REAL-TIME NONLINEAR OPTICAL INFORMATION PROCESSING

by

Ahmad Armand

June 1979

Department of Electrical Engineering
Image Processing Institute
University of Southern California
Los Angeles, California 90007

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER USCIPI Report 880	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* REAL-TIME NONLINEAR OPTICAL INFORMATION PROCESSING		5. TYPE OF REPORT & PERIOD COVERED Technical Report June 1979
		6. PERFORMING ORG. REPORT NUMBER USCIPI Report 880
7. AUTHOR(s) Ahmad Armand		8. CONTRACT OR GRANT NUMBER(s) AFOSR-77-3285
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of Electrical Engineering Image Processing Inst., Univ. of So. Cal. University Park, L.A., Calif. 90007		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research Bldg. 410, Bolling AFB Washington, D.C. 20332		12. REPORT DATE June 1979
		13. NUMBER OF PAGES 137
14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)* AS ABOVE		15. SECURITY CLASS. *(of this report)* UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*

Approved for release: distribution unlimited

The United States Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation hereon.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Optical information processing, Nonlinear optical processing, Real-time optical processing, Optical computing

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*

Existing methods for nonlinear optical processing involve an intermediate photographic step which prevents real-time operation. Although optical information processors have a large space-bandwidth product, the problem of real-time input has been a major limitation in taking advantage of this processing ability. In this research, the realization of nonlinear optical processing in real time is investigated. Several techniques for performing such

operations are presented. One approach realizes the halftone method of nonlinear optical processing in real time by replacing the conventional photographic recording medium with a real-time image transducer. In the second approach halftoning is eliminated and the real-time device is used directly. In this case the nonlinearity is obtained by the inherent nonlinear characteristic of the real-time device. A third method uses the ability of certain real-time devices to perform an intensity-to-spatial frequency conversion.

Following a review of halftoning and real-time devices, a general analysis of the halftone process considering the nonideal characteristic of the recording medium is presented. From this analysis one can predict the amount of degradation of the output due to different parts of the recording medium characteristic curve for any nonlinear transformation. Specific results for logarithmic and level slice processes are obtained.

The problem of designing the halftone screen cell shape is considered for a piecewise linear recording medium by solving a nonlinear integral equation. It is shown that the solution can be achieved for certain monotonic functions including exponential and power transformations. To obtain the solution in general, an approximate method which considers a discrete halftone screen density profile is described. This gives the halftone screen density profile for any form of recording medium characteristic curve and any type of nonlinearity by minimizing in mean-square sense the difference between desired and degraded outputs. The results of computer simulation for logarithmic and level slice functions are shown. Experimental results are obtained for an optimized halftone screen which matches the characteristic of a liquid crystal light valve (LCLV). An overall logarithmic transfer function is produced in real time.

The procedure for obtaining nonlinearities without halftone preprocessing is explained. Experimental results using a special birefringent liquid crystal device which produces a 3-bit parallel A/D converter in real time are shown.

The variable grating mode (VGM) of liquid crystal light valve systems and its application in real-time nonlinear optical processing is discussed. A detailed analysis of the variable phase grating is made and from this the processing limitations of the VGM are determined. Experimental results showing real-time level-slicing with a VGM liquid crystal device are given.

Possible extensions and limitations of this work are discussed.

ACKNOWLEDGEMENTS

The author wishes to thank the members of his dissertation committee, Dr. A.A. Sawchuk, Dr. T.C. Strand, and Dr. Z. Vorel. The advice and guidance of Dr. Sawchuk, the committee chairman and director of the Image Processing Institute, is most sincerely acknowledged.

Special thanks goes to Dr. Strand for his many helpful suggestions and interest in this work.

The assistance of Mr. B.H. Soffer and Mr. D. Boswell of the Hughes Research Laboratories in performing some of the experiments is deeply appreciated.

A number of people have contributed to the physical preparation of this dissertation and it is a pleasure to acknowledge them. Mr. Ray Schmidt of the USC Image Processing Institute and his staff provided photographic services, and Mr. Doyle Howland of USC was responsible for drafting and artwork. The expert typing was done by Ms. Amy Yiu of USC.

Finally a personal note of gratitude goes to my wife for her patience and constant encouragement during the

course of this research.

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

Existing methods for nonlinear optical processing involve an intermediate photographic step which prevents real-time operation. Although optical information processors have a large space-bandwidth product, the problem of real-time input has been a major limitation in taking advantage of this processing ability. In this research, the realization of nonlinear optical processing in real time is investigated. Several techniques for performing such operations are presented. One approach realizes the halftone method of nonlinear optical processing in real time by replacing the conventional photographic recording medium with a real-time image transducer. In the second approach halftoning is eliminated and the real-time device is used directly. In this case the nonlinearity is obtained by the inherent nonlinear characteristic of the real-time device. A third method uses the ability of certain real-time devices to perform an intensity-to-spatial frequency conversion.

Following a review of halftoning and real-time devices, a general analysis of the halftone process considering the nonideal characteristic of the recording

medium is presented. From this analysis one can predict the amount of degradation of the output due to different parts of the recording medium characteristic curve for any nonlinear transformation. Specific results for logarithmic and level slice processes are obtained.

The problem of designing the halftone screen cell shape is considered for a piecewise linear recording medium by solving a nonlinear integral equation. It is shown that the solution can be achieved for certain monotonic functions including exponential and power transformations. To obtain the solution in general, an approximate method which considers a discrete halftone screen density profile is described. This gives the halftone screen density profile for any form of recording medium characteristic curve and any type of nonlinearity by minimizing in mean-square sense the difference between desired and degraded outputs. The results of computer simulation for logarithmic and level slice functions are shown. Experimental results are obtained for an optimized halftone screen which matches the characteristic of a liquid crystal light valve (LCLV). An overall logarithmic transfer function is produced in real time.

The procedure for obtaining nonlinearities without halftone preprocessing is explained. Experimental results using a special birefringent liquid crystal device which

produces a 3-bit parallel A/D converter in real time are shown.

The variable grating mode (VGM) of liquid crystal light valve systems and its application in real-time nonlinear optical processing is discussed. A detailed analysis of the variable phase grating is made and from this the processing limitations of the VGM are determined. Experimental results showing real-time level-slicing with a VGM liquid crystal device are given.

Possible extensions and limitations of this work are discussed.

CHAPTER 1

INTRODUCTION

1.1 Introduction

Optical information processing is of great interest because of its capability for parallel processing of two dimensional data [1-1]. This inherent characteristic of optical processing makes it even more desirable for operation on two-dimensional data in the form of continuous-tone images. The analog nature of optical processing eliminates the sampling and quantization steps usually required before digital processing [1-2]. Linear operations such as convolution, correlation and Fourier transforming can be easily performed with optical processing systems [1-3].

To extend the usefulness of optical processors, other operations which are not necessarily linear are needed. Several efforts have been made to extend the flexibility of the optical processors to nonlinear operations such as logarithms, exponentiation, level slicing, pseudocolor, and analog-to-digital conversion. Marquet and Tsujiuchi [1-4] were the first to note that the halftone screen method could be used for nonlinear processing. Kato and Goodman

[1-5] obtained a logarithmic transformation using this method. Later Sawchuk and Dashiell [1-6] performed a level slice transformation with the halftone method. Lohmann and Strand [1-7] used this method to obtain analog-to-digital conversion. Pseudocolor with the halftone method was shown by Goodman and Liu [1-8].

None of the above nonlinear operations has been obtained in real time. This means that the nonlinear transformation of an input scene cannot be done without significant time delay. This is due to the fact that the input and output of these processors often rely on photographic techniques which are not a real-time realization. Recent developments have simplified the output problem: television and solid state devices are available to efficiently make use of the two-dimensional processed output. In many situations, the human eye or observer is the end user of the information, so that optical systems with their inherent two-dimensional nature are ideally suited to process pictorial information intended eventually for the human observer. The major difficulty lies with real-time input to these processors. Flexible real-time optical input modulators which can convert image information into a form for input to the processing system are needed. Significant research over the last several years has made progress on certain aspects of this problem. For ordinary linear optical processing a

2

great deal of progress has been made in developing fast, sensitive and easily usable real-time controllable replacements for photographic film. There are many different materials and systems, each with its own particular characteristics and tradeoffs [1-9 to 1-13]. One common characteristic of nearly all current devices is that they are intended to operate linearly with moderate contrast (photographic gamma of 3) over a broad dynamic range (densities of 0 to 2 or 3).

Our main goal in this research is to study the feasibility of doing nonlinear optical processing using a real-time image converter and present methods which successfully perform such operations. We will take two different approaches to this problem: 1) realization of the halftone process in real time; and 2) obtaining nonlinearities without halftone preprocessing. In the following we first introduce these two methods and then present the organization of this dissertation and its research contributions.

1.2   Real-Time Nonlinear Processing with Halftone Screens

In halftone nonlinear processing the continuous tone input picture is transformed into a binary picture by contact printing the continuous input data through a halftone screen onto a high-contrast recording medium. The product of the input and halftone screen transmittances is

clipped in the process, giving an array of binary dots whose size is a function of clip level, the input transmittance, and the halftone transmittance profile. The periodic nature of the halftone screen causes each subregion of the binary image corresponding to a constant input intensity to become quasi-periodic. When placed in the usual coherent optical filtering system, multiple diffraction orders appear in the Fourier transform plane because of the sampled input. The procedure for producing nonlinearities involves the use of one diffraction order combined with specially made halftone screens. A filter is placed in the Fourier plane that transmits the light around one diffraction order and blocks everything else. This in effect demodulates the image [1-3]. After the filtered diffraction order in the Fourier plane is inverse transformed, the continuous nonlinearly transformed output appears.

To realize this process in real time the photographic hard-clipping step should be replaced with some real-time method. This means that a real-time image converter is needed instead of the photographic film. Among presently developed devices, the liquid crystal light valve (LCLV) appears to be one of the most promising for use in optical processing systems [1-12,1-13]. It offers the advantages of 1) relatively good modulation transfer function (to 50 or 60 cycles/mm.); 2) fast, reusable, simple operation; 3)

room temperature operation; and 4) physical compactness with minimal controlling electronics required. This is the type of device that has been chosen for use in real-time halftone processing experiments.

Like most other real-time devices, the LCLV has generally been optimized for a large linear dynamic range, in contrast to the hard-clipping characteristic required for halftone processing. Given the choice of a linear LCLV for real-time halftone input, methods of predicting the degradation and compensating for the nonideal characteristic have been studied. Dashiell and Sawchuk have investigated some implications of non-ideal recording materials [1-14]. They found methods to compensate for finite gamma and finite saturation of the photographic film for monotonic nonlinearities.

We will consider a recording medium with a general shape of characteristic curve and try to compensate for its nonideal characteristic for general nonlinear functions.

1.3 Real-Time Nonlinear Processing Without Halftone Screens

We present two different methods for obtaining nonlinearities without halftone screens; a) direct; b) variable grating mode.

1.3.1 Direct Nonlinear Processing

To achieve nonlinear functions without halftone screens the inherent nonlinear characteristic of a recording medium or real-time transducer is directly used. This invovles the proper biasing and selection of operating points on a nonlinear curve to directly achieve a point-by-point intensity transformation. The processing may take place with incoherent illumination, avoiding the problems of speckle, phase noise, and the possible necessity for a laser source. It also requires a lower space-bandwidth product on the real-time input device because the halftone screen is eliminated. One drawback is that the flexibility of this technique is considerably more limited than halftoning. Tai, Cheng and Yu [1-15] have obtained a logarithmic nonlinearity directly (not in real time) using photographic film with special processing. .

1.3.2 Variable Grating Mode Nonlinearities

This method involves the conversion of different input intensities to a local phase-modulated grating whose spatial frequency is a function of the brightness. When the variable grating is placed in the front focal plane of a coherent Fourier transform processor, the difference in local spatial frequency should cause different input levels to be effectively placed at different points in the transform plane. By selective filtering and recombination of the transform components, various nonlinearities should

be possible. This method relies on the behavior of certain liquid crystal real-time devices which have shown the variable grating mode in experiments [1-16].

