary CRT displays.

## 2.3 Summary

In the absence of significant printer distortions, e.g. for display on most binary CRTs, the best of the currently used techniques is error diffusion [7]. Ordered dither, which is simpler but inferior, is used when the amount of computation is an issue. When there is substantial dot overlap, as with laser printers, the clustered ordered dither schemes have been the only available choice [7]. As we discussed, the clustered dot approach is successful in reducing the effects of dot overlap, but it sacrifices spatial resolution and generates more low-frequency artifacts. As we will see in the next chapter, model-based halftoning techniques can correct for the effects of dot overlap without any sacrifices in spatial resolution and low-frequency artifacts. Actually, they can even exploit printer distortions to increase the gray-scale resolution. The key to such techniques is an accurate printer model, whose parameters are adapted to each individual printer.

# Chapter 3

## Model-Based Halftoning

The correct modeling of printer behavior and human perception can be used to improve the quality of printed images. This is the goal of model-based halftoning techniques which are examined in this chapter. We are primarily interested in laser printers, which generate “distortions” such as “dot overlap.” Conventional methods, such as clustered-dot ordered dither, resist distortions at the expense of spatial and gray-scale resolution. Model-based methods, on the other hand, rely on printer models to exploit printer distortions in order to increase both gray-scale and spatial resolution.

In the following sections, we consider two model-based techniques. The *modified error diffusion* algorithm [1] uses the properties of the eye implicitly, while the *least-squares model-based* method [2, 5] uses an explicit eye model. Both methods use a printer model explicitly. We will begin our discussion on model-based techniques by presenting the eye models and the printer models.

### 3.1 Eye Models

Human perception has long been a topic of great interest. Much research has been devoted toward understanding the behavior of the eye. The theory behind human perception is a complex one, and no one model of the eye can fully account for the complexity of human perception. However, for the purpose of this work, a simple

model of human perception is enough to fulfill our need.

When viewing an image, the human eye has the peculiar characteristic of averaging the brightness over a small area. This area depends on the viewing distance. The farther the viewer is away from the image, the larger this area becomes. The physical behavior of the eye can be characterized in the spatial frequency domain as follows. Essentially, the eye acts as a low-pass filter. Thus, the eye filters out the high-frequency components of the image and passes the low-frequency components. The frequency domain description of the eye characteristic should reflect the fact that its response depends on the viewing distance. To accomplish this, spatial frequency is measured in terms of cycles per degree.

Numerous researchers have estimated the spatial frequency sensitivity of the eye, often called the modulation transfer function (MTF). Typical of such is the estimate which Mannos and Sakrison developed.

$$H(f) = 2.6 (0.0192 + 0.114 f) \exp \left\{ -(0.114 f)^{1.1} \right\} \quad (3.1)$$

where  $f$  is in cycles/degree. This MTF is plotted in Figure 3-1. As indicated by Equation (3.1), the eye is most sensitive to frequencies around 8 cycles/degree. According to this model, the eye is a band-pass filter. However, it is the attenuation of the high frequencies that is the most critical to halftoning. For model-based halftoning, we adopted the eye model that was used in [5]. It is an FIR filter whose frequency response closely resembles the Mannos and Sakrison model. However, instead of a band-pass filter, this filter is low-pass. The impulse and frequency responses of this filter are shown in Figure 3-2. A two-dimensional eye filter can be obtained as a separable combination of two one-dimensional filters, as in [2].

## 3.2 Printer Models

The purpose of a printer model is to predict the “gray” levels produced by a printer. Accurate and objective predictions of gray level make it possible, not only to correct

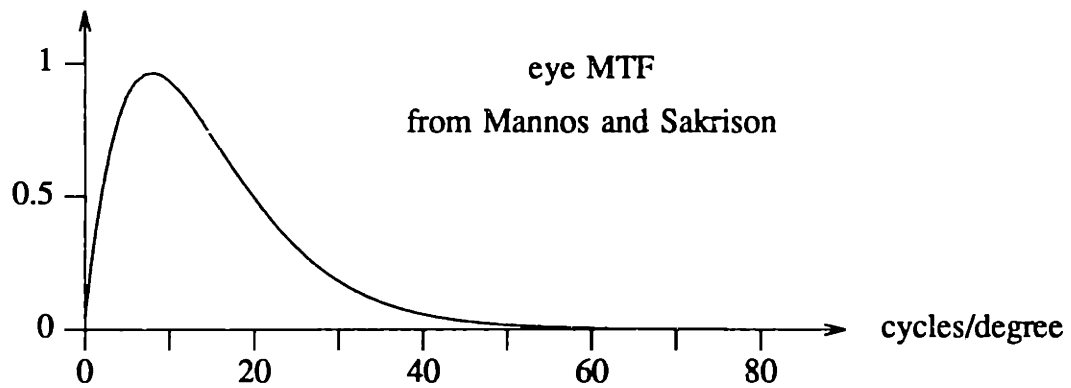


Figure 3-1: The spatial frequency sensitivity of the eye

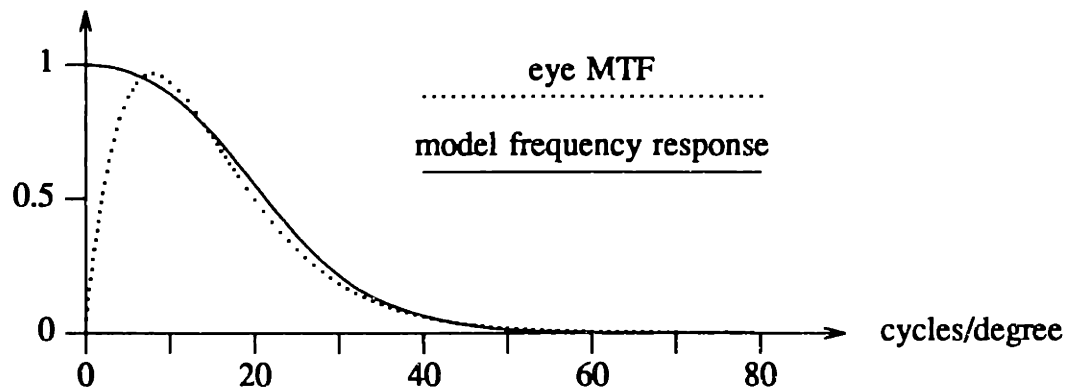
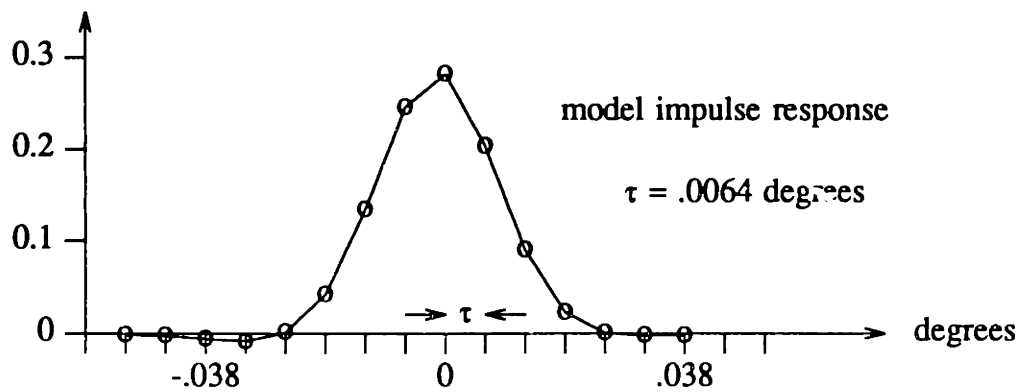


Figure 3-2: Eye filter

for the effects of printer distortions, but also to take advantage of them to produce more gray levels. To obtain good printer models, we first need to understand the different factors that affect the behavior of the printer.

The main function of a printer is to place ink dots in designated areas, usually on a Cartesian grid. We will refer to grid locations with ink as black dots and grid locations without ink as white dots. There are two types of printers which perform this function: write-black printers and write-white printers. The difference between the two is that write-black printers print dots of black ink against a white background, while write-white printers, effectively, “print” dots of white ink against a black background. In reality, both printer types use black ink for printing. The actual difference lies in the way the drum of the printer and the particles of carbon or ink are charged. The nature of the distortions in these two types of printers is quite different. There are many causes for these distortions. They include the spreading of the laser beam, interactions of the laser and the charge applied to the drum, the type of toner particles used, and the heat finishing. The printer distortions have a significant effect on the actual darkness of a printed image. The purpose of printer models is to be able to predict these distortions.

In our description of printer models, we will introduce the following notation. The printer is driven by a binary image  $[b_{i,j}]$ , where  $b_{i,j} = 1$  means that an ink dot is to be printed at pixel  $(i, j)$  and  $b_{i,j} = 0$  means that no ink dot is to be printed. As a result of printer distortions, the gray level  $p_{i,j}$  produced in the vicinity of  $(i, j)$  depends in some complicated way on  $b_{i,j}$  and the neighboring pixel values. The printer model takes the form

$$p_{i,j} = \mathcal{P}(W_{i,j}) \tag{3.2}$$

where  $W_{i,j}$  consists of  $b_{i,j}$  and the bits in its neighborhood. This is the most general form of the printer model proposed in [2]. Our task is to find the function  $\mathcal{P}(\cdot)$ . For the model-based techniques, it is essential that the window  $W_{i,j}$  is finite.

To obtain such a printer model, two approaches can be taken. The first approach relies on the mathematical formulation of the physical behavior of printers. The



second approach, which is the goal of this thesis, relies on measurements. In the following subsections, brief discussions of each of the two approaches will be presented.

### 3.2.1 Circular Dot-Overlap Model

One approach for modeling the behavior of printers is through a mathematical formulation of the physical phenomena that affect it. One model which was proposed in [1] is the “circular dot-overlap model.” In this model, each ink dot is assumed to be circular with uniform distribution of ink.

The radius of the dots produced by a printer must be at least  $T/\sqrt{2}$ , where  $T$  is the spacing of the Cartesian grid, so that a page can be blackened entirely [1]. We will refer to a printer that produces dots of minimal size as the “ideal” printer. Actual printers produce dots that are larger than the minimal size. We will use  $\rho$  to denote the ratio of the actual dot radius to the radius of the dot of the ideal printer. The effective gray level of a printed pixel is assumed to be the percentage of the area of the pixel that is covered by ink. If an appropriate value of  $\rho$  is chosen, then the effective gray level of all 2-D patterns can be predicted by calculating the area of each pixel that is covered by ink.

The amount of ink spreading at each pixel can be expressed in terms of the parameters  $\alpha$ ,  $\beta$ , and  $\gamma$ , shown in Figure 3-3. These parameters are the ratios of the areas of the shaded regions shown in the figure to  $T^2$ , the area of the pixel. They can easily be expressed in terms of the radius  $\rho$  [1]. In terms of these parameters, the circular dot-overlap model becomes

$$p_{i,j} = \mathcal{P}(W_{i,j}) = \begin{cases} 1, & \text{if } b_{i,j} = 1 \\ f_1\alpha + f_2\beta - f_3\gamma, & \text{if } b_{i,j} = 0 \end{cases} \quad (3.3)$$

where  $f_1$  is the number of horizontally and vertically neighboring dots that are black,  $f_2$  is the number of diagonally neighboring dots that are black and not adjacent to any horizontally or vertically neighboring black dot, and  $f_3$  is the number of pairs of neighboring black dots in which one is a horizontal neighbor and the other is a vertical neighbor. Since the parameter  $\alpha$  is the largest of the three parameters, we

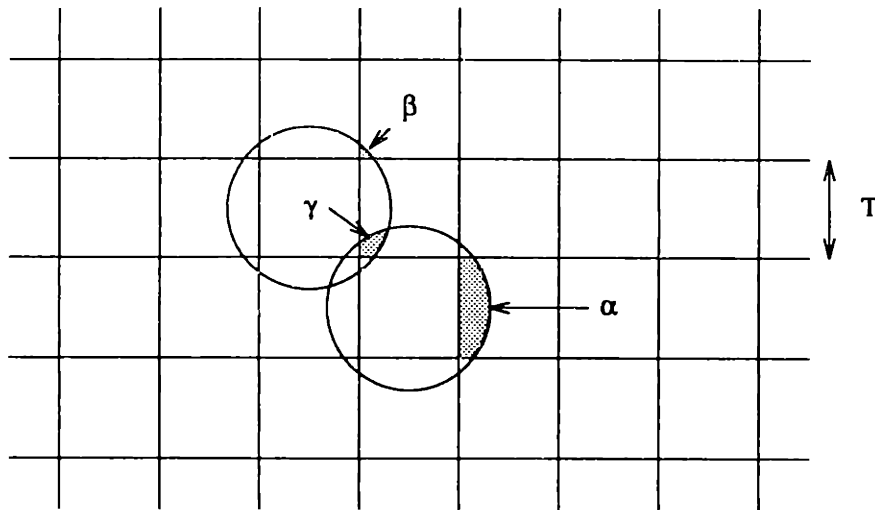


Figure 3-3: Definition of  $\alpha$ ,  $\beta$  and  $\gamma$ .

also refer to this model as  $\alpha$  dot overlap.