1.4 Organization of the Dissertation and Its Research Contributions

In the second chapter we review halftone nonlinear processing and characteristics of a liquid crystal light valve (LCLV) as a real-time image converter. We then explain how these two techniques can be combined to produce nonlinearities in real time. Chapter 3 presents a formulation of the halftone process with a nonbinary recording medium. This formulation relates the input and output intensities through the halftone screen parameters and the recording medium characteristic curve. The results for the binary recording medium are obtained as a special case of the general recording medium.

In Chapter 4 we use the formulation of chapter three to predict degradations due to the nonbinary characteristic of the recording medium. The result of computer simulations for different parameters of the recording medium characteristic are shown.

In Chapter 5 we try to compensate for the degradations by analytical methods. It is shown that for some nonlinearities we can obtain the halftone screen parameters

without any precompensation for the nonbinary recording medium.

.Chapter 6 gives a new analysis of the halftone process in which the halftone screen density function is considered as a discrete function. With such analysis we describe new optimized methods of halftone screen design. Results of computer simulation for different parameters of the recording medium characteristic are shown.

In Chapter 7 we discuss methods of making halftone screens and present experimental results for a real-time logarithmic process using an optimized halftone screen with a LCLV.

Chapter 8 explains how the nonlinear characteristic of a special birefringent LCLV can be used to generate an analog-to-digital conversion nonlinearity directly without any halftone preprocessing. The result of an experiment producing a 3-bit parallel A/D conversion in real time is shown.

In Chapter 9 the variable grating mode VGM of the liquid crystal and its application in real-time nonlinear processing is discussed. In a detailed analysis we determine the processing limitations of the VGM effect. Preliminary experimental results using a VGM liquid crystal device for real-time level slicing is presented.

Chapter 10 discusses conclusions and possible extensions of this work.

The specific research contributions of this work are now described. The analysis of the halftone nonlinear process has been extended to predict the nonlinear input-output curve with any recording medium characteristic curve shape and any halftone screen cell shape. New analytical design formulas for halftone cell shape with a piecewise linear recording medium and certain nonlinearities have been derived. New design procedures for halftone screen cell shape are presented which produce screens that compensate for the nonideal characteristic of the recording medium with any nonlinearity. New methods have been developed for obtaining real-time nonlinear functions without halftone preprocessing using the birefringent liquid crystal device and the variable grating mode (VGM) liquid crystal device. Experimental demonstrations of logarithm, 3-bit parallel A/D conversion, and level slice functions in real time have been made.

CHAPTER 2

REVIEW OF HALFTONE NONLINEAR PROCESSING AND

THE LIQUID CRYSTAL LIGHT VALVE

In this chapter we review nonlinear halftone processing and the liquid crystal light valve (LCLV) as a real-time image converter. We then discuss how they can be combined to produce parallel nonlinear transformations in real time.

2.1 Nonlinear Optical Processing with Halftone Screens

Halftone nonlinear optical processing methods rely on a pulse - width modulation intermediate step, in which the input data is coded, and a filtering part that recovers a continuous-tone picture from the coded one [2-1]. The pulse-width modulation step, often called the halftone process, transforms a continuous-tone picture into a picture with only two levels of density. The resulting picture is called a halftoned picture. It consists of groups of dots or bars depending on whether the halftone screen used in the process is one or two-dimensional. Corresponding to each density level in any part of the original picture a group of equal-area dots or a series of

equal-width bars is present in the halftoned picture. The area of the dots or the width of the bars is a function of the corresponding density level in the original picture and the density profile of the halftone screen. Thus, the information about the original picture density levels is encoded into areas of the dot groups or the widths of bars present in the halftoned picture. The method is shown in Fig. 2.1. The input illumination passes through the original picture transparency and falls on a periodic array of partially absorbing dots or lines (halftone screen). The light transmitted by the halftone screen falls upon a hard-clipping photograpic film which acts as a thresholding device and records a binary image. The operation for a one dimensional halftone screen is shown in Fig. 2.2. Note that the densities of the input picture transparency and the halftone screen add together to effectively partially absorb this illumination at any point. Depending on whether this sum of densities is less than or greater than the threshold density of the hard-clipping film, it will or will not be activated. This causes the amplitude transmittance for the encoded picture transparency to look like Fig. 2.2(e). From the amplitude transmittance of the halftoned picture we see that the effect of the halftone process on the input picture is equivalent to sampling. If this sampled picture is used as an input to a coherent optical processing system as shown in Fig. 2.3, different

INPUT ILLUMINATION

```
┌──────────────────────────────┐
│  ORIGINAL  TRANSPARENCY      │
├──────────────────────────────┤
│  HALFTONE  SCREEN            │
├──────────────────────────────┤
│  HARD-CLIPPING  FILM         │
└──────────────────────────────┘
```
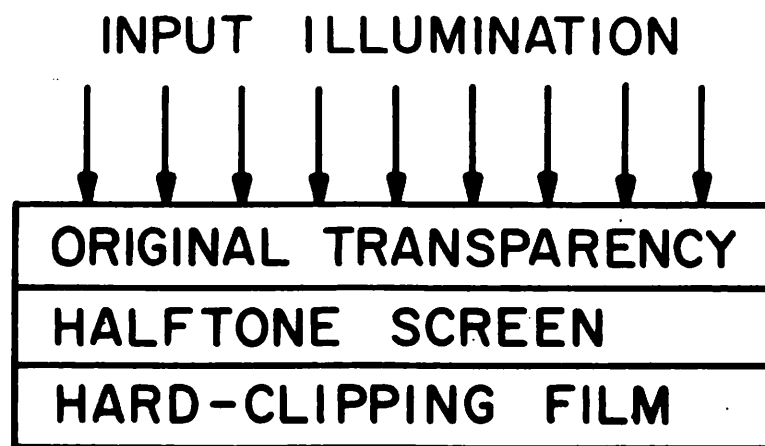
Figure 2.1   Production of halftoned picture
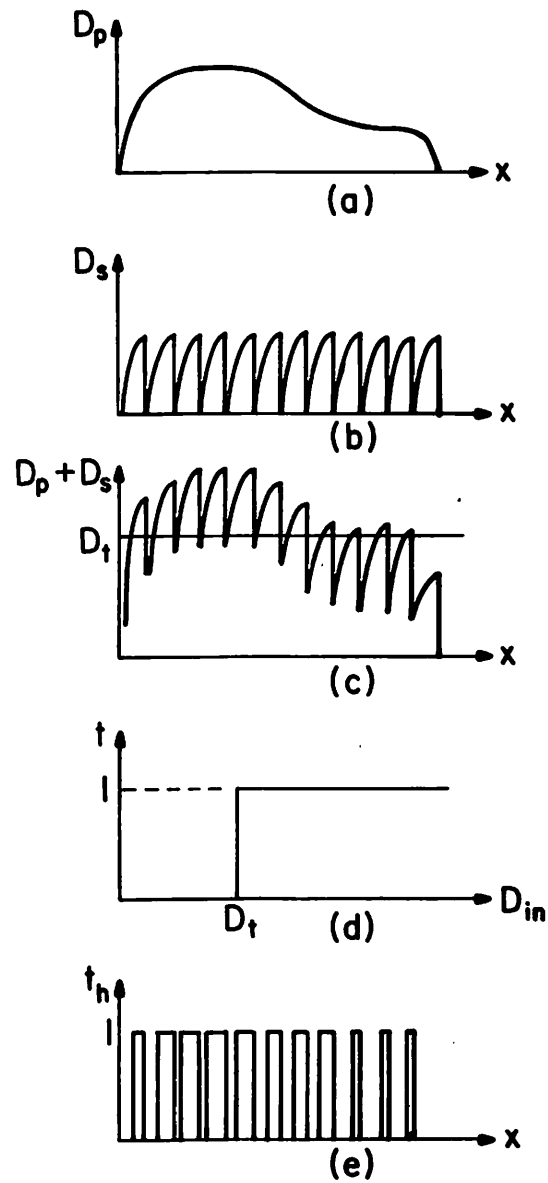
Figure 2.2 The halftone encoding process
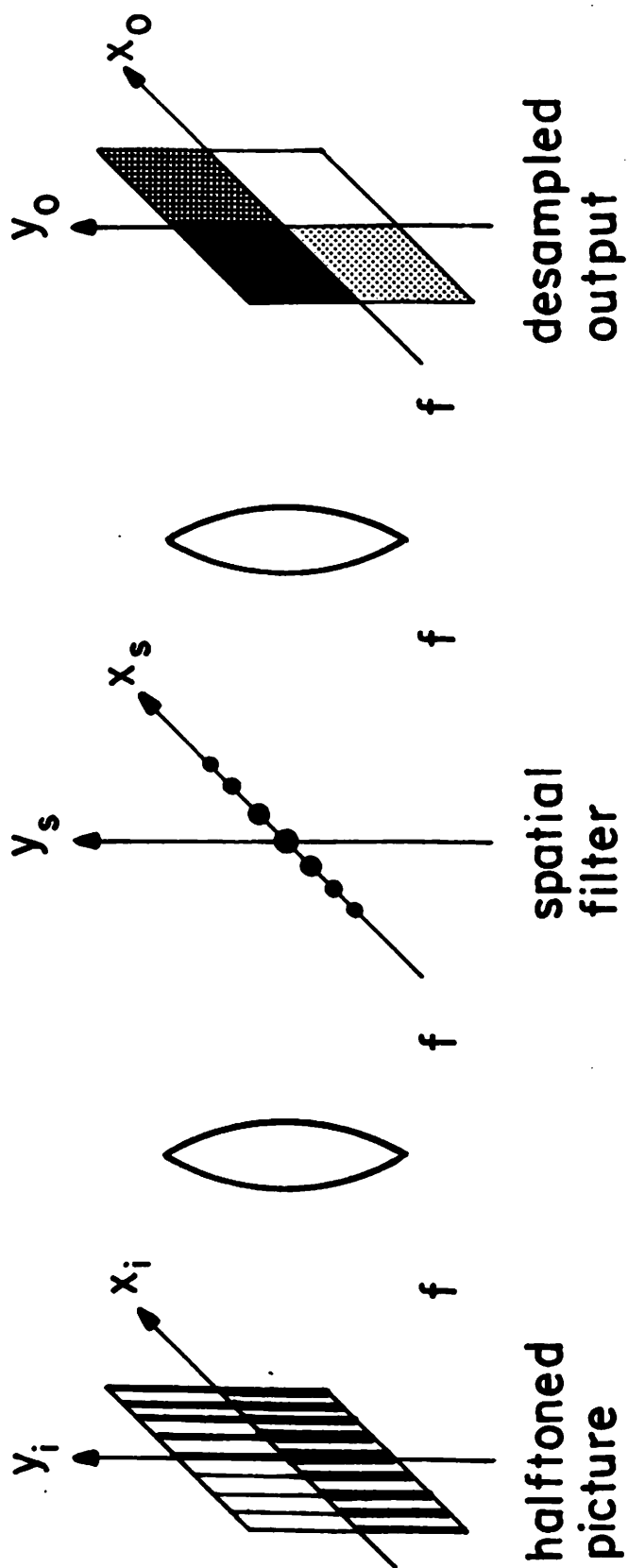
Figure 2.3 Desampling of the halftoned picture in a coherent optical processing system

diffraction orders appear in the filter plane. To be able to successfully filter out any single order, the spatial frequency of the halftone screen should be much greater than the maximum spatial frequency present in the input picture. Assuming that such a requirement holds, we can put a simple spatial filter in the Fourier plane and select one of the orders. The selected order is then transformed back to give a desampled image in the output plane of the processor. The intensity at any point of this output picture is then a nonlinear function of the corresponding point in the input. The form of this nonlinearity is determined by the shape of the halftone screen density profile and the diffraction order selected in the spatial filtering system. Details of the mathematical relationship between the output and input intensities will be worked out in the following chapters.

2.2 Liquid Crystal Light Valve

The term "liquid crystal" is applied to substances whose rheological behavior is similar to that of fluids but whose optical behavior is similar to the crystalline state over a given temperature range [2-2]. One class of liquid crystals called nematics is the most common type used in devices whose properties interest us. The nematic liquid crystal phase is formed by many organic compounds. The most significant common feature of all these compounds is

that they have long rodlike molecules [2-3]. When an electrical field is applied to a liquid crystal, the material is reoriented in space by both the field and ionic conduction effects. If the conductivity of the liquid crystal is sufficient, conduction effects will predominate, whereas in the absence of significant conductivity effects, the nematic is simply reoriented by the field [2-4].

The field effect operation of a LC device is based on the electrically controlled birefringence characteristic of the nematic liquid crystal. The LC layer is normally placed between a transmitting electrode and a reflecting mirror. An incident polarized beam of light on the transmitting electrode will be reflected back from the mirror while traveling through the liquid crystal as shown in Fig. 2.4.

Depending on the voltage across the liquid crystal, the reflected light will experience some phase retardation due to reorientation of the LC molecules by the field. This phase retardation can be detected as an intensity variation if we pass the reflected light through a second polarizer that is either parallel or perpendicular to the one in the incident beam. With such arrangement we have a light valve in which the electric field variation across the LC controls the intensity variation of the output beam (reflected light). Now, if we introduce a photoconductor

OFF

M   TE   LC   TE   P

LIGHT IN

A

ON   P

LIGHT IN

LIGHT OUT

A

θ = ELLIPTICALLY POLARIZED LIGHT
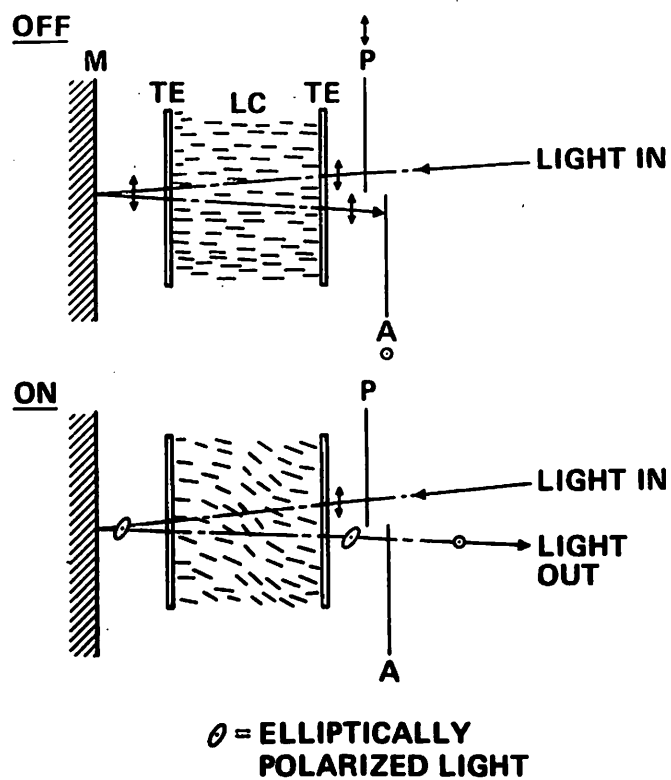
Figure 2.4   Diagram of a tunable birefringence field effect cell (P=Polarizer, A=Analyzer, TE=Transparent electrode, M=Mirror, LC=Liquid crystal

layer in the system we can transform the intensity variation of a second light source (input light) to voltage variation across the photoconductor that in turn can be transferred to voltage variation across LC. Hence, the end result is a light controlled light valve. This is shown in Fig. 2.5 with a light blocking layer added to separate the input and output beams. If the LC is twisted by 45° we could get a dark off-state and the resulting device would be the so-called standard linear LCLV [2-5]. This device can be made to operate with a speed of about 30 frames/second or TV frame rate.

2.3 Real-Time Nonlinear Optical Processing with Halftone Screens

To realize the halftone process in real time both the halftoning step and output readout should be performed rapidly. As described in chapters 1 and 2 the output is easy to obtain in real time. To achieve the halftoning in real time the photographic hard-clipping medium should be replaced with a real-time image converter. The ideal real-time device should have two properties: 1) the ability to convert an incoherent image to a coherent image; 2) infinite gamma. Everywhere in this work we define the parameter gamma as the slope of the linear portion of the output amplitude transmittance vs. log of input intensity of the real-time image transducer. Our definition of gamma
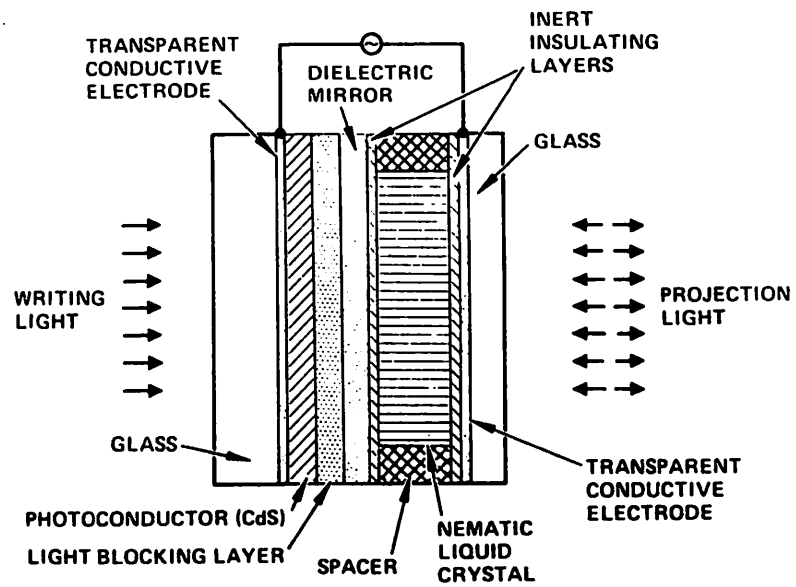
Figure 2.5   Cross-sectional schematic of LCLV

should not be confused with the photographic definition of gamma as the slope of the linear part of density vs. log of the input intensity. The first property is common to many image transducers and the liquid crystal light valve (LCLV) described shares this property. On the other hand, almost all real-time image converters are optimized for linear processing and do not possess the second property. This can be seen from Fig. 2.6 where a typical characteristic of a standard linear LCLV is shown.

In the following chapters we investigate this problem by developing a new formulation of the halftone process and presenting methods for the compensating for nonideal characteristic of real-time image converters.
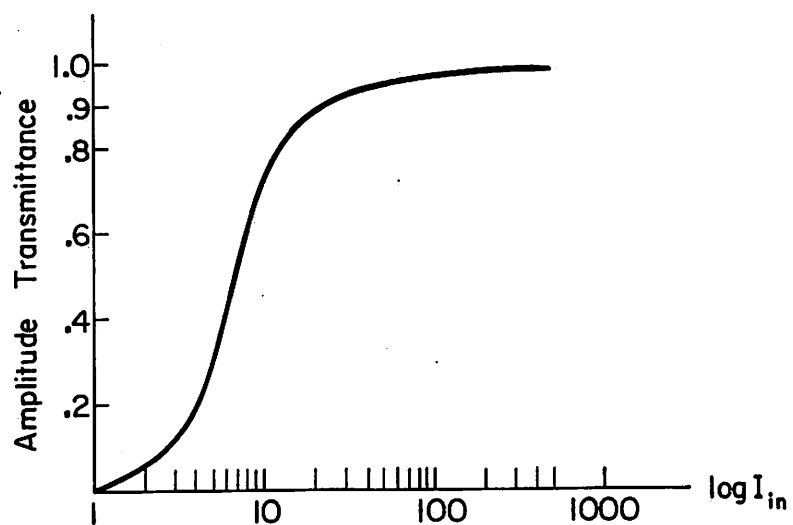
Figure 2.6   Amplitude transmittance vs log exposure
characteristic of a standard linear LCLV

CHAPTER 3

HALFTONE PROCESSES WITH GENERAL RECORDING MEDIUMS

Halftone nonlinear processes, described in chapter two, have been formulated with the assumption that a binary recording medium is used in the halftoning step [3-1]. With this assumption, once the output intensity is expressed as a function of input intensity and the halftone screen density profile (analysis), we can easily invert the problem and get the halftone screen density profile given the relationship between the output and input intensities (synthesis). Unfortunately, almost all recording media deviate from the binary assumption. This deviation which is quite small for high contrast photographic films, is quite noticeable for any real-time image converter at the present time [3-2,3-3]. Consequently, to appreciate the halftone processes fully, we should remove the assumption of a binary recording medium from the very beginning of our mathematical formulation of these processes. Dashiell and Sawchuk have considered the effects of finite gamma and saturation density of the recording medium on halftone process by modifying the formulas for the ideal recording medium [3-4].

This formulation does not predict the effects of the curved portions of the recording medium characteristic curve on the overall nonlinear transformation. Moreover the formulation is restricted to monotonic halftone cell shapes. In this chapter we present a formulation of the halftone process which considers a recording medium with a characteristic curve of general shape and predict the final degradation of the output due to such characteristic curve for any halftone screen cell shape. This formulation is used for the case of a binary recording medium and its result is compared with previous derivations.

3.1 Mathematical Formulation

For a general recording medium in the halftone process we should replace the binary characteristic curve of Fig. 2-2a by a curve like Fig. 3.1. This causes the amplitude transmittance of the halftoned picture to consist of pulses that are no longer rectangular as shown in Fig. 3.2. The amplitude, width and shape of these pulses depend on the input picture density levels, halftone screen density profile, and the shape of the characteristic curve of the recording medium. Each group of these pulses corresponds to a constant intensity subregion in the input picture. The period L (Fig. 3.2) of the halftone screen is chosen to be small in comparison with the period of the highest spatial frequency component in the input picture,
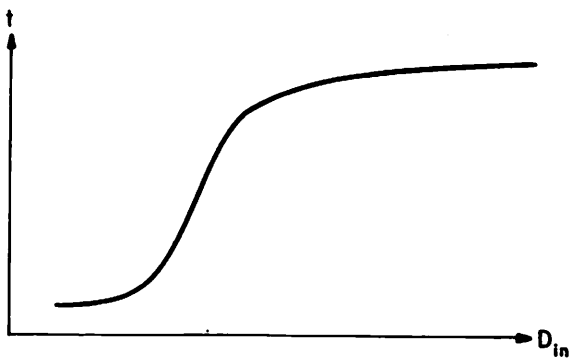
Figure 3.1   Transmittance vs input illumination density
for a general recording medium
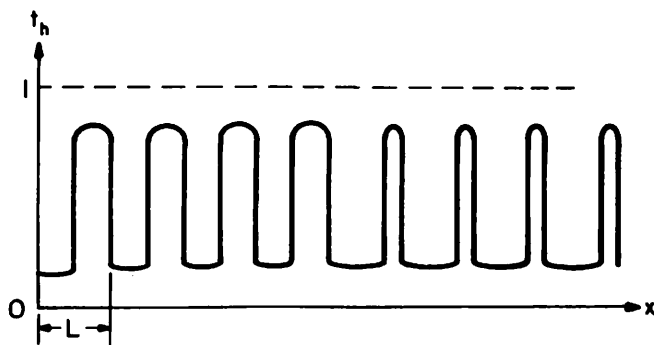
Figure 3.2   Amplitude transmittance of the halftoned
transparency made with a recording medium
like that in Figure 3.1

so any local region of the amplitude transmittance of the halftoned transparency is approximately a periodic sequence of pulses and can be expanded in complex Fourier series

$$t_h(x) = \sum_{k=-\infty}^{+\infty} \underline{B}_k \exp(-j2\pi kx/L) \qquad (3.1)$$

where

$$\underline{B}_k = \frac{1}{L} \int_0^L t_h(x) e^{j\frac{2\pi kx}{L}} dx. \qquad (3.2)$$

In the above sum each term represents a grating order, and when we take the Fourier transform of the halftoned picture in the coherent optical processor, these orders appear in the Fourier plane as isolated spectral islands. The spatial filter in this plane selects a single order. Hence, the resulting intensity distribution at the processor output is

$$
\begin{aligned}
I_0(I_{in}, k) &= |\underline{B}_k|^2 \\
&= \left| \frac{1}{L} \int_0^L t_h(x) e^{j\frac{2\pi kx}{L}} dx \right|^2
\end{aligned} \qquad (3.3)
$$

which relates the intensity at any point of the output picture to the amplitude transmittance of the halftoned picture and the selected order. Now $t_h(x)$ can be related to the input intensity as follows. Let the local input picture intensity that produced the above train of pulses

on the halftoned picture be denoted by I_{in}. If the density variation of one period of the halftone screen is represented by $f(x)$, then the intensity transmitted by the halftone screen is $I_{in} \cdot 10^{-f(x)}$ [3-5]. If we let the amplitude transmittance versus log exposure curve of the recording medium be described by $g(logE)$, then we can write

$$t_h(x) = g\{log[I_{in} \cdot 10^{-f(x)}]\} = g[logI_{in} - f(x)] \qquad (3.4)$$

and replacing it in Eq. (3.3), we have

$$I_o(I_{in}, k) = \left| \frac{1}{L} \int_0^L g[logI_{in} - f(x)] e^{j\frac{2\pi kx}{L}} dx \right|^2 \qquad (3.5a)$$

which relates the intensity at any point of the output picture to the intensity of the corresponding point in the input picture through a nonlinear integral relationship. When the specific forms of $g(logE)$ and $f(x)$ are substituted in this relationship and the integral is solved, the overall relation between the output intensity and the input intensity is nonlinear. This nonlinearity depends on $g(logE)$, $f(x)$ and the value of the order selected.