The circular dot-overlap model is an idealization of the physical behavior of printers. It provides a good first order approximation for the behavior of many printers. The main advantage of the circular dot-overlap model is its simplicity. However, it does not adequately explain all the printer distortions that occur. Often there are significant discrepancies between predictions of the model and the measured values. Plate 3 shows an example of some of these discrepancies. For each pattern we list the gray level that would result if no dot overlap were present (square dots), the gray level predicted by the circular dot-overlap model with  $\alpha = .33$ , and the gray level measured by our reflection densitometer.

More importantly, for some printers the circular dot-overlap model cannot explain the effects of printer distortions. In order to accurately predict the behavior of any printer, we need to consider the measurement approach.

### 3.2.2 Measurement Approach

An alternative approach for predicting printer behavior is by direct measurement of the printer parameters. The parameters of the printer model can be obtained directly from measurement of the gray level of various printed test patterns. As the

measuring device we used a reflection densitometer, the MACBETH RD922 (Answer II). The measurement approach makes minimal assumptions on the printing process and can apply to any write-black or write-white printer. The development of such a measurement approach is the main focus of this thesis, and will be discussed in detail in Chapter 4.

The measured printer model parameters can be used by various halftoning techniques to enhance their accuracy of gray-scale rendition. Two such model-based techniques are the *modified error diffusion* algorithm and the *least-squares model-based* method. We now consider these two techniques.

### 3.3 Modified Error Diffusion

The standard error diffusion algorithm, as mentioned earlier, is sensitive to the effects of ink spreading. By incorporating a printer model into error diffusion, the modified error diffusion algorithm [1], not only corrects for the effects of printer distortions, but also takes advantage of them to produce more gray levels.

In the modified error diffusion algorithm, the printer model is used to correctly calculate the error values. In the standard error diffusion algorithm, the errors are calculated under the assumption that the printed pixels are either perfectly black or perfectly white, and thus their gray value is equal to the binary value assigned to the output pixels. However, in actual printers this assumption does not hold. The gray value of what is printed is affected by dot overlap. The printer model is used to predict the gray values of the printed pixels. A block diagram of the modified error diffusion algorithm is shown in Figure 3-4. The notation is the same as that used for the standard error diffusion algorithm in the previous Chapter.

The modified error diffusion equations are [1]

$$v_{i,j} = y_{i,j} - \sum_{m,n} h_{m,n} e_{i-m,j-n}^{i,j} \quad (3.4)$$

$$b_{i,j} = \begin{cases} 1, & \text{if } v_{i,j} > t \\ 0, & \text{otherwise} \end{cases} \quad (3.5)$$

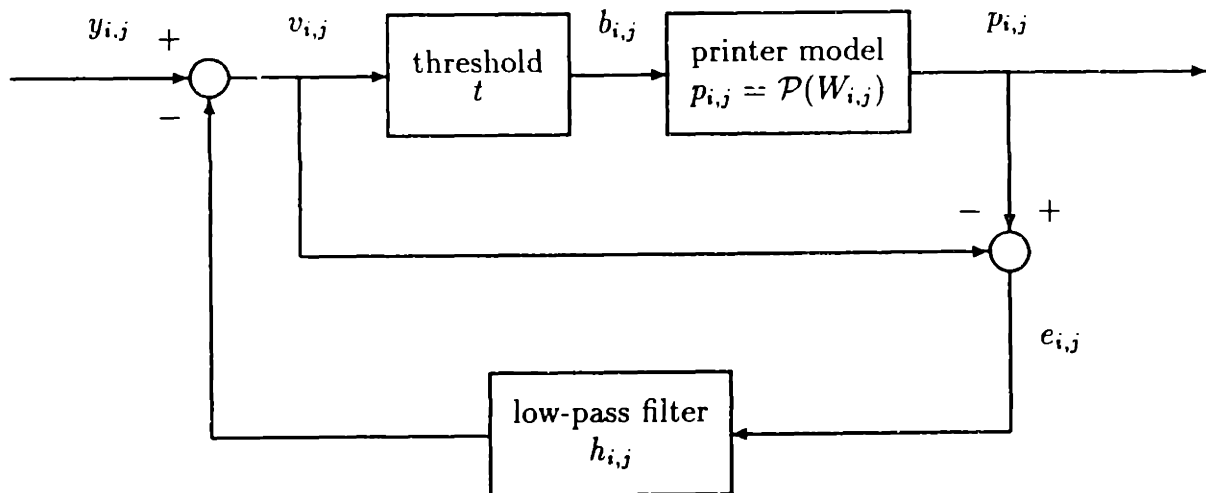


Figure 3-4: Error Diffusion With Ink-Spreading Compensation

$$e_{m,n}^{i,j} = p_{m,n}^{i,j} - v_{m,n} \quad \text{for } (m,n) \prec (i,j) \quad (3.6)$$

where  $(m,n) \prec (i,j)$  means  $(m,n)$  precedes  $(i,j)$  in the scanning order and

$$p_{m,n}^{i,j} = \mathcal{P}(W_{m,n}^{i,j}) \quad \text{for } (m,n) \prec (i,j) \quad (3.7)$$

where  $W_{m,n}^{i,j}$  consists of  $b_{m,n}$  and its neighbors as in Equation (3.2). Notice, however, that in this case the neighbors  $b_{k,l}$  have been determined only for  $(k,l) \prec (i,j)$ ; they are assumed to be zero for  $(k,l) \succeq (i,j)$ . Since only the dot-overlap contributions of the “past” pixels can be used in (3.7), the “past” errors keep getting updated as more binary values are computed. Hence the dependence of the error and the printer model output on the “instant”  $(i,j)$ .

Plate 2(b) shows the results of the modified error diffusion algorithm with the circular dot-overlap printer model applied to the test images. Recall that Plate 2 was printed on a write-black HP LASERJET II. Observe that the images are sharper than any of the screening techniques, have fewer low-frequency artifacts, and also have no apparent distortion of the gray-scale.

In the presence of dot overlap, the error diffusion algorithm is no longer causal. The

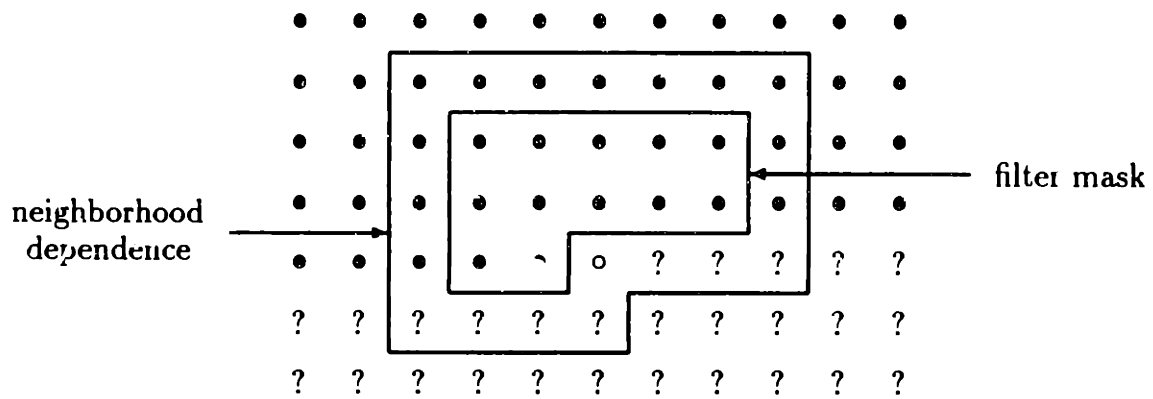


Figure 3-5: Error calculation in modified error diffusion

gray level at each printed pixel depends on its own pixel value plus the neighboring values. Figure 3-5 shows the binary values that are necessary for the calculation of the error for all the points in the error diffusion filter mask. For the points on the edge of the filter mask, the error depends on future points which have not yet been determined. As we will see in Chapter 5, this can cause detrimental instabilities in the algorithm if the wrong initial conditions are assumed.

### 3.4 Least-Squares Model-Based Method

The least-squares model-based (LSMB) method is an alternative halftoning technique that takes advantage of both an eye model and a printer model [2]. The least-squares method attempts to produce an “optimal” halftone image by minimizing the error between the original gray-scale image as seen through the eye and the printed image as seen through the eye. Thus the computation of the error between the two images requires modeling the behavior of the eye and the behavior of printer. This is illustrated in the block diagram of Figure 3-6.

Given a gray-scale image  $[y_{i,j}]$ , the LSMB method finds the binary image  $[b_{i,j}]$  that minimizes the sum of the squares of the differences between the two perceived

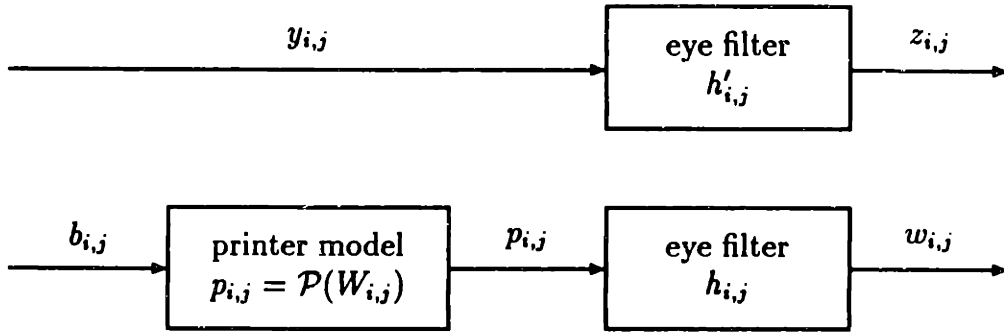


Figure 3-6: The Visual and Printer Perceptual Models.

images  $[z_{i,j}]$  and  $[w_{i,j}]$

$$E = \sum_{i=1}^{N_W} \sum_{j=1}^{N_H} (z_{i,j} - w_{i,j})^2 \quad (3.8)$$

where, as illustrated in Figure 3-6,

$$z_{i,j} = y_{i,j} * h'_{i,j} \quad (3.9)$$

$$w_{i,j} = p_{i,j} * h_{i,j} = \mathcal{P}(W_{i,j}) * h_{i,j} \quad (3.10)$$

and  $*$  indicates convolution. As in [2], we have allowed different impulse responses  $[h_{i,j}]$ ,  $[h'_{i,j}]$  for the eye filters corresponding to the continuous-tone and halftone images. In fact, in the examples of Chapter 5, we dropped the filter  $[h'_{i,j}]$  completely as this results in sharper halftone images [2].

In principle, the optimal solution can be obtained by generating all the possible combinations of “1”s and “0”s for the entire image and finding which combination of binary values minimizes the error. Unfortunately, this approach is not computationally feasible. For instance, let us consider an image with 512 x 512 pixels. If all the possible combinations of “1”s and “0”s were generated for this image, there would be  $2^{512 \times 512}$  possible binary images. This makes the search for the minimum error next to impossible.

Although the search through all the possible combinations of binary values is not

feasible, there are alternative ways to approach the problem. One approach is to update only one pixel at a time. For each image pixel, the error is computed for each of the two possible values of the binary pixel in question, and the algorithm picks the binary value which results in the lower error value. Since the assignment of a binary value at this point only affects the gray value of that point and the gray values of its surrounding neighbors, the minimum error calculation can be performed by simply re-computing the error values for the pixels that are affected. In fact, the best results are obtained when the error is computed only at the point being updated (see [2]).

Another strategy for implementing the least squares method is to increase the number of points updated at each particular instant. Instead of updating one point at a time, four or nine points can be updated at a time. However, the computational complexity increases exponentially with the number of points being updated.

The above minimization must be repeated for all pixels in the image, and the whole procedure should be repeated through several iterations until convergence. Thus, the least-squares method is iterative. The starting point can be an all white or all black image, or the result of some other halftoning algorithm, e.g. the modified error diffusion. The number of iterations depends on the starting point. Typically 5-10 iterations are required for convergence.

### 3.4.1 One-dimensional Least-squares

In the one-dimensional case, each row or column of the image is halftoned independently. One-dimensional halftoning is simpler than two-dimensional halftoning and easier to analyze, but is seldom used in practice because it does not exploit the two-dimensional properties of the eye. One-dimensional least-squares halftoning can be implemented, in closed form, with the Viterbi algorithm, which is a dynamic programming algorithm [5]. Unfortunately, the Viterbi algorithm cannot be used in two dimensions. Therefore, no closed form solution exists for the two-dimensional least-squares problem and, as we saw above, iterative techniques are necessary. In addition to their theoretical interest, the one-dimensional results are a useful guide to understanding the two-dimensional problem.

# Chapter 4

## Measurement of Printer Model Parameters

The model-based halftoning techniques outlined in Chapter 3 depend on an accurate representation of printer behavior. The characteristics of different printers vary considerably and have a significant effect on the darkness of the printed images. Thus, a model which accurately predicts the behavior of a printer is necessary for the effectiveness of the model-based halftoning techniques.

One printer model that is based on the physical understanding of printer behavior is the circular dot-overlap model discussed in Chapter 3. The circular dot-overlap model is a good model for predicting the effects of printer distortions for many printers, including one of our test vehicles, the HP LASERJET II. However, this model does not apply to all printers. Rather than trying to understand the behavior of each individual printer or class of printers, we develop an alternative approach which is based on direct measurement of printer characteristics and can be applied to any printer.

The main goal of this thesis is to develop a method for obtaining accurate estimates of the printer model parameters. This method is based on direct measurement of the reflectance of a set of printed test images. As was suggested in [6], the measured reflectance can be related to the printer model parameters by a set of linear equations. We then formulate a constrained optimization problem which incorporates various



constraints on the model parameters. This optimization problem can be solved by standard iterative techniques. We will refer to the resulting printer model as the “measurement model.” The approach we develop is general and makes very few assumptions about the laser printer.

In this chapter, we present the formulation of the problem. We first describe the approach for one-dimensional printer models, which are used in one-dimensional model-based halftoning. This is done to test the feasibility of the measurement approach. Next, we consider the two-dimensional problem. The two-dimensional measurement model is the most important in practice and will be the main focus of this work.

As our test vehicles, we used an HP LASERJET II, which is a write-black printer with 300 dpi resolution and a DATA PRODUCTS LZR 2665, which is a write-white printer. As a measuring device, we used a MACBETH RD922 (Answer II) Reflection Densitometer. This densitometer is designed to measure the average reflectance over an area whose diameter is approximately 4 mm.

The reason for choosing an instrument which measures the average reflectance of the test patterns, instead of using an instrument which measures the gray level of individual pixels, is that we do not have to estimate the transfer function of the instrument. Estimating the characteristics of the measuring device and the printing device at the same time would be very difficult, if at all possible. Also, measurement of average reflectance does not require precise alignment of the measuring device with the printed patterns. Finally, measuring the average reflectance over a small area of the image is in fact similar to what the human eye is doing. Thus, to some extent, our measuring device resembles the behavior of human perception.

## 4.1 One-dimensional Measurement Model

There are many reasons for studying the one-dimensional problem, even though it is of limited practical significance. First, the problem is simpler in one dimension in terms of the number of unknown parameters and measurements involved. It is

thus easier to test the various assumptions that we make and to look for potential problem areas before attempting to tackle the two-dimensional problem. Second, the model can be tested using the Viterbi algorithm to obtain the solution to the one-dimensional least-squares model-based halftoning problem. As we saw in the previous chapter, no closed form solution exists for the two-dimensional least-squares problem. We thus have to rely on approximate solutions to the least-squares problem and the modified error diffusion algorithm, in order to test the two-dimensional model.

The goal of a printer model is to predict the “gray” level of printed pixels. The gray level of each printed pixel depends in a complicated way on the pixels in its surrounding neighborhood. The advantage of the measurement approach is that we can measure the effect of the surrounding pixels directly, instead of trying to predict it based on the physical understanding of printer behavior. In the following section, we will describe the procedure for obtaining the one-dimensional measurement model.

#### 4.1.1 Formulation

One-dimensional images are assumed to be invariant in the horizontal (or vertical) direction. An one-dimensional binary image is specified as a sequence of bits.

. . . 0 1 0 1 1 . . .

The horizontally (or vertically) invariant image is obtained by repeating the pattern several times.

. . . 0 1 0 1 1 . . .  
. . . 0 1 0 1 1 . . .  
. . . 0 1 0 1 1 . . .

This sequence of bits is the data that is sent to the printer for printing. At each location of the image, the printer prints a dot of ink where a “1” is specified and prints no ink dot where a “0” is specified. The goal of the one-dimensional printer model is to predict the gray level at each printed pixel.

The gray value of each printed pixel is affected by the neighboring pixels. According to the one-dimensional version of the printer model of Equation (3.2), the “gray” level  $p_k$  produced by the pixel at site  $k$  is

$$p_k = \mathcal{P}(W_k) \quad (4.1)$$

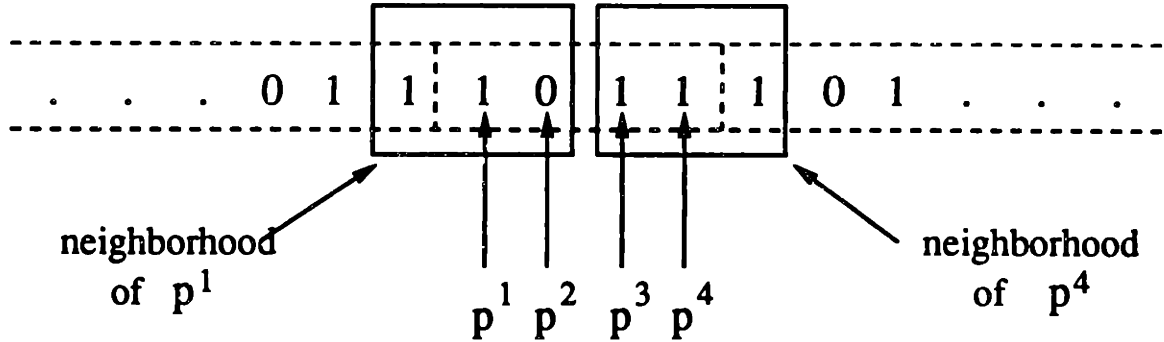
where  $W_k$  is a window consisting of  $b_k$  and the bits in its neighborhood. For the model-based halftoning techniques described in Chapter 3, it is essential that the window be finite. The first step in the procedure of obtaining a printer model is to select the window size. In the following discussion, for simplicity, we choose a printer model of window size 3.

Once the window size is selected, the mapping  $\mathcal{P}$  has to be specified for every possible pattern  $W_k$ . For a printer model with window size 3, there are 8 possible patterns for  $W_k$ :

$$\begin{aligned} W^1 &= 0\ 0\ 0 \\ W^2 &= 0\ 0\ 1 \\ W^3 &= 0\ 1\ 0 \\ W^4 &= 0\ 1\ 1 \\ W^5 &= 1\ 0\ 0 \\ W^6 &= 1\ 0\ 1 \\ W^7 &= 1\ 1\ 0 \\ W^8 &= 1\ 1\ 1 \end{aligned}$$

The mapping  $\mathcal{P}$  assigns a “gray” level  $p^i$  to each pattern  $W^i$ . In other words, it determines the effect of the neighboring pixels on the pixel at the center of the window. To determine the mapping  $\mathcal{P}$ , we must first establish relationships between the “gray” values  $p^i$  corresponding to the above set of patterns and the reflectance measurements of the test patterns. Then we can solve for the unknown “gray” values  $p^i$ .

To relate the measurement of the average reflectance to the unknowns, we consider the following strategy. First, we generate a set of test images that are printed



$$\text{Sample equation: } (p^1 + p^2 + p^3 + p^4)/4 = m^1$$

Figure 4-1: Sample periodic pattern with period "1011"

and measured. As an example, let us consider the test pattern that results when the pattern "1011" is repeated periodically as shown in Figure 4-1. The average reflectance of the printed test image is equal to the average reflectance of all of the individual pixels in the periodic pattern. Thus, an equation relating the unknowns  $p^i$  to the measurements  $m^j$  can be obtained. The equation corresponding to the sample pattern of Figure 4-1 is also shown in the figure.

Note that the periodicity of printed test patterns does not have to equal the size of the window of the printer model. In principle, the periodicity of the test patterns can take any value. In practice, however, we choose patterns with low periodicity. The main reason for avoiding patterns with high periodicity is that the reflection densitometer can only measure reflectance over a finite area. If the periodicity is too high, then relationships between the measurements and the unknowns cannot be established. In this example, the periodicity of the pattern is 4, and the window size of the unknowns is 3.

Based on the procedure outlined above, many different patterns of various periodicities can be generated and measured. This results in additional equations relating

the unknowns to the measurements.

$$\begin{aligned}
 (a^{1,1} * p^1 + a^{1,2} * p^2 + \dots + a^{1,I} * p^I) / d^1 &= m^1 \\
 (a^{2,1} * p^1 + a^{2,2} * p^2 + \dots + a^{2,I} * p^I) / d^2 &= m^2 \\
 &\vdots \\
 (a^{J,1} * p^1 + a^{J,2} * p^2 + \dots + a^{J,I} * p^I) / d^J &= m^J
 \end{aligned}$$

The above equations can be summarized in matrix form.

$$A * P = M$$

The number of equations  $J$  can exceed the number of unknowns  $I$ . Once the problem is placed in matrix form, we can use constrained optimization to solve for the unknown vector  $P$ . However, before we do that, we first seek ways to reduce the size of the problem. There are two methods that can be used to reduce the number of unknowns. Both assume some basic understanding of the printing mechanism. One is based on symmetry and the other is based on some limitations on the amount of ink spreading.

In the one-dimensional model with window size 3, there are originally 8 unknowns. First, we can eliminate patterns that are the same based on symmetry. For example, the pattern “001” is the mirror image of the pattern “100.” It is reasonable to assume that the “gray” level of a white dot surrounded by one black and one white dot does not depend on what side the dots are on. Thus, only one of these patterns must be included in the set of unknowns. We are in effect setting an equivalence relation between the pattern “001” and the pattern “100”. The only other patterns in the one-dimensional model with window size 3 that are mirror images of each other are “110” and “011.” The set of unique patterns after symmetry reduction is shown below.

0 0 0  
 0 0 1  
 0 1 0  
 1 0 1  
 1 1 0  
 1 1 1

We can further reduce the set of unique unknown patterns based on the fact that the dot overlap (black or white) is limited to the immediate neighbors. For instance, we can assume that a white dot surrounded by two white dots is white, and a black dot surrounded by two black dots is black. Thus, the “gray” levels assigned to the patterns “000” and “111” are 0 and 1, respectively. For most printers, this is a reasonable assumption. Thus, for the one-dimensional printer model with window size 3, only 4 unknowns must be determined. The set of unique patterns and their unknown “gray” level parameters is the following.