The above formulation could be performed in terms of the intensity transmittance, $\tau(x)$, of the halftone screen rather than its density profile. In this case, for simplicity, we also need the transmittance versus exposure curve of the recording medium, $T(E)$, rather than $g(logE)$. If we go through similar arguments as above we get

$$I_o(I_{in}, k) = \left| \frac{1}{L} \int_0^L T[I_{in} \cdot \tau(x)] e^{j\frac{2\pi kx}{L}} dx \right|^2 . \qquad (3.5b)$$

3.2 Binary Recording Medium

The characteristic curve of a binary recording medium is shown in Fig. 3.3.   Ideally a=0, and b=1. Note that this form of characteristic curve is applicable to a positive transparency.   We could also consider the more familiar negative transparency although the basic results remain the same. We will choose the positive transparency curve because they are more similar to the characteristic curves for real-time devices. We now simplify the general relationship of Eq. (3.5a) using the characteristic curve of Fig. 3.3.

3.2.1 Zero order

For the zero order case k=0, and Eq. (3.5a) becomes

$$I_o(I_{in}, 0) = \left\{ \frac{1}{L} \int_0^L g[\log I_{in} - f(x)] dx \right\}^2. \qquad (3.6)$$

Considering Fig. 3.3 we can write

$$\text{if} \quad \log I_{in} - f(x) < \log I_r \quad \text{then} \quad g(\log E) = a \qquad (3.7)$$

and

$$\text{if} \quad \log I_{in} - f(x) \geq \log I_r \quad \text{then} \quad g(\log E) = b. \qquad (3.8)$$
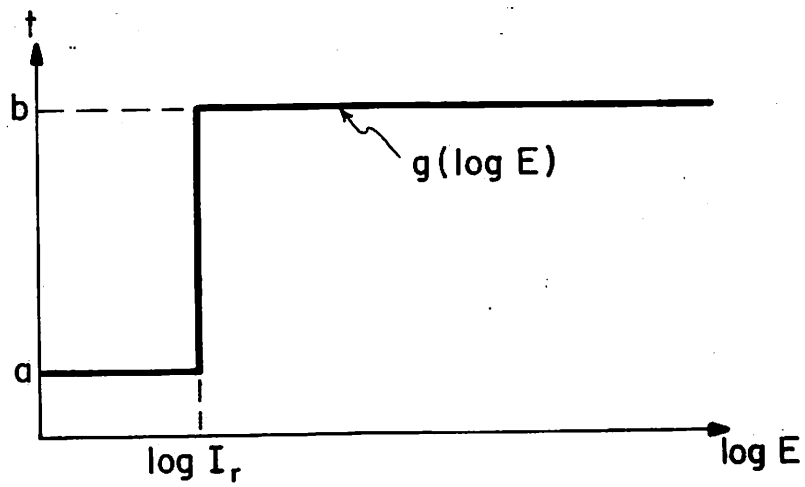
Figure 3.3   Characteristic curve of a binary
recording medium

Assuming f(x) to be a monotonically increasing function we have

$$\text{if} \quad x > f^{-1}(\log\frac{I_{in}}{I_r}) \qquad \text{then} \qquad g(\log E) = a \qquad (3.9)$$

and

$$\text{if} \quad x \leq f^{-1}(\log\frac{I_{in}}{I_r}) \qquad \text{then} \qquad g(\log E) = b. \qquad (3.10)$$

When used in Eq. (3.6), this gives

$$I_o(I_{in},k) = \{\frac{1}{L}[\int_0^{f^{-1}(\log\frac{I_{in}}{I_r})} b\,dx + \int_{f^{-1}(\log\frac{I_{in}}{I_r})}^0 a\,dx]\}^2 \qquad (3.11)$$

or after simplification

$$I_o(I_{in},k) = [a+\frac{(b-a)}{L}f^{-1}(\log\frac{I_{in}}{I_r})]^2 \qquad (3.12)$$

which is the same as the results obtained previously for the case a=0, b=1 [3-1] except for the fact that it is obtained in a more straightforward manner and can easily be generalized.

3.2.2 Nonzero order

For k=0, the above simplifications for the characteristic curve can be used in Eq. (3.5) to obtain

$$I_o(I_{in},k) = |\frac{1}{L}[\int_0^{f^{-1}(\log\frac{I_{in}}{I_r})} be^{j\frac{2\pi kx}{L}}\,dx + \int_{f^{-1}(\log\frac{I_{in}}{I_r})}^L ae^{j\frac{2\pi kx}{L}}\,dx]|^2. \qquad (3.13)$$

let

$$x_1 = f^{-1}(\log\frac{I_{in}}{I_r})$$

then

$$I_o(I_{in}, k) = \left| \frac{(b-a)}{2\ kj}[e^{j\frac{2\pi kx_1}{L}} - 1] \right|^2$$

$$= \frac{(b-a)^2}{4\pi^2 k^2}\ (e^{j\frac{2\pi kx_1}{L}} - 1)(e^{-j\frac{2\pi kx_1}{L}} - 1)$$

$$= \frac{(b-a)^2}{4\pi^2 k^2}\ (1 - \cos\frac{\pi kx_1}{L})$$

(3.14)

$$= \frac{(b-a)^2}{\pi^2 k^2}\ \sin^2\frac{\pi kx_1}{L}$$

after replacing   for x_1 we have

$$I_o(I_{in}, k) = \frac{(b-a)^2}{\pi^2 k^2}\ \sin^2[\frac{\pi k}{L}\ f^{-1}(\log\frac{I_{in}}{I_r})].$$ (3.15)

This also agrees with the previously obtained results [3-1].

3.3 Halftone Screen Density Functions for Some Useful Nonlinearities

The halftone screen density function for different nonlinearities with the general recording medium will be determined in later chapters. Here, for our reference we will find the halftone screen density functions for the

binary recording medium for several specified nonlinearities.

3.3.1 Logarithmic Transformation

This transformation can be obtained in the zero order. Hence from Eq. (3.14)

$$I_0(I_{in}, 0) = [a + \frac{(b-a)}{L} f^{-1} (\log \frac{I_{in}}{I_r})]^2. \qquad (3.16)$$

For a logarithmic transformation

$$I_0(I_{in}, 0) = K \log(I_{in}/I_r) \qquad (3.17)$$

where K is a constant. Combining Eq. (3.17) and Eq. (3.16), we get

$$K \log(\frac{I_{in}}{I_r}) = [a + \frac{(b-a)}{L} f^{-1} (\log \frac{I_{in}}{I_r})]^2. \qquad (3.18)$$

If we let $\log \frac{I_{in}}{I_r}$ be represented by $f(x)$ and note that $f^{-1}[f(x)] = x$ then Eq. (3.18) can be written as

$$f(x) = \frac{1}{K}[a + \frac{b-a}{L} x]^2 \qquad (3.19)$$

which gives the density profile of the halftone screen for a logarithmic transformation.

3.3.2 Exponential Transformation

This transformation can also be obtained in the zero order. We want

$$I_o(I_{in}, 0) = \alpha(\beta)^{\frac{I_{in}}{I_r}} \qquad (3.20)$$

where α and β are constants. Combining Eq. (3.20) and Eq. (3.16) we get

$$\alpha(\beta)^{\frac{I_{in}}{I_r}} = [a + \frac{(b-a)}{L} f^{-1}(\log\frac{I_{in}}{I_r})]^2 \qquad (3.21)$$

or

$$\alpha(\beta)^{\left[10^{\log(\frac{I_{in}}{I_r})}\right]} = [a + \frac{(b-a)}{L} f^{-1}(\log\frac{I_{in}}{I_r})]^2. \qquad (3.22)$$

From this we can write

$$\alpha(\beta)^{\left[10^{f(x)}\right]} = [a + \frac{(b-a)}{L} x]^2. \qquad (3.23)$$

Taking the logarithm of both sides twice and simplifying, we obtain

$$f(x) = \log\{2\log[a + \frac{(b-a)}{L} x] - \log\alpha\} - \log\log\beta. \qquad (3.24)$$

This gives the density profile of the halftone screen for an exponential transformation.

3.3.3 Power Transformation

This is another example of a transformation possible in the zero order. We want

$$I_o(I_{in}, 0) = \gamma(\frac{I_{in}}{I_r})^\lambda \qquad (3.25)$$

where γ and λ are constants. Combining this with Eq. (3.16) we have

$$\gamma(\frac{I_{in}}{I_r})^\lambda = [a + \frac{(b-a)}{L} f^{-1}(\log\frac{I_{in}}{I_r})]^2 \qquad (3.26)$$

or

$$\gamma(10)^{\lambda\log\frac{I_{in}}{I_r}} = [a + \frac{(b-a)}{L} f^{-1}(\log\frac{I_{in}}{I_r})]^2 \qquad (3.27)$$

from which we can write

$$\gamma(10)^{\lambda f(x)} = [a + \frac{(b-a)}{L} x]^2 . \qquad (3.28)$$

Taking the logarithm of both sides of Eq. (3.28) and simplifying we obtain

$$f(x) = \frac{2}{\lambda}\log[a + \frac{(b-a)}{L} x] - \frac{1}{\lambda}\log\gamma . \qquad (3.29)$$

This gives the density profile of the halftone screen for a power transformation.

3.3.4 Level Slice Transformation

In this case the desired input and output intensities are

33

$$I_o = \begin{cases} 0 & \text{for } I_{in} < I_1 \\ c^2 & \text{for } I_1 \leq I_{in} < I_2 \\ 0 & \text{for } I_2 \leq I_{in} \end{cases} \qquad (3.30)$$

and this equation is shown in Fig. 3.4. To obtain this transformation we should use the first order, i.e., k=1 in Eq. (3.15). We have

$$I_o(I_{in}, 1) = \frac{(b-a)^2}{\pi^2} \sin^2[\frac{\pi}{L}f^{-1}(\log\frac{I_{in}}{I_r})] \qquad (3.31)$$

Combining Eq. (3.30) and Eq. (3.31) we can see that

$$f^{-1}(\log\frac{I_{in}}{I_r}) = 0 \qquad \text{for} \qquad I_{in} < I_1,$$

$$\qquad (3.32)$$

$$f^{-1}(\log\frac{I_{in}}{I_r}) = \frac{L}{\pi}\sin^{-1}\frac{c\pi}{b-a} \qquad \text{for } I_1 \leq I_{in} < I_2,$$

and

$$f^{-1}(\log\frac{I_{in}}{I_r}) = L \qquad \text{for} \qquad I_{in} \geq I_2.$$

These are summarized in Fig. 3.5 where

$$r = \frac{L}{\pi}\sin^{-1}\frac{c\pi}{b-a} \qquad (3.33)$$

From Fig. 3.5 we can obtain the density profile of the halftone screen for the level slice transformation as shown in Fig. 3.6.
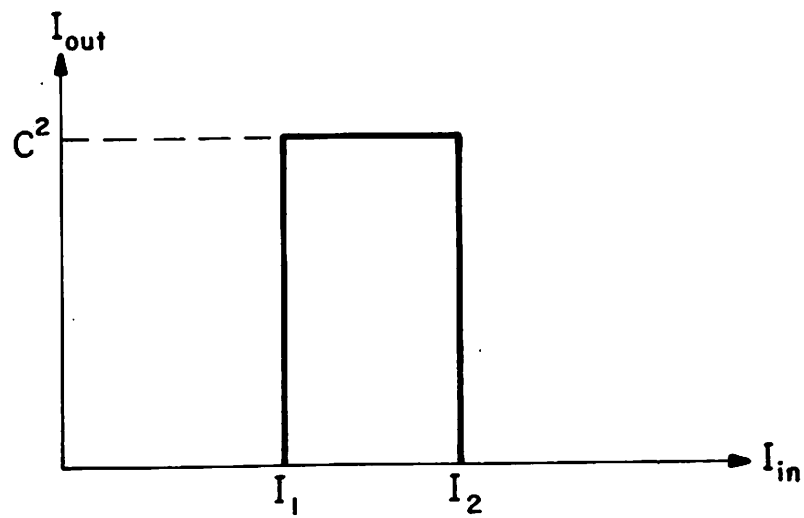
Figure 3.4 Level slice transformation
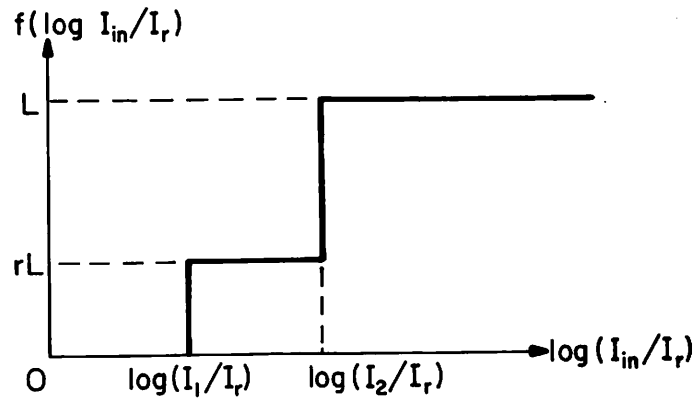
Figure 3.5   Inverse function of the density profile of
             the halftone screen which performs the level
             slice transformation of Figure 3.4
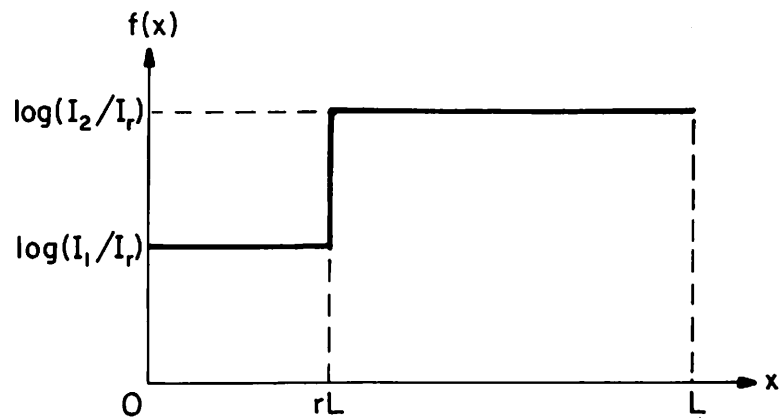
Figure 3.6   Density profile of the halftone screen which
             performs the level slice transformation
             of Figure 3.4

CHAPTER 4

DEGRADATION OF THE OUTPUT INTENSITY IN THE

HALFTONE PROCESS

The nonbinary characteristic of the recording medium in the halftone process is the main source of degradation of the output intensity. This analysis theoretically predicts the departure of the nonlinearity achieved from the ideal so that a compensation for this effect can be included in the design of the halftone screen. A typical characteristic curve of a nonbinary recording medium is shown in Fig. 4.1. Comparing this figure with the characteristic curve of a binary medium shown in Fig. 3.3, we can see that there are three distinct regions in these two curves that make them different. As shown in Fig. 4.1, they are usually referred to as the toe, the linear part, and the shoulder of the characteristic curve. The effect of finite gamma and finite saturation of the recording medium on the halftone process has been considered before [4-1]. In the following analysis, we will show how each of these three regions of the characteristic curve contributes to the final degradation of the output intensity. The analysis will be done separately for the zero and nonzero
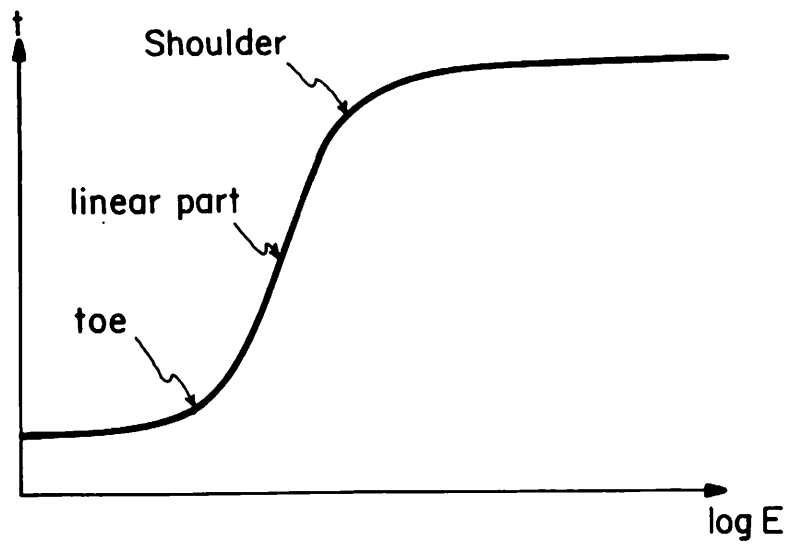
Figure 4.1   Typical characteristic curve of a
             nonbinary recording medium

orders.

4.1 Zero Order

The general formula relating the output intensity in the zero order to the input intensity from Eq. (3.6) is

$$I_o(I_{in}, 0) = \{\frac{1}{L} \int_0^L g[\log I_{in} - f(x)]dx\}^2. \qquad (4.1)$$

To find out how different parts of the nonbinary recording medium characteristic curve affect $I_o(I_{in}, 0)$, we can use the halftone screen density profile that is designed for the binary recording medium and some model characteristic curves which are binary in every respect except for their toe, linear part, or shoulder as shown in Fig. 4.2, Fig. 4.3, and Fig. 4.4. This has been done for the logarithmic process as explained in the following sections.

4.1.1 Degradation Due to the Linear Toe of the Recording Medium Characteristic Curve in the Logarithmic Process

In this case

$$I_o(I_{in}, 0) = K \log(I_{in}/I_r) \qquad (4.2)$$

where K is a constant and I_r is shown in Fig. 3.3. For the ideal recording medium the halftone screen density profile from Eq. (3.19) is

$$f(x) = \frac{1}{K}[a + \frac{(b-a)}{L}x]^2. \qquad (4.3)$$
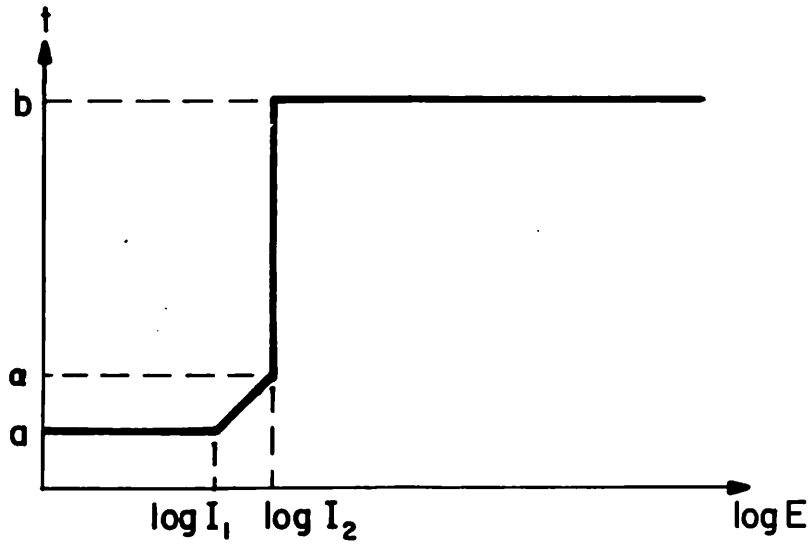
Figure 4.2   Characteristic curve of a recording
medium with a linear toe
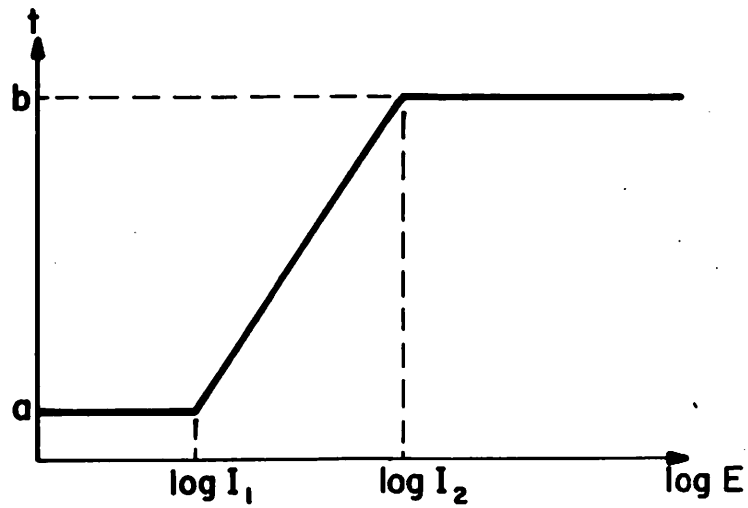
Figure 4.3   Characteristic curve of a recording
             medium with a linear part
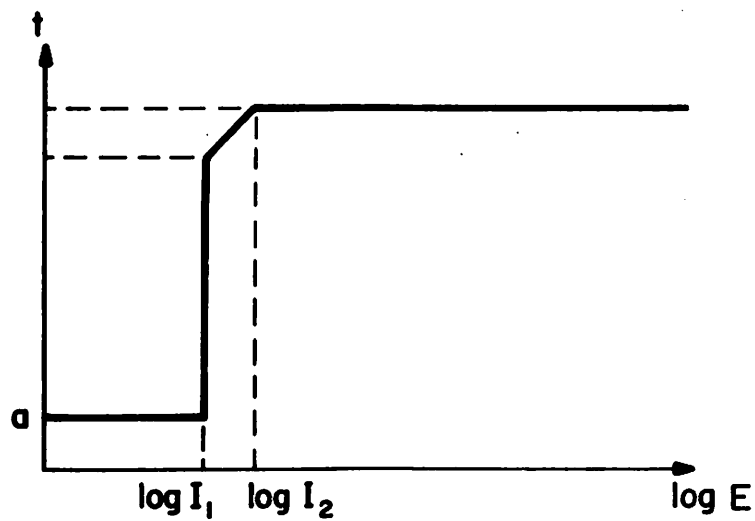
Figure 4.4   Characteristic curve of a recording
             medium with a linear shoulder

41

Let

$$t = \log(I_{in}/I_r)$$

and   (4.4)

$$\varepsilon = \log(I_2/I_1)$$

and now use Eq. (4.2), Eq. (4.3) and the characteristic curve of Fig. 4.3 in Eq. (4.1). For simplicity assume a=0, b=1. After some simplification the result is

$$
I_0(I_{in},0)=\begin{cases}
\dfrac{4\alpha^2 K}{9\varepsilon^2}(t+\tfrac{\varepsilon}{2})^3 & \text{for} \quad \log I_{in} < \log I_2 \\[4mm]
\{\dfrac{2\alpha K^{1/2}}{3\varepsilon}[(t+\tfrac{\varepsilon}{2})^{3/2}-(t-\tfrac{\varepsilon}{2})^{3/2}]+K^{1/2}(1-\alpha)(t-\tfrac{\varepsilon}{2})^{1/2}\}^2 \\[2mm]
\qquad\qquad\qquad \text{for} \quad \log I_2 \le \log I_{in} < \dfrac{1}{K} \\[4mm]
[K^{1/2}(1-\alpha)(t-\tfrac{\varepsilon}{2})^{1/2}+\dfrac{\alpha}{\varepsilon}(t-\tfrac{\varepsilon}{2})-\dfrac{2\alpha K^{1/2}}{3\varepsilon}(t-\tfrac{\varepsilon}{2})^{3/2}-\dfrac{\alpha}{3\varepsilon K}+\alpha]^2 \\[4mm]
\qquad\qquad\qquad \text{for} \quad \dfrac{1}{K} \le \log I_{in}
\end{cases}
\qquad (4.5)
$$

The quantity α is defined in Fig. 4.3. These equations have been calculated for $\alpha =0.1$ and different ε's. This means that the height of the toe is kept constant while its slope changes. This way the input intensity range over which the linear toe exists is extended to higher values of ε. The result is shown in Fig. 4.5. From this figure it is seen that increasing ε causes the transfer curve to lose some dynamic range while its shape is preserved. If we assume a curved toe rather than a linear toe, we expect to get some small change in the shape of the transfer curve

42

Figure 4.5 Logarithmic transfer function for a
recording medium with a linear toe

(a)   binary medium ε=0;   (b)   ε=1.0;
(c)   ε=0.33;   (d)   ε=0.20;

but same loss of dynamic range.