0 0 0  $\rightarrow$  0  
 1 0 0  $\rightarrow$   $p^1$   
 0 1 0  $\rightarrow$   $p^2$   
 1 1 0  $\rightarrow$   $p^3$   
 1 0 1  $\rightarrow$   $p^4$   
 1 1 1  $\rightarrow$  1

The reduction in the number of unknowns simplifies the problem considerably, especially for larger window sizes.

### 4.1.2 Window Size

In the previous section, we outlined the measurement approach for the one-dimensional printer model with window size 3. Equations relating the unknown parameters to the reflectance measurements were established. The next step is to use constrained optimization to solve for these unknown parameters to build our measurement model.

However, before solving for these unknown parameters, we must first determine which window size should be used.

For the one-dimensional printer model, we considered several window sizes. An obvious choice would be to use window size 3 since it involves a relatively small number of unknowns. However, one can argue that window size 3 can not capture all the effects of printer distortion in one dimension. Thus, larger window sizes, such as 5 and 7, were investigated.

For window size 3, there are only 6 unique patterns after symmetry reduction. For window size 5, after symmetry reduction, there are 20 unique patterns, and for window size 7 there are 72. As saw above, for window size 3, the total number of unknown parameters can be further reduced based on the properties of ink spreading. Thus, patterns such as “111” and “000” can be assumed to have gray values of 1 and 0, respectively. This reduces the total number of unknown parameters to 4. Similarly, for window size 5, patterns such as “x111x” and “x000x” can be assumed to have gray values of 1 and 0, respectively. Here, “x” is used to indicate that the pixel value at that location can be either a “1” or a “0”. For window size 7, patterns such as “xx111xx” and “xx000xx” can be assumed to have gray values of “1” and “0,” respectively. The total number of unknown parameters for window size 5 becomes 14, and window size 7 has 52 unknown parameters.

The procedures used to generate equations for window size 3 are the same as the procedures used to generate equations for window sizes as 5 and 7. However, more unknowns and equations are involved in the larger window sizes.

### **4.1.3 Selection of Equations**

For each window size we chose many different patterns of different periodicities to obtain a set of equations for the unknown parameters. For window size 3, we used patterns of periodicity 2, 3, and 4 to generate the set of equations. Patterns of periodicity 2, 3, 4, 5, 6, and 7 were used to generate equations for window size 5. Window size 7 equations were obtained using patterns of periodicity 2, 3, 4, 5, 6, 7, and 8.

Patterns of different periodicity can be generated through an exhaustive search through all the possible combinations of “1”s and “0”s in the given periodicity. To avoid redundancy, equations which are repeated are eliminated from the matrix. The elimination of redundant equations is performed by comparing the coefficients in each equation and dropping equations with the same coefficients. For window size 3, the total number of equations used is 6. For window size 5, the total number of equations used is 22. Window size 7 used 49 equations.

In the one-dimensional printer model, the performance of the model can be tested by using the Viterbi algorithm. This will be done in the next chapter.

## 4.2 Two-dimensional Measurement Model

The development of the measurement model in two dimensions is very similar to the one-dimensional case, in terms of generating the unknowns and selecting the test patterns. As we discussed, the goal of the printer model is predict the “gray” level of a printed pixel based on the binary value of the surrounding pixels. We now describe how this can be done using reflectance measurements of various test patterns.

### 4.2.1 Formulation

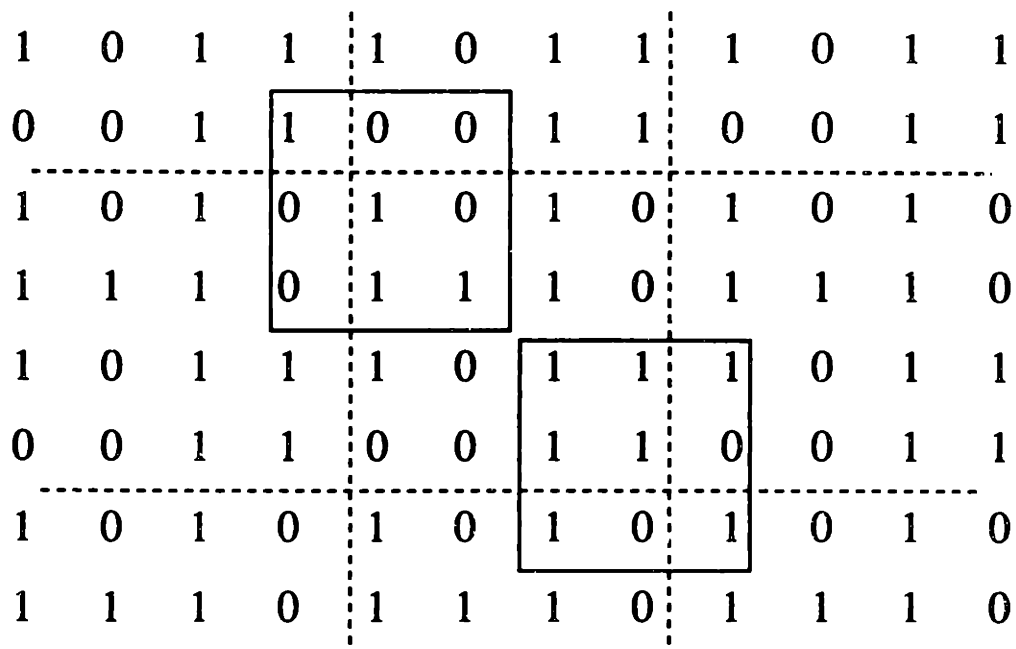
According to the two-dimensional printer model of Equation (4.2), the “gray” level  $p_{i,j}$  produced at site  $(i,j)$  is given by

$$p_{i,j} = \mathcal{P}(W_{i,j}) \quad (4.2)$$

where  $W_{i,j}$  is a window consisting of  $b_{i,j}$  and the bits in its neighborhood. The first step in the procedure for obtaining the measurement model is to determine the size and shape of the window. In the following discussion, we shall first consider a  $3 \times 3$  square window.

Once the window size and shape is selected, we must determine the function  $\mathcal{P}$  for all possible patterns  $W_{i,j}$ . Thus, given a pattern  $W^i$ , we must find the





Sample equation:  $(2 * p^1 + 4 * p^5 + 6 * p^3 + 4 * p^{24})/16 = m^1$

Figure 4-2: Sample periodic pattern and corresponding equation

corresponding “gray” level  $p^i$ . For example, one possible pattern is the following

1 0 1  
0 0 1  
1 0 0

For a  $3 \times 3$  window, the total number of such patterns that can be obtained is equal to 512.

To determine the mapping  $\mathcal{P}$ , we must relate the unknown “gray” values  $p^i$  to reflectance measurements of various test patterns. As in the one-dimensional case, we can find linear equations relating the reflectance measurements and the unknowns. Figure 4-2 shows an example of a test pattern. It is obtained by periodically repeating a  $4 \times 4$  pattern. The basic pattern is repeated both vertically and horizontally.

The measured reflectance of the printed test pattern is equal to the average reflectance of the 16 individual pixels contained in one period of the periodic pattern.

The reason this is true, as explained earlier, is that the densitometer measures the average reflectance over an area. Since each of the 16 pixels in the  $4 \times 4$  pattern corresponds to one of the unknown variables, equations relating the unknown parameters to measurements can be established.

Using the above procedure, additional patterns based of the same and different periodicities can be generated, printed, and measured to provide more equations. As in the one-dimensional case, the periodicity of test pattern does not have to be the same as the window size. In principle, the periodicity of the test patterns can be of any value. However, practical considerations limit the periodicity of test patterns to small values.

After equations relating the unknowns and the measurements are generated, we can then use constrained optimization to solve for the unknowns. However, before we do this, we again consider various possibilities for reducing the size of the problem. This will make the solution more computationally tractable.

In light of the large number of unknowns, several steps are taken to reduce the set of unknowns to a smaller set. First, the total number of unknowns can be reduced using symmetry. We shall assume here that patterns that are reflected or rotated are essentially the same. This is illustrated in Figure 4-3. Symmetry reduces the working set of 512 unknowns to 102 unique patterns. This reduction is performed by computer analysis.

The set of 102 unknowns can be further reduced to a smaller set by assuming that the "gray" level of a white dot surrounded by white dots is 0, and the "gray" level of a black dot surrounded by black dots is 1. Thus,

$$\begin{array}{ccc} 0 & 0 & 0 \\ 0 & 0 & 0 \rightarrow 0 \\ 0 & 0 & 0 \end{array}$$

$$\begin{array}{ccc} 1 & 1 & 1 \\ 1 & 1 & 1 \rightarrow 1 \\ 1 & 1 & 1 \end{array}$$

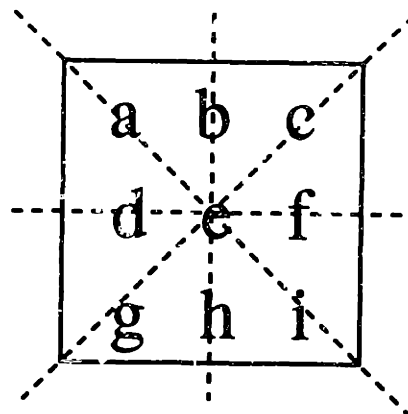


Figure 4-3: Symmetry reduction

This eliminates two more unknowns.

Further reduction in the working set can be performed based on some additional understanding of printer behavior. In write-black printers, black dots of ink tend to spread beyond the pixel boundaries. It is thus reasonable to assume that the “gray” value corresponding to all patterns with a “1” in the center is 1. That is, all patterns with the following form should be given the gray value of 1.

$$\begin{array}{ccc} x & x & x \\ x & 1 & x \\ x & x & x \end{array} \rightarrow 1$$

Conversely, in write-white printers, “white” dots tend to spread beyond the pixel boundaries. Thus, we can reasonably assume that the “gray” value corresponding to all patterns with a “0” on the center is 0.

$$\begin{array}{ccc} x & x & x \\ x & 0 & x \\ x & x & x \end{array} \rightarrow 0$$

This reduces the working set. The final working set contains 50 unknowns.

The assumptions made here are based on some understanding of the printing mechanism. These assumptions are necessary to reduce the size of our problem and to make the solution computationally tractable. However, by making these assumptions, we have also limited our scope to write-black and write-white printers only. This is one potential drawback in our formulation. Fortunately, most laser printers fall in one of these two categories. Also, for reasons which we will see later, these assumptions also help to maintain stability in the modified error diffusion algorithm.

#### 4.2.2 Window Shape and Size

In the previous section, we considered a printer model with a  $3 \times 3$  square window. The  $3 \times 3$  window is intended to capture the distortions caused by the pixels in the immediate neighborhood of each printed pixel. However, one might question if the  $3 \times 3$  window is enough to capture all the possible distortions. The answer to this is that the  $3 \times 3$  window does in fact capture most of the first order distortions. It neglects, however, certain second and third order effects. This leads us to investigate printer models with different window size and shape.

An alternative window shape that we consider is the cross-shaped window. Two cross-shaped windows are shown in Figure 4-4. The procedure for generating test patterns and equations for the cross-shaped windows is similar to the procedure outlined above for the  $3 \times 3$  square window. The only difference is in the number of unknowns and the number of equations involved. For the  $3 \times 3$  square window, there are 512 possible patterns. For the cross-shaped windows with lengths 3 and 5, there are 32 and 512 possible patterns, respectively.

In principle, the best results should be obtained with a printer model with a  $5 \times 5$  square window, shown in Figure 4-4. Unfortunately, the number of unknowns involved is too large. There is a total of 33,554,432 possible patterns for this window. This makes the problem intractable.

1	0	1
0	0	1
0	0	1

total patterns = 512

	0	
0	0	1
	1	

total patterns = 32

0	1	0	0	1
1	0	1	0	1
0	0	0	1	0
0	1	1	0	0
1	1	0	0	1

total patterns = 33,554,432

		1		
		0		
0	1	0	1	1
		1		
		0		

total patterns = 512

Figure 4-4: Different window sizes and shapes

### 4.2.3 Selection of Equations

For each window size and shape, we selected test patterns with different periodicities to generate equations relating the unknowns to the reflectance measurements. For the  $3 \times 3$  square window, we used patterns of periodicity  $2 \times 2$ ,  $2 \times 3$ ,  $3 \times 3$ , and a subset of the  $4 \times 4$  patterns. Again, we generated all the possible patterns for each periodicity, and eliminated redundant equations by comparing the coefficients of the equations. This resulted in a total of 200 patterns for the  $3 \times 3$  window. For the  $5 \times 5$  cross-shaped window, patterns of periodicity  $2 \times 2$ ,  $2 \times 3$ ,  $3 \times 3$ , and a subset of  $5 \times 5$  were used. The total number of equations generated was 244.