4.1.2 Degradation Due to the Linear Part of the Recording Medium Characteristic Curve in the Logarithmic Process

Using the assumptions of section 4.1.1 and the characteristic curve of Fig. 4.4 in Eq. (4.1) we get after some simplification

$$I_o(I_{in}, 0) = \begin{cases} \frac{4K}{9\varepsilon^2}(t-\frac{\varepsilon}{2})^3 & \text{for} \quad \log I_{in} < \log I_2 \\ \frac{4K}{9\varepsilon^2}[(t+\frac{\varepsilon}{2})^{3/2}-(t-\frac{\varepsilon}{2})^{3/2}]^2 & \text{for } \log I_2 \leq \log I_{in} \leq \frac{1}{K} \\ [\frac{1}{\varepsilon}(t-\frac{\varepsilon}{2})-\frac{2K^{1/2}}{3\varepsilon}(t-\frac{\varepsilon}{2})^{3/2}-\frac{1}{3\varepsilon K}+1]^2 & \text{for } \frac{1}{K} \leq \log I_{in} \end{cases} \quad (4.6)$$

where we assume a=0 and b=1. These equations have been calculated for different values of ε. The result is shown in Fig. 4.6. From this figure, we can see that extending the linear part of the recording medium characteristic curve to larger input intensity values causes the transfer curve to lose dynamic range and also to change shape. Although the loss in dynamic range is undesirable, the transfer curve shape change is even less desirable.

4.1.3 Degradation Due to the Linear Shoulder of the Recording Medium Characteristic Curve in the Logarithmic Process
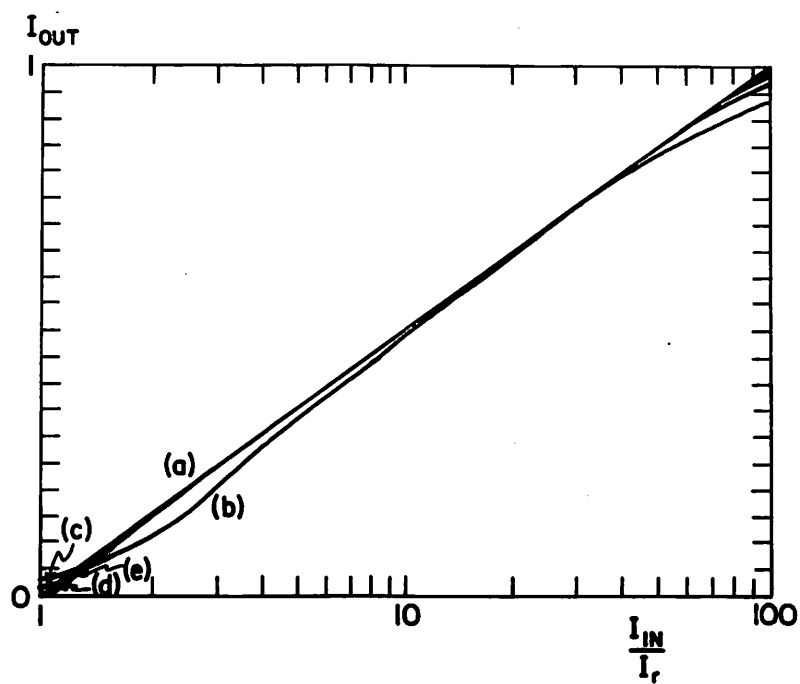
Figure 4.6   Logarithmic transfer function for a
              recording medium with a linear part

(a) binary medium ε=0;   (b)   ε=1.0;
(c)   ε=0.33;   (d)   ε=0.20;   (e)   ε=0.10

Using the assumptions of section 4.1.1 and the characteristic curve of Fig. 4.5 in Eq. (4.1) we get after some simplification

$$I_o(I_{in}, 0) = \begin{cases} K[\frac{2(1-\beta)}{3\epsilon}(t+\frac{\epsilon}{2})^{3/2} + \beta(t+\frac{\epsilon}{2})^{1/2}]^2 & \text{for } \log I_{in} < \log I_2 \\[2em] K\{\frac{2(1-\beta)}{3\epsilon}[(t+\frac{\epsilon}{2})^{3/2} - (t-\frac{\epsilon}{2})^{3/2}] + \beta(t+\frac{\epsilon}{2})^{1/2}\}^2 \\ \qquad \text{for } \log I_2 \le \log I_{in} < 1/K \\[2em] [\frac{(1-\beta)}{\epsilon}(t-\frac{\epsilon}{2}) - \frac{2(1-\beta)}{3\epsilon}(t-\frac{\epsilon}{2})^{3/2}]^2 \\ \qquad \text{for } \frac{1}{K} \le \log I_{in} \end{cases} \qquad (4.7)$$

where we assume a=0 and b=1 as before. The result of the above equations for $\beta = .9$ and different values of ϵ are shown in Fig. 4.7. The transfer curve change is seen to be similar to the one for the linear toe.

Noting that the input intensity range over which the toe and shoulder of the recording medium characteristic curve are present is usually much smaller than the range for the linear part, we can conclude that the linear part is responsible for most of the degradation in the output intensity for a monotonic nonlinearity like the logarithmic process.

4.2 Nonzero Order

The general formula relating the output intensity in the nonzero order to the input intensity from Eq. (2.5) is
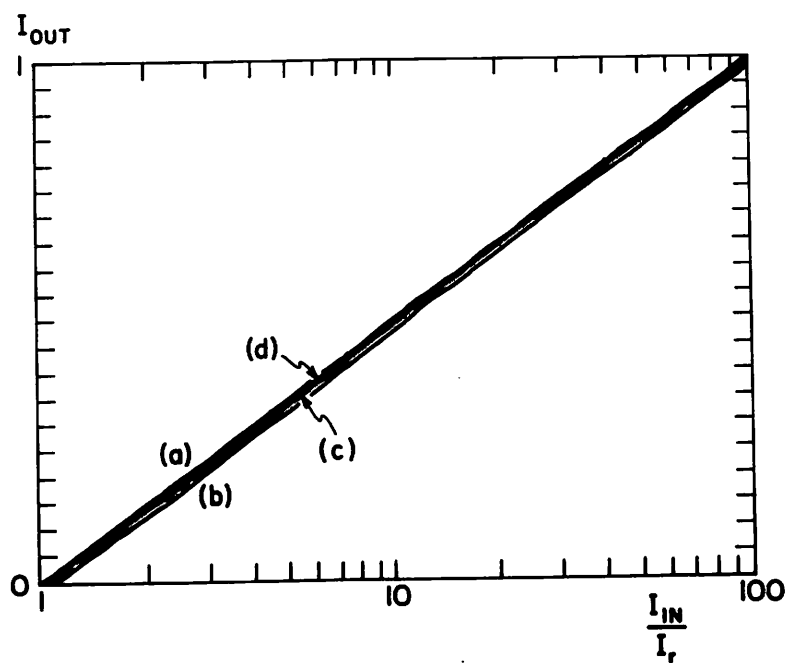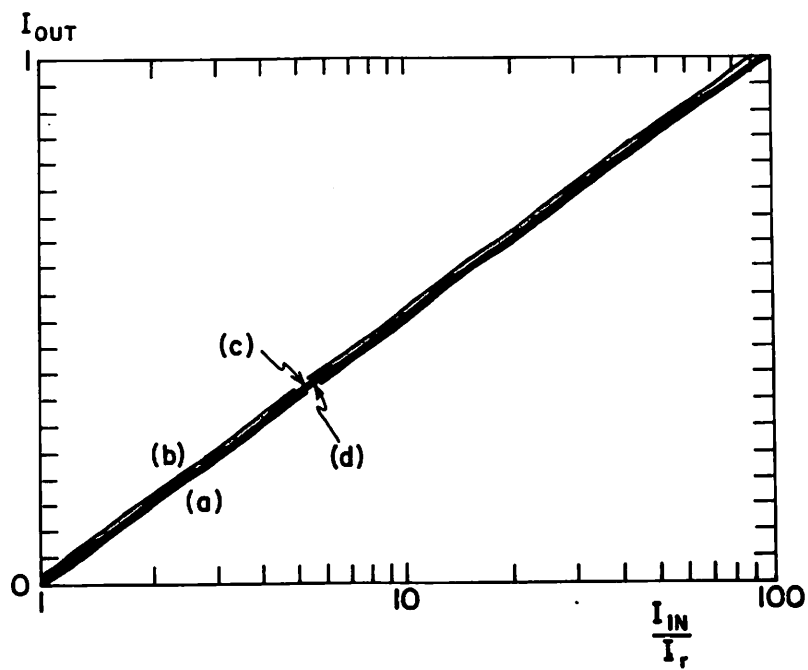
Figure 4.7   Logarithmic transfer function for a
recording medium with a linear shoulder

(a)   binary medium $\varepsilon=0$;   (b) $\varepsilon=1.0$
(c) $\varepsilon=0.33$;   (d) $\varepsilon=0.20$

$$I_o(I_{in}, k) = \left| \frac{1}{L} \int_0^L g[\log I_{in} - f(x)] e^{j\frac{2\pi kx}{L}} dx \right|^2 \qquad (4.8)$$

where k is the order selected and the rest of the parameters are explained in section 4.1.1. To find the degradation due to each part of the characteristic curve of the recording medium, we can go through the same sort of procedures as in the zero-order case. The specific example we consider here is the level slice function. It turns out in this case that the calculations simplify and the different parts of the recording medium characteristic curve do not have to be considered separately. This is shown in the following sections.

4.2.1 Degradation of the Output Intensity Due to the Nonbinary Characteristic of the Recording Medium for a Level-Slice Process

The level slice function is a nonmonotonic nonlinearity and hence we have to use a nonzero order to obtain this function. Since it has only one sign change in its slope the first order should be used. We want to have

$$I_o(I_{in}, 1) = \begin{cases} 0 & \text{for } I_{in} < I_a \\ c^2 & \text{for } I_a \leq I_{in} < I_b \\ 0 & \text{for } I_b \leq I_{in}. \end{cases} \qquad (4.9)$$

For simplicity we choose $c^2 = \frac{1}{\pi^2}$. The form of this function

48

is shown in Fig. 4.8. With an ideal recording medium shown in Fig. 3.3 and a=0, b=1, the halftone screen density from Fig. 3.6 is

$$f(x) = \begin{cases} \log(I_a/I_r) & 0 \le x < \frac{L}{2} \\ \\ \log(I_b/I_r) & \frac{L}{2} \le x \le L \end{cases} \tag{4.10}$$

as shown in Fig. 4.9. From Eq. (4.8) the output intensity for a general recording medium in the first order is

$$I_o(I_{in},1) = \left| \frac{1}{L} \int_0^L g[\log I_{in}-f(x)]e^{j\frac{2\pi x}{L}} dx \right|^2 \tag{4.11}$$

and substituting for f(x) from Eq. (4.10) we have

$$I_o(I_{in},1) = \left| \frac{1}{L}\left\{ \int_0^{L/2} g[\log I_{in}-\log(I_a/I_r)]e^{j\frac{2\pi x}{L}} dx + \right. \right.$$

$$\left. \left. \int_{L/2}^L g[\log I_{in}-\log(I_b/I_r)]e^{j\frac{2\pi x}{L}} \right\} \right|^2 \tag{4.12}$$

or

$$I_o(I_{in},1) = \frac{1}{\pi^2}\{g[\log I_{in}-\log(I_a/I_r)]$$
$$-g[\log I_{in}-\log(I_b/I_r)] \}^2. \tag{4.13}$$

From this equation we can say that the output intensity is obtained by shifting the recording medium characteristic curve by amounts $\log(I_a/I_r)$ and $\log(I_b/I_r)$ (determined from the level slice characteristic), subtracting these shifted curves from each others, and then squaring the result. This shows how a perfect level slice function is obtained
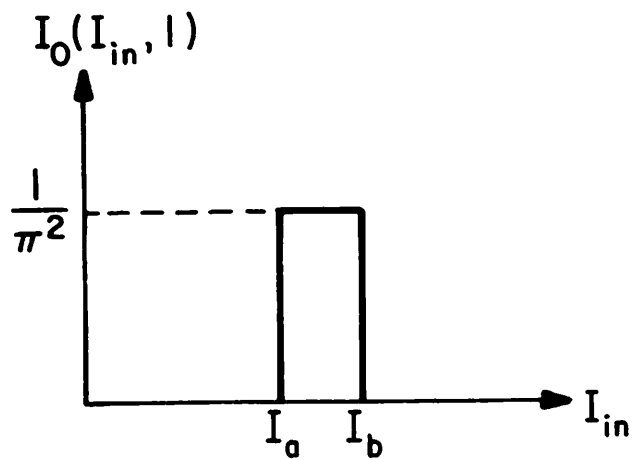
Figure 4.8   Level slice function
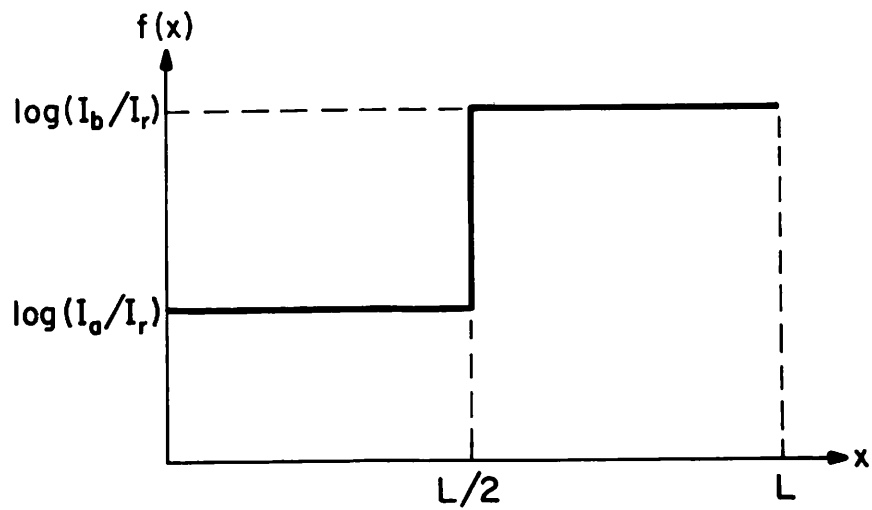
Figure 4.9   Halftone screen density for the
level slice function of Figure 4.8

when we have a binary recording medium with a characteristic curve shown in Fig. 3.3. On the other hand, for a nonbinary recording medium with a characteristic curve like Fig. 4.1, the toe and shoulder of this curve makes the sharp corners of the level slice function rounded and the linear part of this curve make the rise and fall of the level slice function linear rather than abrupt.

CHAPTER 5

PIECEWISE LINEAR RECORDING MEDIUM

In the previous chapter we saw how the nonbinary characteristic of the recording medium affects the nonlinear process. We conclude that for smooth nonlinearities the linear part of the recording medium characteristic is the main source of degradation for the output of the process.

In this chapter, we design halftone screen density profiles that compensate for the nonbinary characteristic of the recording medium. This means that we want to find $f(x)$ from Eq. (3.5) once $I_o(I_{in}, k)$, the form of $g(\cdot)$, and k are specified. This is a nonlinear integral · equation problem which may not have a general solution. Batten and Everett [5-1] in their study of the control of film characteristics have obtained closed form solutions to a similar integral equation for certain types of nonlinear functions. Unfortunately their method cannot be applied to our formulas for the important cases of logarithmic and exponential processes. This occurs because the integral on the right-hand side of Eq. (3.5) is raised to a power of two while in their equation, they only have a simple

integral. So, the solutions that they have obtained for the logarithmic and exponential processes are not applicable in our formulation.

To simplify somewhat the above-mentioned integral equation, we assume that the recording medium has a piecewise linear characteristic as shown in Fig. 4.4. This assumption, as mentioned above, is a good one for monotonic nonlinearities. Noting that monotonic nonlinearities are possible in the zero order, we then have k=0 in Eq. (3.5) leading to added simplification. With this procedure we get an integral equation which is easier to work with at the expense of limiting ourselves to monotonic nonlinearities.

In the followings we attempt to find f(x) from Eq.(3.5) with the above assumptions. The results that we obtain work only for monotonic nonlinear functions. Compensation methods for other forms of nonlinearities will be discussed in the next chapter.

5.1 Halftone Screen Density Function for Monotonic Nonlinearities Using a Piecewise Linear Recording Medium

The piecewise linear recording medium characteristic curve is shown in Fig. 4.4. From this figure we can write

$$g(\cdot) = a, \quad \text{when } \log I_{in} - f(x) < \log I_1 \qquad (5.1)$$

$$g(\cdot) = \frac{1}{\log(I_2/I_1)}\{(b-a)[\log I_{in} - f(x)] + a\log I_2 - b\log I_1\},$$

$$(5.2)$$

$$\text{when } \log I_1 \leq \log I_{in} - f(x) < \log I_2$$

and

$$g(\cdot) = b, \quad \text{when } \log I_2 \leq \log I_{in} - f(x). \quad (5.3)$$

Noting that we are solving for monotonic nonlinearities, $f(x)$ can also be assumed to be a monotonic increasing or decreasing function, depending on the type of nonlinearity. Here, we assume $f(x)$ to be a monotonically increasing function which is applicable to the increasing logarithmic and exponential nonlinearities. If we let the inverse function of $f(x)$ be represented by $h(\cdot)$, then we have

$$y = \log(I_{in}/I_1) \quad (5.4a)$$

$$\varepsilon = \log(I_2/I_1) \quad (5.4b)$$

and

$$\log(I_{in}/I_2) = y - \varepsilon. \quad (5.4c)$$

Then the conditions of Eq. (5.1) through Eq. (5.3) are equivalent to

$$g(\cdot) = a, \quad \text{when } x > h(y) \quad (5.5)$$

$$g(\cdot) = \frac{1}{\varepsilon}[(a-b)f(x) + by - a(y-\varepsilon)],$$

$$(5.6)$$

$$\text{when } h(y-\varepsilon) < x \leq h(y)$$

54

and

$$g(\cdot) = b, \qquad \text{when } x \leq h(y-\varepsilon).\qquad (5.7)$$

These equations specify the characteristics of a piecewise linear recording medium. Now, rewriting the general formula in the zero order from Eq. (3.6) in the form

$$I_o(I_{in},0) = \{\frac{1}{L}\int_0^L g[logI_{in}-f(x)]dx\}^2 \qquad (5.8)$$

and using the above requirements for $g(\cdot)$ in it, we get

$$\int_0^L g[logI_{in}-f(x)]dx = \int_0^{h(y-\varepsilon)} bdx + \frac{1}{\varepsilon}\int_{h(y-\varepsilon)}^{b(y)}[(a-b)f(x)+by-a(y-\varepsilon)]dx$$

$$(5.9)$$

$$+ \int_{h(y)}^L adx$$

which after simplification becomes

$$\int_0^L g[logI_{in}-f(x)]dx = \frac{b-a}{\varepsilon}\int_{y-\varepsilon}^Y h(u)du+aL. \qquad (5.10)$$

The integral on the right-hand side of the above equation can be simplified again as described in Appendix A. The result is

$$\int_{h(y-\varepsilon)}^{h(y)} f(x)dx = yh(y)-(y-\varepsilon)h(y-\varepsilon)-\int_{y-\varepsilon}^Y h(u)du \qquad (5.11)$$

and when used in Eq. (5.10) gives

$$\int_0^L g[logI_{in}-f(x)]dx = \frac{b-a}{\varepsilon}\int_{y-\varepsilon}^Y h(u)du+aL. \qquad (5.12)$$

This combined with Eq. (5.8) reduces to

$$I_0(I_{in}, 0) = [\frac{b-a}{\epsilon} \int_{y-\epsilon}^{y} h(u) du + aL]^2. \qquad (5.13)$$

Taking the square root of both sides of Eq. (5.13) gives

$$+\sqrt{I_0(I_{in}, 0)} = \frac{b-a}{\epsilon} \int_{y-\epsilon}^{y} h(u) du + aL. \qquad (5.14)$$

The positive sign has been chosen for the square root since both y and $h(\cdot)$ take on nonnegative values. The right-hand side of Eq. (5.14) is a function of I_{in}, which in turn is a function of y through Eq. (5.4). This dependence can be shown through

$$z(y) = +\sqrt{I_0(I_{in}, 0)} \qquad (5.15)$$

and when used in Eq. (5.14) results in

$$z(y) = \frac{b-a}{\epsilon} \int_{y-\epsilon}^{y} h(u) du + aL. \qquad (5.16)$$

To remove the integral, we can differentiate both sides of this equation with respect to y to get

$$z'(y) = \frac{b-a}{\epsilon} [h(y) - h(y-\epsilon)]. \qquad (5.17)$$

This is a difference equation in terms of $h(\cdot)$. One method to solve such an equation is to use the Laplace transform [5-2]. Considering the definition of y in Eq. (5.4), we

56

see that for a physically realizable problem this parameter is nonnegative. So, we can take the Laplace transform of both sides of Eq. (5.17) to obtain

$$Z'(s) = \frac{b-a}{\varepsilon}[H(s) - H(s)e^{-\varepsilon s}] \tag{5.18}$$

where

$$Z'(s) = \mathcal{L}[z'(y)]$$

$$\tag{5.19}$$

$$H(s) = \mathcal{L}[h(y)].$$

Equation. (5.17) can be written as

$$H(s) = \frac{\varepsilon}{b-a} \cdot \frac{Z'(s)}{1-e^{-\varepsilon s}} \cdot \tag{5.20}$$

This equation gives the Laplace transform of the function $h(\cdot)$. Before transforming back to obtain $h(\cdot)$, note that

$$|e^{-\varepsilon s}| < 1 \tag{5.21}$$

hence

$$\frac{1}{1-e^{-\varepsilon s}} = 1 + e^{-\varepsilon s} + e^{-2\varepsilon s} + e^{-3\varepsilon s} + \ldots \tag{5.22}$$

which permits Eq. (5.20) to be written as

$$H(s) = \frac{\varepsilon}{b-a}[Z'(s) + Z'(s)e^{-\varepsilon s} + \ldots]. \tag{5.23}$$

Now taking the inverse Laplace transform of both sides of Eq. (5.23) gives

$$h(y) = \frac{\varepsilon}{b-a}[z'(y) + z'(y-\varepsilon) + z'(y-2\varepsilon) + \ldots] \tag{5.24}$$

which is easily seen to satisfy Eq. (5.16) except for a constant. When we include this constant, the complete solution is

$$h(y) = \frac{\varepsilon}{b-a}[z'(y)+z'(y-\varepsilon)+\ldots]+\frac{aL}{b-a} \qquad (5.25)$$

which expresses the inverse function of the halftone screen density in terms of the derivative of the nonlinear function that we want to generate. To get a solution from this formula, the nonlinear function should be differentiable, a condition which is true for the usual smooth monotonic nonlinearities in practice. On the other hand, the solution for $h(\cdot)$ should be monotonic and hence invertible to give the density function. In the following sections we consider specific examples and find out when a satisfactory $h(\cdot)$ exists.

5.2 Examples

In this section we consider the power, logarithmic, and exponential transformations separately. In each case, the desired input-output relationship will be used in the results of the previous section to obtain the proper halftone screen density function.