The equations can be placed in matrix form  $A * P = M$ . To solve for the solution vector  $P$ , we now introduce a constrained optimization technique.

### 4.3 Constrained Optimization

In the previous section we described how we can obtain a set of equations relating the unknown printer parameters to the reflectance measurements. These equations can be placed in matrix form  $A * P = M$ , where  $M$  is the vector of measurements and  $P$  is the unknown vector of “gray” values that correspond to the set of window patterns. Our goal is to solve for the vector  $P$ , which specifies the mapping  $\mathcal{P}$  which, in turn, specifies the printer model.

One approach to obtaining the solution vector  $P$  is to solve the linear system directly. Unfortunately, there are two problems associated with this approach. The first problem is the possible rank deficiency of matrix  $A$  and the second problem is discrepancies in measurements.

The first problem, where the matrix  $A$  is not full rank, occurs when not enough independent equations can be generated. Thus, we cannot solve for  $P$  directly. It is, in fact, difficult to obtain enough independent equations to guarantee that the matrix  $A$  has full rank. It is our speculation that this problem is caused by unknowns that occur in pairs. It is possible that, in the set of printed patterns, sets of unknowns always occur together. This causes matrix  $A$  to be rank deficient. For the  $3 \times 3$  square window, 3 of the singular values of the matrix are zero.

The second problem is that discrepancies in measurements exist. Again, here we cannot solve for  $P$  directly. One explanation for the discrepancies in measurements is the fact that the window size is too small to capture all the effects of printer distortions. For instance, for the  $3 \times 3$  square window, it is quite possible that ink dots which are two pixels away from a pixel affect its “gray” value. Thus, the limitations in the window size result in inconsistencies within the set of linear equations.

The first problem suggests that we need some additional constraints, and the second problem suggests that we need to find a best fit to the measurements. A reasonable set of constraints on the vector  $P$  is that its components  $p^i$ , which are gray-scale values, must be between 0 and 1. We can then minimize the error between the reflectance measurements of the test patterns and the reflectance predicted by the

parameter vector  $P$ . The minimization will be done in least-squares sense. Thus, we arrive at the following constrained optimization problem:

$$\begin{aligned} \text{minimize : } & \| A * P - M \|_2^2 \\ \text{subject to : } & 0 \leq p^1 \leq 1 \\ & 0 \leq p^2 \leq 1 \\ & \vdots \\ & 0 \leq p^n \leq 1 \end{aligned}$$

We now turn to the solution of the constrained optimization problem.

### 4.3.1 Solution Vector

The main issue in a constrained optimization problem is whether it has a unique solution and whether the algorithm we use can find it. Unfortunately, it is not always possible to find a unique solution.

To solve the constrained optimization problem, we used an optimization routine by Fletcher and Harwell from the Harwell Subroutine Library [13]. It is based on Davidson's method [12] which uses an approximation to the inverse Hessian matrix. The linear inequality constraints are dealt with by projection techniques. An initial estimate of the solution which satisfies the constraints must be provided. Given the starting point  $P'$ , an initial value of the error  $A * P' - B$  is calculated. At each iteration, a new point  $P''$  that satisfies the constraints and produces a lower error value is found. The algorithm continues for several iterations until it is unable to move toward a point with a lower error value. At this point, the algorithm stops and returns the current vector as the answer.

If there is a unique minimum, then the algorithm converges to that point. Unfortunately, often, there are several local minima in the solution space, and it is not always possible to determine the global minimum. However, if the matrix  $A$  is full

rank, then there exists a unique minimum in the solution space, and the constrained optimization routine returns the optimal solution [14, pp. 180–182]. When  $A$  is not full rank, there are two possibilities. If enough of the constraints become active, then a unique solution vector exists and the algorithm converges to that vector. Otherwise, there may be several local minima, and the starting point determines which solution vector is returned by the algorithm. Even though this can happen in practice, we have found that the performance of the printer model is not strongly affected by which solution is used to construct our printer model. This will become evident when we present our results in the next chapter.

As mentioned earlier, in many cases the matrix  $A$  is not full rank. For the one-dimensional printer models, the matrix  $A$  becomes full rank when the following assumption is made. If we assume that the gray value of all patterns with “1” at the center is 1, then the matrix  $A$  becomes full rank and a global minimum can be found. This assumption is valid for write-black printers only. For write-white printers, if we assume that the gray value of all patterns with “0” at the center is 0, then the matrix  $A$  becomes full rank, and the algorithm converges to the optimal value.

For the two-dimensional printer model with window size  $3 \times 3$ , the matrix  $A$  is not full rank, even when we make all the assumptions we made earlier to reduce the number of unknowns. For the cross-shaped window with length 3 the matrix  $A$  is full rank, while for the cross-shaped window with length 5 the matrix  $A$  is not full rank.

Using the techniques which we have outlined in this chapter, we are able to construct one-dimensional and two-dimensional printer models based on the measurement approach. The accuracy of these models will be tested in the next chapter when we apply them to the model-based halftoning techniques.



# Chapter 5

## Application

In the previous chapter, we described a procedure for estimating the printer model parameters based on measurements of printed test patterns. The resulting measurement model can be used to predict printer behavior for any write-black and write-white printer. In this chapter, we apply the procedure to obtain the parameters of the printer model and use the model-based halftoning techniques of Chapter 3 to evaluate its performance.

In our experiments, we used the HP LASERJET II write-black printer and the DATA PRODUCTS LZR 2665 write-white printer as our test vehicles. The characteristics of these two printers are very different. As our measurement device, we used the MACBETH RD922 (Answer II) Reflection Densitometer. Based on the procedures outlined in Chapter 4, we estimated the parameters of each printer model and then used them with the modified error diffusion and the least-squares model-based halftoning algorithms. A number of important issues arose in the process. The first issue is related to the reflectance measurements.

### 5.1 Reflectance Measurements

In the development of our measurement model, we assumed that the reflection densitometer measures the average reflectance of the printed test images. In our initial experiments, we estimated the printer model parameters based on these reflectance

measurements. The printer model was then used in model-based halftoning. The results of our experiments indicated that something was wrong. The images produced by the model-based techniques were too dark. This led us to question the reliability of the measurements. We suspected that the reflectance measurements need some type of calibration.

The only way to verify this hypothesis was to print a pattern whose reflectance is known and compare it to the reflectance measured by the densitometer. We decided to print a 32 level gray-scale ramp on the HP LaserJet. Since the HP LaserJet is a write-black printer with a fair amount of dot overlap, we tried to minimize the effects of dot overlap by using the classical screen of Figure 2-1(a) and printing at a resolution of 100 dpi. The 100 dpi resolution was simulated on the 300 dpi printer by repetition of ink dots. The combined effect of the lower resolution image and the use of the classical screen for halftoning minimizes the effects of ink spreading.

The reflectance of the 32 level ramp was then measured using the densitometer. If the instrument is calibrated correctly, then the measured reflectance values should be very close to the expected values. The expected values of the measurements lie along a straight line with a slope of  $-1$  and intersect at the y-axis and the x-axis at coordinates  $(0,1)$  and  $(1,0)$ , respectively. Figure 5-1 shows a comparison of the expected values versus the measured values. Observe that the values obtained from the densitometer are consistently lower. This explains why the model-based halftoning algorithms using the measurement model are producing images that are too dark. The inconsistency between the measured and expected reflectance values had to be resolved for the measurement model to be of use in model-based halftoning. This prompted us to investigate a phenomenon known as multiple internal reflection.

### **5.1.1 Multiple Internal Reflection**

Early studies of light reflection off highly scattering material done by Callahan [15] showed that the reflection of light is influenced by ink dots that are placed on top of the medium. Later studies by Clapper and Yule [16] showed that internal reflections of light within these materials account for a significant portion of the reflected light.

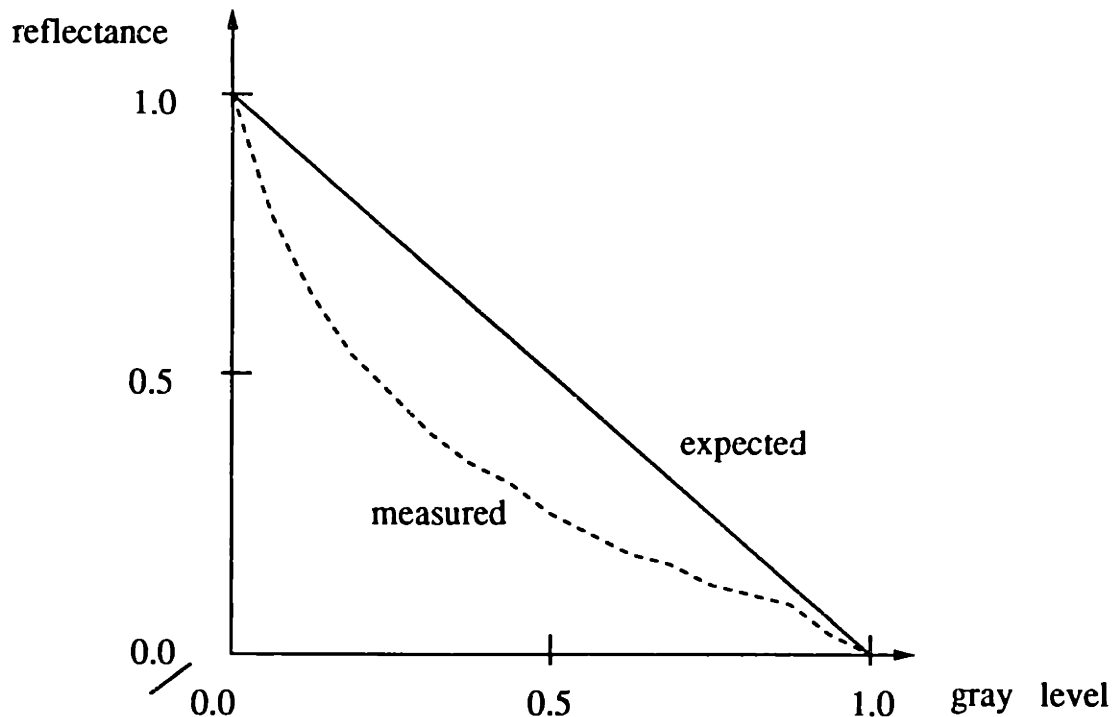


Figure 5-1: Reflectance calibration

The term *multiple internal reflection* is used to refer to this phenomenon.

Multiple internal reflection occurs when light multiply reflects within a medium such as paper. This phenomenon occurs when light strikes a medium at one point and emerges from the medium at several other points. Since light emerges at several points, ink dots that are placed on top of the medium influence the degree to which light is reflected. The phenomenon of multiple internal reflection is illustrated in Figure 5-2.

The degree to which multiple internal reflection occurs in our experiments depends on several factors. The first factor is the material of the paper. The second is the ink used for printing. The third is the intensity of the incident light. In the measurements of our test patterns, the third factor seems to play a large role in skewing our measurements.

The reflectance densitometer used in our experiments measures the ratio between the reflected light and the incident light. These measurements are obtained by shining a beam of light directly over the patterns and then measuring the amount of light

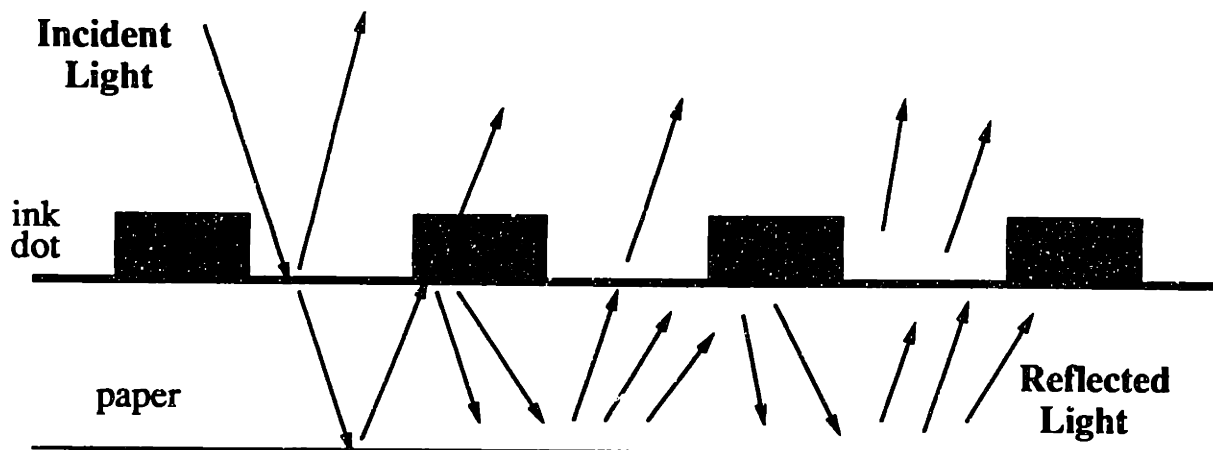


Figure 5-2: Multiple internal reflection

reflected. This beam of light is of high intensity and is placed approximately 1 mm away from the paper. In this setup, there is strong evidence which suggests that the high intensity and the close proximity of the light source magnify the effects of multiple internal reflection. These two combined effects cause more light to be reflected than is predicted. This explains the bias in the measurements.

### 5.1.2 Densitometer Calibration

To correct for the discrepancy between the measured and expected reflectance values, we used the measurements from the 32 level ramp to calibrate the reflectance densitometer. We used the graph shown in Figure 5-1 to correct our measurements. We then used the calibrated values to estimate the printer model parameters. The results of the model-based halftoning techniques indicate that the calibration removes the bias toward producing images that are too dark.

Thus, calibration of the reflectance measurements is an important step in the accurate measurement of printer model parameters. The accuracy of printer model parameters strongly affects the performance of model-based halftoning. In the following sections, we will examine the performance of our measurement models in model-based

halftoning.

## 5.2 One-Dimensional Least-Squares Model-Based Halftoning

In one-dimensional halftoning, each row or column of the image is halftoned independently. One-dimensional halftoning is simpler and easier to analyze, but is seldom used in practice because it does not exploit the two-dimensional properties of the eye. Nevertheless, the study of the one-dimensional techniques, which use the one-dimensional printer model, provides a good testing ground for our measurement approach.

As we saw in Chapter 3, least-squares model-based (LSMB) halftoning finds an “optimal” halftoned reproduction, by minimizing the squared error between the perceived intensity of the original gray-scale image and the perceived intensity of the printed halftoned image. One-dimensional LSMB halftoning can be implemented, in closed form, with the Viterbi algorithm. Since no closed form solution exists for the two-dimensional least-squares problem, the one-dimensional problem offers a unique opportunity to isolate the performance of the printer models from that of the optimization algorithm.

In testing one-dimensional printer models, i.e. the accuracy of gray-level rendition, the best test image to use is a gray-scale ramp. Plate 4 shows a gray-scale ramp halftoned using a two-dimensional classical screening technique, gray-scale ramps produced by the Viterbi algorithm using different printer models, and a gray scale ramp halftoned by a one-dimensional equivalent of the classical screen. The resolution of each gray-scale ramp is  $1200 \times 285$  pixels. Since the classical screening technique is fairly robust to printer distortion, the two-dimensional classical ramp is used as a reference for the performance of the different one-dimensional techniques.

The second ramp in Plate 4 was produced using the Viterbi algorithm and the circular dot overlap model with  $\rho = 1.25$ . The third, fourth, and fifth ramps were produced using the Viterbi algorithm and measurement models with window sizes 3,

5, and 7 respectively. The last ramp shown in Plate 4 was halftoned using a one-dimensional equivalent of the classical screen. The comparison with one-dimensional classical screening demonstrates the advantage of model-based techniques in terms of the number of gray levels produced. The other advantage of model-based techniques, namely that they produce sharper images, is not obvious here since we have a smooth image. In fact, it is the second advantage that is the most important, provided of course that the accuracy of gray-level rendition is at least preserved.

From Plate 4, we see that the gray scale of the ramp that was generated using the circular dot-overlap model is not monotonic. The problem appears near the top of the ramp.<sup>1</sup> This degradation in gray-level rendition is caused by the idealization of printer behavior. The gray-scale rendition of the ramp that was generated using the measurement model with window size 3 is considerably better than that produced using the circular dot-overlap model. However, it is still not perfect. There is still a minor problem with the monotonicity of the tone scale.

In the gray scale ramps produced using the models with window sizes 5 and 7, we see that there is very little improvement over the gray-scale ramp produced using the window of size 3. This indicates that a measurement model with window size 3 is enough to capture most of the effects of printer distortions in one dimension. Thus, there is no need to use larger window sizes to construct our measurement models. This result can also be used to argue that the  $3 \times 3$  window size is sufficient for constructing the two-dimensional measurement models.

The most important conclusion from these comparisons is that the measurement approach for estimating printer parameters works. The rendition of gray scale is fairly accurate and LSMB halftoning with our measurement model produces more gray levels than classical screening. We are now ready to test the two-dimensional measurement model using the modified error diffusion algorithm and the least-squares method.

---

<sup>1</sup>The reader is cautioned that the Mach band effect may make the halftoned ramp look more nonmonotonic than it is.

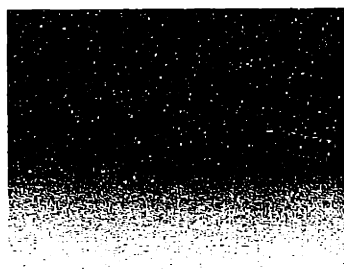


Figure 5-3: Mountain Effect

### 5.3 Modified Error Diffusion

In chapter 3, we outlined the modified error diffusion algorithm. It uses a printer model to estimate the gray level of the printed pixels. An accurate printer model is essential in the calculation of the error values. However, as we pointed out earlier, there is a causality problem when using either the measurement model or the circular dot-overlap model. Since this algorithm is not causal (as opposed to the standard error diffusion algorithm which is causal), steps need to be taken to maintain stability and prevent any corruption in the halftoned images.

For a write-black printer, if the assumption that all printed “1”s take on the gray level of 1 is relaxed, then corruption in image will occur. Similarly, for a write-white printer, if the assumption that all printed “0”s take on the gray value of 0 is relaxed, then corruption in the output image will also occur. An illustration of the type of corruption that results is shown in Figure 5-3. The image shown in this figure is the output image of a gray scale ramp processed using the modified error diffusion algorithm with a measurement model that relaxes the assumptions mentioned above. The corruption in image occurs at the top part of the ramp where a big portion of the image is black. We shall refer to this effect as the “mountain effect.”

The occurrence of the mountain effect is caused by instability of the error filter. To prevent the occurrence of the mountain effect, we propose two methods. The first is to make the constraints tighter. The second method is to use multi-pass error

diffusion. These two methods are described in the following sections.

### 5.3.1 Causality Constraints

The first method for preventing the occurrence of the mountain effect is to make stricter assumptions in the measurement model. To maintain stability, we need to assume that all “1”s have a gray value of 1 for write-black printers and all “0”s have a gray value of 0 for write-white printers. These assumptions were originally made when building the printer model with the  $3 \times 3$  window size, in order to reduce the number of unknowns and the degrees of freedom. We see here that these assumptions can also be used to guarantee stability.

In addition to these assumptions, the initial background of the output image must also be fixed to specified values. For write-black printers, we found that the initial background should be white. For write-white printers, the initial background should be black. The reason that the initial background affects the stability of the algorithm is that when the printer model attempts to calculate the error  $e_{i,j}$ , it needs to know the pixel value of some of the pixels that have not yet been determined, as was shown in Figure 3-5. If we assume the proper background before the initial processing of the image, then we can prevent any instability of the algorithm.

The two assumptions mentioned above prevent the occurrence of the mountain effect. This is the reason that no instabilities were observed in [1]. However, it is evident that the initial background assumption causes some error in the algorithm, since assuming the background to be either black or white will bias the algorithm in one direction. Fortunately, this error is very small.

The modified error diffusion algorithm was tested on several images. The test images include the “Lena” and “Building” images shown in Plate 5, and the gray-scale ramp in Plate 10. As we saw in Chapter 2, the resolution of the first two images is  $460 \times 460$  pixels. We upsampled by a factor of 2 to obtain  $920 \times 920$  pixels that result in the  $3.07 \times 3.07$  inch printed images. As we mentioned in Chapter 2, we used an interpolation scheme consisting of an *expander* [8, pp. 105–109] and an equiripple low-pass FIR filter [8, pp. 462–480]. The resolution of the gray-scale ramp



is  $1200 \times 240$  pixels.

Plate 5 shows the results of the modified error diffusion algorithm with two  $3 \times 3$  measurement models, obtained based on two different starting points. Plate 5 was printed on the write-black HP LASERJET II. Plate 10 was also printed on the write-black HP LASERJET II and compares the performance of different printer models and model-based techniques.

The test images in Plates 5 and 10 show that the performance of the modified error diffusion using our measurement model is quite good. The rendition of gray scale is relatively accurate, and the sharpness of the images is maintained. However, one problem with the quality of these images is the occurrence of some low-frequency artifacts. These low-frequency artifacts are especially apparent in the ramp of Plate 10 (third from the left). The funny patterns close to the top of the ramp are not very appealing. These low-frequency artifacts can also be seen on the arch in the “Lena” picture in Plate 5.

Plate 5 also demonstrates that the performance of the measurement model is not strongly influenced by which starting point is used to obtain the solution vector. In Section 4.3 we found that several solution vectors  $P$  may exist, and which one is chosen by the algorithm depends on the starting point. The images in Plate 5 show that the difference between the models obtained using two different starting points is small.

Plate 9 was printed on the write-white DATA PRODUCTS LZR 2665 printer. It shows the results of the modified error diffusion algorithm with a  $3 \times 3$  measurement model and compares them to classical screening. Again, the rendition of gray scale is relatively accurate, and the performance of error diffusion (sharpness and few low-frequency artifacts) is maintained. We have thus demonstrated that our procedure for estimating printer model parameters works on two different printers. A comparison of Plates 2(b) and 5 shows that the performance of the circular dot-overlap model and the measurement model is comparable on the write-black printer. However, the circular dot-overlap model does not apply to the write-white printer and, thus, we have to use the measurement model.

### 5.3.2 Multi-pass Error Diffusion

An alternative approach to solving the stability problem is to consider a multi-pass modified error diffusion algorithm. The original algorithm requires one-pass through the data. We suggest a multi-pass version which uses the image obtained from the previous pass as the background of the next iteration. More specifically, this procedure can be illustrated as follows.

During the first pass through the image, we assume the background to be completely white. The original image is processed using modified error diffusion based on this assumption. During the second iteration, the binary image obtained from the first pass is used as the background. Similarly, on the third pass, the image produced on the second pass is used as the background. The algorithm converges after a few iterations.

The multi-pass error diffusion algorithm offers several advantages. The first advantage is that it eliminates the instability problem. The second advantage is that the multi-pass algorithm also eliminates some of the low-frequency artifacts that the one-pass algorithm produces. This can be seen by comparing the third and fifth image from the left in Plate 10. Note that the multi-pass algorithm makes a difference only in the presence of dot overlap or other printer distortions. Also, a careful examination of Plates 6 and 10, shows that the multi-pass algorithm produces images that are slightly lighter than those produced by the one-pass algorithm. The difference is very small and appears to be insignificant. However, when the multi-pass error diffusion result is used as a starting point for the least-squares approach, then the resulting error is much lower than that obtained when the one-pass result is used as a starting point [2]. More importantly, the image with the lower error preserves the (visually pleasant) error diffusion texture, while the image with higher error is very grainy. This is because the one-pass error diffusion produces images that are darker than the original continuous-tone image. As the iterative least-squares algorithm tries to modify the binary image to restore the correct brightness, it destroys the error diffusion patterns [2]. Finally, we should point out that the images produced by the multi-pass error diffusion appear to be slightly grainier than those produced by the

one-pass algorithm.

## 5.4 Least-Squares Model-Based Halftoning

A different approach to model-based halftoning is the least-squares method. This method takes advantage of both an eye model and a printer model in an attempt to produce an “optimal” halftoned image. The goal in this approach is find the binary array that causes the perceptual printing model to produce an output that is as close as possible to the response of the visual perception to the original image.

As we discussed in Chapter 3, only approximate solutions to the least-squares problem can be found. Thus, the images produced using this method might only be local optima of the least-squares problem. There could in fact be several local minima for a given image. Here we demonstrate the approach that updates only one pixel at a time, using the modified error diffusion result as a starting point. In our experiments, we found that this version gives the best results. Examples using this technique are shown in Plate 8(b). In these examples, the initial background is the output image from 5 iterations of multi-pass error diffusion, shown in Plate 8(a). The number of iterations for the least-squares method is 5.

From the sample outputs, we see that the images produced using the least-squares method are sharper than those of the modified error diffusion algorithm. Also, the least-squares approach maintains the texture of error diffusion, which is known to be visually pleasant [7]. This happens only when the multi-pass error diffusion result is used as the starting point. Using an all-white image as the starting point results in images that are very grainy. In fact, as we discussed in the previous section, even when the one-pass error diffusion result is used as a starting point, the least-squares iterations result in grainy images.

More importantly, the gray-scale rendition of the halftoned images is good. We have thus demonstrated that when our measurement model is used with the model-based techniques of Chapter 3, it produces the correct gray scale and also maintains the overall performance of these techniques.

# Chapter 6

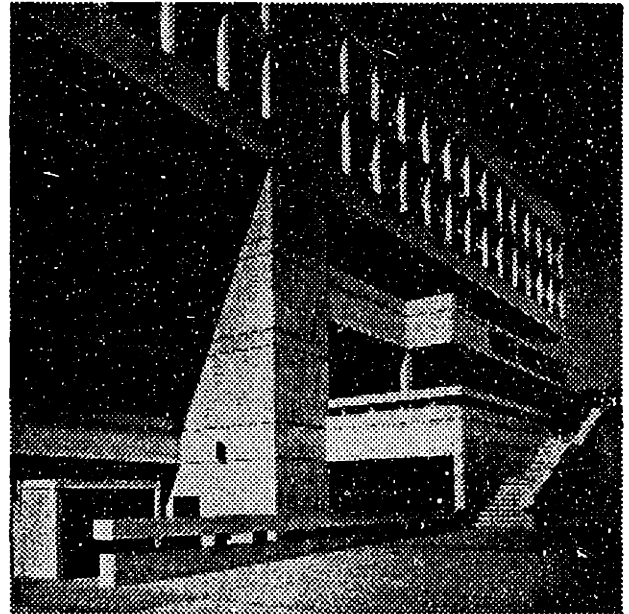
## Conclusions

The main focus of this work has been the development of an approach for estimating printer model parameters based on direct measurements of the reflectance of test patterns. The printer models predict the effects of printer distortions. They are an essential part of model-based halftoning techniques depend on accurate printer models to produce high quality images using standard laser printers.

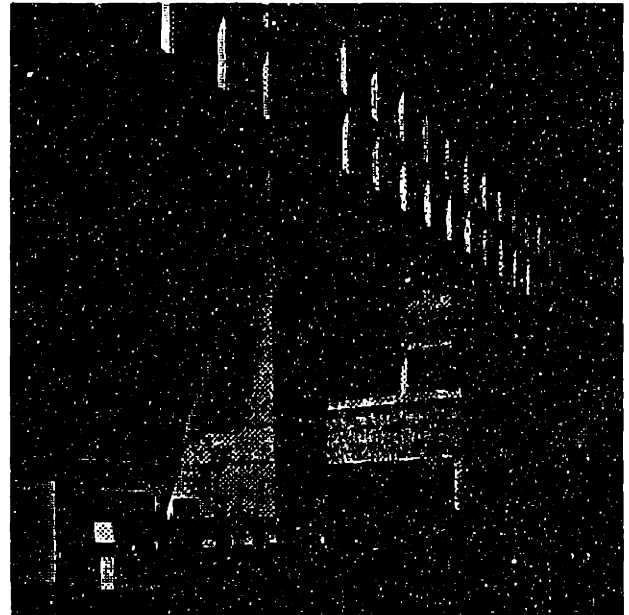
We first considered the problem of estimating the parameters of a one-dimensional printer model. Even though it is of limited practical significance, the one-dimensional problem is simpler, and the model can be tested using the Viterbi algorithm to obtain the solution to the one-dimensional least-squares model-based halftoning problem. Our results indicate that the measurement model performs better than the “circular dot-overlap model,” which assumes that the printer produces circularly shaped dots of ink.

The performance of two-dimensional printer models was tested using two model-based halftoning techniques, the modified error diffusion and the least-squares method. Both of these halftoning techniques use printer models to correct for the effects of printer distortions. Our results using a write-black printer indicate that the performance of the measurement model and the circular dot-overlap model is comparable. The measurement model was the only one that could be applied to a write-white printer, however. We thus demonstrated the necessity of a general approach that can be applied to a wide variety of laser printers.

We also derived an iterative (multi-pass) version of the modified error diffusion algorithm that improves its performance.

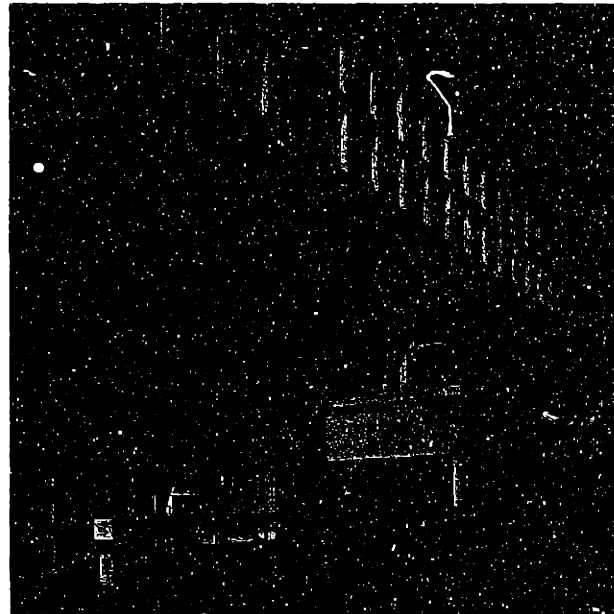


(a) Classical

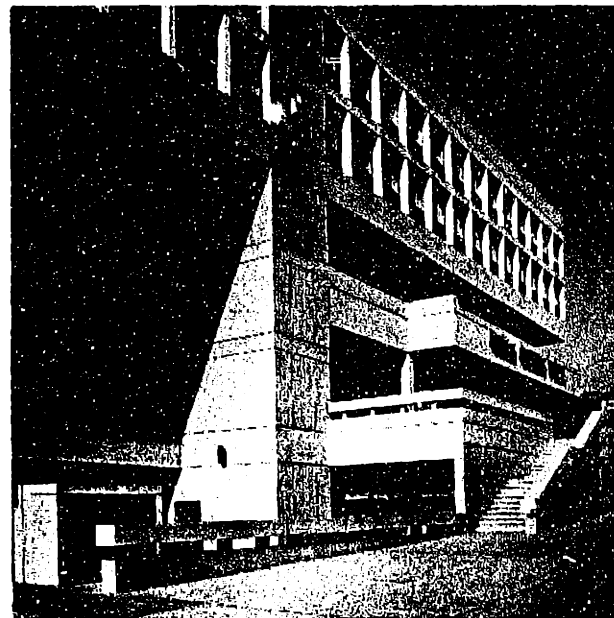


(b) Bayer

Plate 1: Dispersed vs. clustered ordered dither



(a) Standard error diffusion



(b) Modified error diffusion with circular dot-overlap model



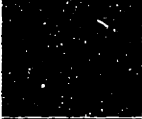
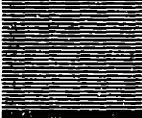

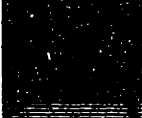






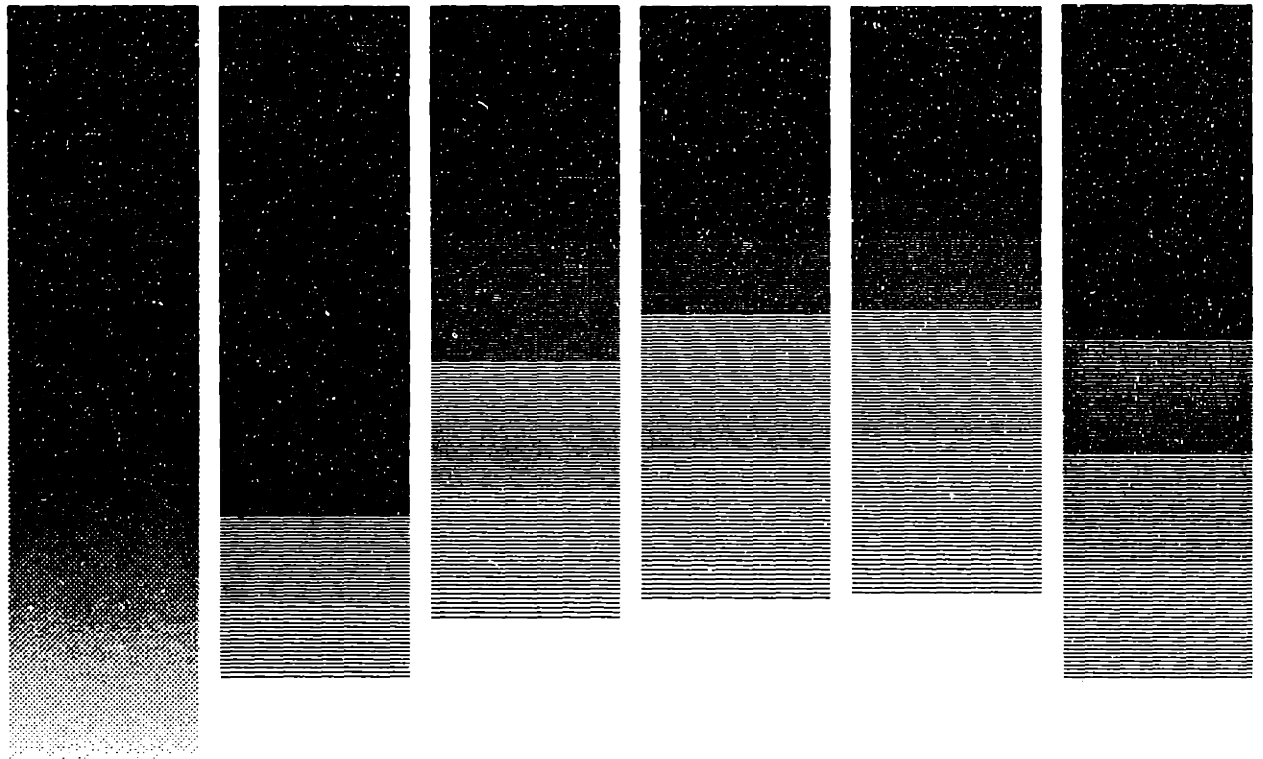
Pattern		Frequency of 1's	Darkness predicted by dot-overlap model window 3 (a = 0.33)	Measured darkness
100000		.17	.28	.22
100100		.33	.55	.60
101000		.33	.55	.40
110000		.33	.44	.30
101010		.5	.83	.91
101100		.5	.72	.56
111000		.5	.61	.36
110110		.67	.89	.87
101110		.67	.89	.86
111100		.67	.78	.47
111110		.83	.94	.81
111111		1.0	1.0	1.0

Plate 3: 300 dots/inch printing





Classic

Circular  
dot-  
overlap  
model

Window 3

Window 5

Window 7

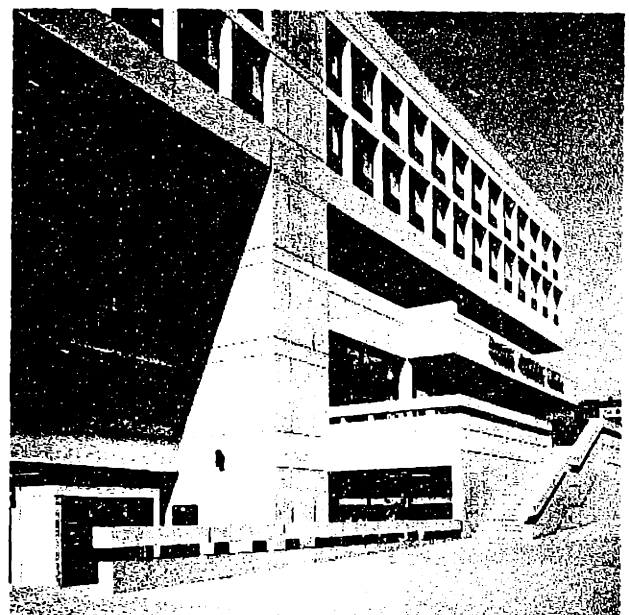
Measurement model

1-D  
clustered  
ordered

Plate 4: One-dimensional least-squares model-based halftoning



(a) Measurement model - initial point A



(b) Measurement model - initial point B



(a) 1 iteration (CDOM)



(b) 1 iteration (MM)



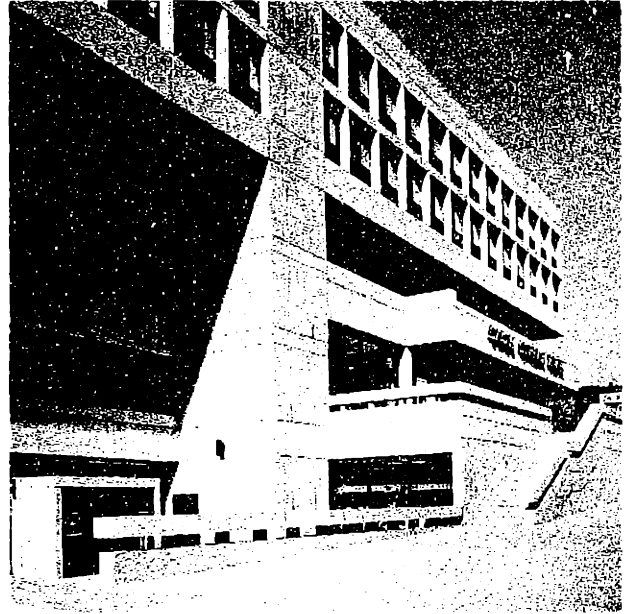
(c) 5 iterations (CDOM)



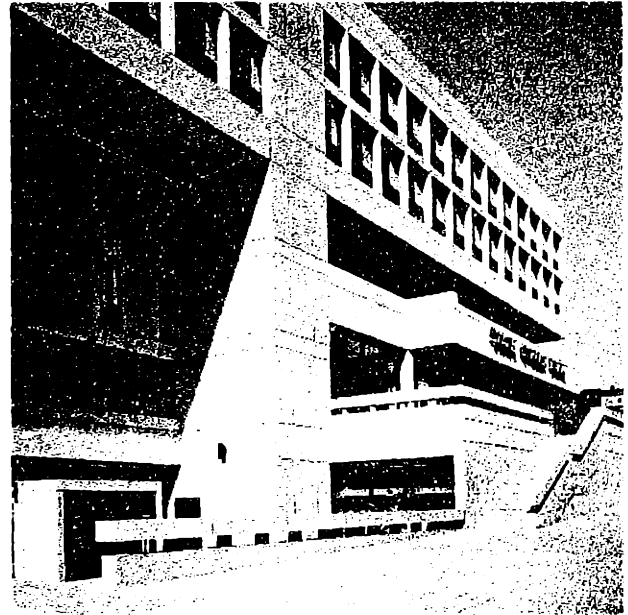
(d) 5 iterations (MM)

**Plate 6: Modified error diffusion**

**Circular dot-overlap model (CDOM) vs. Measurement model (MM)**



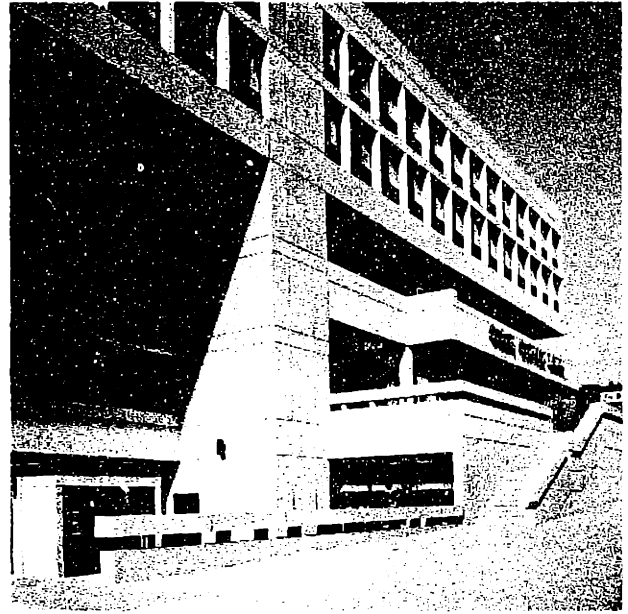
1 iteration



5 iterations

Plate 7: Modified error diffusion

Cross-shaped window model



(a) Error diffusion



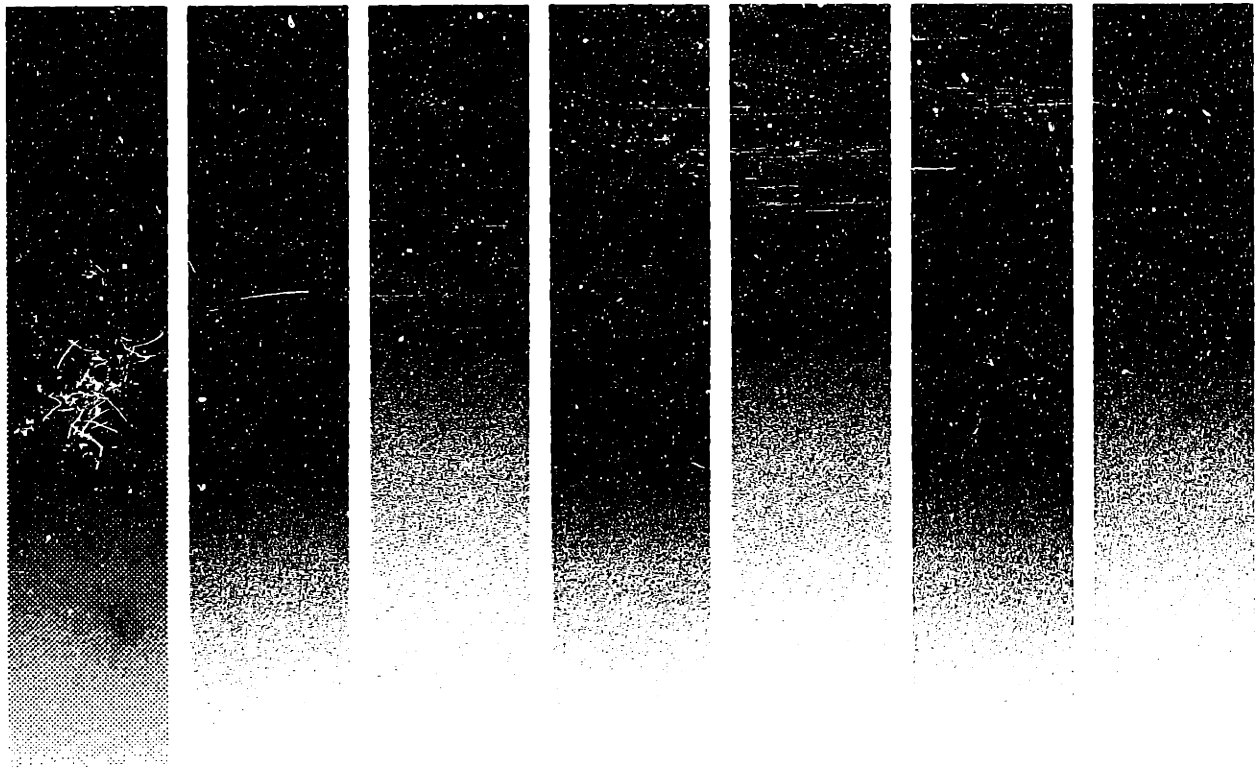
(b) Least-squares model-based



(a) Classical



(b) Modified Error Diffusion



Classic

CDOM

MM

Modified  
error diffusion  
1 iteration

CDOM

MM

Modified  
error diffusion  
5 iterations

CDOM

MM

Least-squares  
model-based  
5 iterations

Plate 10: Circular dot-overlap model (CDOM)  
vs. Measurement model (MM)

# Bibliography

- [1] T.N. Pappas and D.L. Neuhoff, "Model-Based Halftoning," *Proc. SPIE, vol. 1453, Human Vision, Visual Proc., and Digital Display II*, San Jose, CA, Feb. 1991.
- [2] T.N. Pappas and D.L. Neuhoff, "Least-Squares Model-Based Halftoning," *Proc. SPIE, vol. 1666, Human Vision, Visual Proc., and Digital Display III*, San Jose, CA, Feb. 1992.
- [3] D.L. Neuhoff and T.N. Pappas, "Perceptual Coding of Images for Halftone Display," *ICASSP-91*, Toronto, Ontario, Canada, May 1991.
- [4] J.L. Mannos and D.J. Sakrison, "The Effects of a Visual Fidelity Criterion on the Encoding of Images," *IEEE Trans. Inf. Theory*, vol. IT-20, no. 4, pp. 525-536, July 1974.
- [5] D.L. Neuhoff, T.N. Pappas, and N. Seshadri, "One-Dimensional Least-Squares Model-Based Halftoning," *ICASSP-92*, San Francisco, CA, March 1992.
- [6] D.L. Neuhoff, T.N. Pappas, and N. Seshadri, "One-Dimensional Least-Squares Model-Based Halftoning," to appear.
- [7] R. Ulichney, *Digital Halftoning*. The MIT Press, 1987.
- [8] A.V. Oppenheim and R.W. Schaffer, *Discrete-Time Signal Processing*. Prentice Hall, 1989.



- [9] P.G. Roetling and T.M. Holladay, "Tone Reproduction and Screen Design for Pictorial Electrographic Printing," *Journal of Applied Phot. Eng.*, vol. 15, no. 4, pp. 179–182, 1979.
- [10] R.W. Floyd and L. Steinberg, "An Adaptive Algorithm for Spatial Grey Scale," *Proc. SID*, vol. 17/2, pp. 75–77, 1976.
- [11] J.F. Jarvis, C.N. Judice, and W.H. Ninke, "A Survey of Techniques for the Display of Continuous-tone Pictures on Bilevel Displays," *Comp. Graphics and Image Proc.*, vol. 5, pp. 13–40, 1976.
- [12] D. Bertsekas, *Constrained Optimization and Lagrange Multiplier Methods*. Academic Press, 1982.
- [13] M.J. Hopper, ed., *Harwell Subroutine Library: A Catalog of Subroutines*, 55, Oxfordshire, England: A.E.R.E. Harwell, 1979.
- [14] P.E. Gill, W. Murray, and M.H. Wright, *Practical Optimization*. Academic Press, 1981.
- [15] P. Callahan, "Light Scattering in Halftone Prints," *Journal of the Optical Society of America* vol. 42, no. 2, pp. 104-105, Feb. 1952.
- [16] F. Clapper and J. Yule, "The Effect of Multiple Internal Reflections on the Densities of Halftone Prints on Paper," *Journal of the Optical Society of America* vol. 43, no. 7, pp. 600-603, Jul. 1953.