5.2.1 Power Transformation

For this transformation we want

$$I_o(I_{in},0) = \gamma(\frac{I_{in}}{I_1})^\lambda \qquad (5.26)$$

where γ and λ are constants and I_1 is defined in Fig. 4.4. Replacing for $\dfrac{I_{in}}{I_1}$ in Eq. (5.26) from Eq. (5.4) we get

$$I_o(I_{in}, 0) = \gamma \cdot 10^{\lambda y} \qquad (5.27)$$

and combining with Eq. (5.15) gives

$$z(y) = +\gamma^{1/2} \cdot 10^{\frac{1}{2}\lambda y} \qquad (5.28)$$

From this

$$z'(y) = K_1 \cdot 10^{\frac{1}{2}\lambda y} \qquad (5.29)$$

where

$$K_1 = \frac{\lambda}{2} \cdot \gamma^{1/2} \cdot \ell n 10. \qquad (5.30)$$

Now we use Eq. (5.29) in Eq. (5.25) to obtain

$$h(y) = \frac{\varepsilon K_1}{b-a}[10^{\frac{1}{2}\lambda y} + 10^{\frac{1}{2}\lambda y(y-\varepsilon)} + \ldots] + \frac{aL}{b-a} \qquad (5.31)$$

$$= \frac{\varepsilon K_1 \cdot 10^{\frac{1}{2}\lambda y}}{b-a}[1 + e^{-\frac{1}{2}\lambda \varepsilon} + e^{-\frac{1}{2}\lambda \cdot 2\varepsilon} + \ldots]. \qquad (5.32)$$

Noting that the term inside the bracket is a geometric series which sums to

$$\frac{1}{1 - e^{-\frac{1}{2}\lambda \varepsilon}}$$

we can write

$$h(y) = \frac{1}{b-a} \left[\frac{\epsilon K_1 e^{\frac{1}{2}\lambda y}}{1-e^{-\frac{1}{2}\lambda\epsilon}} + aL \right] . \qquad (5.33)$$

This can easily be inverted to obtain $f(x)$ as

$$f(x) = \frac{2}{\lambda} \left[\ln(x - \frac{aL}{b-a}) + \frac{b-a}{\epsilon K_1}(1-e^{-\frac{1}{2}\lambda\epsilon}) \right]. \qquad (5.34)$$

5.2.2 Logarithmic Transformation

For this transformation we want

$$I_0(I_{in}, 0) = K \cdot \log(I_{in}/I_1) \qquad (5.35)$$

where K is a constant and I_1 is defined in Fig. 4.4. Replacing for $\frac{I_{in}}{I_1}$ in Eq. (5.35) from Eq. (4.4) we get

$$I_0(I_{in}, 0) = Ky \qquad (5.36)$$

and combining with Eq. (5.15) gives

$$z(y) = +(Ky)^{1/2}. \qquad (5.37)$$

From this we obtain

$$z'(y) = \frac{1}{2}K^{1/2}y^{-1/2}. \qquad (5.38)$$

When used in Eq. (5.25) this gives us

$$h(y) = \frac{K^{1/2}\epsilon}{2(b-a)}[y^{-1/2} + (y-\epsilon)^{-1/2} + \ldots] + \frac{aL}{b-a}. \qquad (5.39)$$

Noting that $z'(y)$ is a monotonically decreasing function we

see that $h(y)$ is not necessarily monotonic. Hence, in this case, we cannot invert $h(y)$ to obtain $f(x)$, although it will be obtained in the next chapter by another method.

5.2.3 Exponential Transformation

For this transformation we want

$$I_o(I_{in}, 0) = \alpha(\beta)^{\frac{I_{in}}{I_1}} \tag{5.40}$$

where α and β are constants and I_1 is defined in Fig. 4.4. Replacing for $\frac{I_{in}}{I_1}$ in Eq. (5.40) from Eq. (5.4) we have

$$I_o(I_{in}, 0) = \alpha(\beta)^{10^Y}. \tag{5.41}$$

and combining with Eq. (5.15) gives

$$z(y) = +\alpha^{1/2}(\beta)^{1/2 \cdot 10^Y} \tag{5.42}$$

From Eq. (5.42) we have

$$z'(y) = K_2 \cdot 10^Y \cdot \beta^{10^Y} \tag{5.43}$$

where

$$K_2 = \frac{1}{2} \alpha^{1/2} \cdot \ln 10 \cdot \ln \beta. \tag{5.44}$$

Now we use above in Eq. (5.25) to obtain

$$h(y) = \frac{K_2 \varepsilon}{b-a} 10^Y [\beta^{10^Y} + 10^{-\varepsilon} \cdot \beta^{10^{Y-\varepsilon}} + \ldots] + \frac{aL}{b-a}. \tag{5.45}$$

61

It can be seen that $z'(y)$ is a monotonically increasing function which would make $h(y)$ a monotonically increasing function also. Although we cannot invert $h(y)$ analytically, it is possible to do so by numerical means.

5.3 Comments

Equation (5.25) gives the inverse function of the halftone screen density function for a piecewise linear recording medium. To obtain the density function we should be able to invert the function obtained from this equation. The logarithmic process example shows that this inversion is not possible in general. This, of course does not mean that the compensation problem for monotonic nonlinearities should be restricted to a certain class of such functions. The fact is that too many restrictions are placed on our problem to make it manageable with the mathematical formulation presented here. In the next chapter we consider the compensation problem in general and try to find new methods of compensating for the nonbinary characteristic of the recording medium for both monotonic and nonmonotonic nonlinearities.

CHAPTER 6

DISCRETE DENSITY HALFTONE SCREEN

Several methods are available in practice for generating halftone screens with a desired density profile. Some methods are purely optical and involve the photographic recording of geometrical shadows or diffraction patterns from ruled gratings. Although this technique produces continuous halftone screens, it does not have the flexibility to precisely produce arbitrary screens needed for nonlinear processing.

Another type of halftone screen density profile that can be generated in practice is a step function approximation to the desired continuous density. These halftone screens are generated by digital image recorders, plotting microdensitometers or step and repeat cameras. Hence the theoretical accuracy available in designing the halftone screen density profile is limited by the practical limitations in making the screen. This motivates the following analysis which considers the halftone screen density profile as a step function approximation with discrete values. As will be shown in this chapter, this assumption helps to simplify the formulas and allows us to

obtain some conclusions which cannot be drawn from the continuous density analysis. To do so, we formulate with a discrete density halftone screen and derive optimization formulas for the zero.and nonzero orders.

6.1 Zero Order

The general formula relating the output intensity in the zero order to the input intensity is given by

$$I_o(I_{in},0) = \{\frac{1}{L} \int_0^L g[\log I_{in} - f(x)]dx\}^2 \qquad (6.1)$$

as expressed in Eq. (3.6). Let f(x) be approximated as shown in Fig. 6.1, then Eq. (6.1) can be written as

$$I_o(I_{in},0) = \{\frac{1}{L}[\int_0^{L_1} g(\log I_{in} - a_1)dx + \int_{L_1}^{L_2} g(\log I_{in} - a_2)dx +$$

$$(6.2)$$

$$\dots + \int_{L_{N-1}}^L g(\log I_{in} - a_N)dx]\}^2 .$$

Assuming that we have equal dimensions on the x-axis, i.e.

$$L_1 - 0 = L_2 - L_1 = \dots = L_N - L_{N-1} = \frac{L}{N} \qquad (6.3)$$

then

$$I_o(I_{in},0) = \{\frac{1}{L}[(L_1-0)g(\log I_{in} - a_1) + (L_2-L_1)g(\log I_{in} - a_2) +$$

$$(6.4)$$

$$\dots + (L_N - L_{N-1})g(\log I_{in} - a_N)]\}^2$$

or
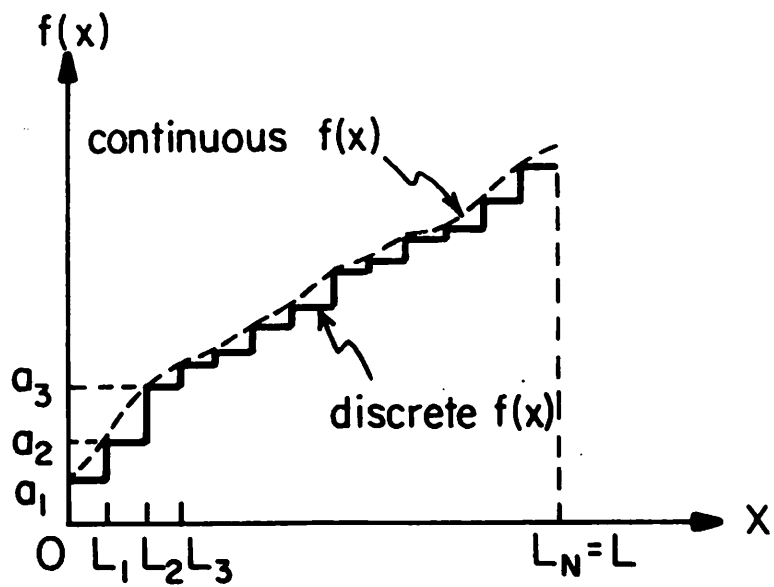
Figure 6.1   Step approximation to halftone
screen density profile

$$I_O(I_{in}, 0) = \{\frac{1}{L} \sum_{i=1}^{N} g(\log I_{in} - a_N)\}^2. \qquad (6.5)$$

The above formula gives the output intensity in the zero order as a function of the discrete grey levels on the halftone screen and the characteristic curve of the recording medium.

6.1.1 Synthesis Problem (Zero Order)

The synthesis problem here is to find a_i, $i=1,N$ in Eq. (6.5) when $g(\cdot)$ and the functional relationship of $I_O(I_{in}, 0)$ in terms of I_{in} is given. Since Eq. (6.5) holds over the input intensity range, the functional relationship of $I_O(I_{in}, 0)$ in terms of I_{in} holds for discrete values of I_{in}, say $I_{in} = I_k$, where $k=1,K$. Thus in discrete form Eq. (6.5) can be written as

$$I_O(I_1, 0) = \{\frac{1}{N} \sum_{i=1}^{N} g(\log I_1 - a_i)\}^2$$
$$\vdots$$
$$I_O(I_k, 0) = \{\frac{1}{N} \sum_{i=1}^{N} g(\log I_k - a_i)\}^2 \qquad (6.6)$$
$$\vdots$$
$$I_O(I_K, 0) = \{\frac{1}{N} \sum_{i=1}^{N} g(\log I_K - a_i)\}^2.$$

One procedure to find the optimum a_i's is to minimize the mean square error expression

$$\min_{a_1,a_2,\ldots,a_N} E = \sum_{k=1}^{K} \{I_o(I_k,0) - [\frac{1}{N}\sum_{i=1}^{N} g(\log I_k - a_i)]^2\}^2. \quad (6.7)$$

This should produce values of a_i's that bring the output intensity in the zero order as close as possible to the desired output intensity in the mean-square sense.

Now, since $a_1,\ldots,a_N$ are the different density values on the halftone screen, their values should be positive and remain between certain practical limits, as expressed by

$$c \leq a_i \leq d \qquad (6.8)$$

where c and d are given constants. This makes the minimization problem a constrained minimization. There is a transformation that simplifies this problem and imbeds the constraint in the expression to be minimized [6-1]. If we let

$$a_i = (d-c)\sin^2 y_i + c \qquad (6.9)$$

this makes the values of a_i's limited to the range [c,d]. Note that

$$\frac{\partial E}{\partial y_i} = \frac{\partial E}{\partial a_i}\frac{\partial a_i}{\partial y_i} \qquad i = 1,N. \qquad (6.10)$$

When we minimize E with respect to y , we set

$$\frac{\partial E}{\partial y_i} = 0 \qquad i = 1,N \qquad (6.11)$$

and combining Eq. (6.10) and (6.11) we have

$$\frac{\partial E}{\partial a_i} = 0 \qquad i = 1, N \qquad (6.12)$$

or

$$\frac{\partial a_i}{\partial y_i} = 0 \qquad i = 1, N. \qquad (6.13)$$

To perform the minimization we want Eq. (6.12) to be satisfied at the same time that Eq. (6.13) is not satisfied in the range [c,d]. To check for this condition note that

$$\frac{\partial a_i}{\partial y_i} = (d-c)\sin 2y_i \qquad (6.14)$$

which when set to zero gives

$$y_i = 0 \implies a_i = c$$
$$\text{or} \qquad\qquad\qquad\qquad\qquad i = 1, N \qquad (6.15)$$
$$y_i = \frac{\pi}{2} \implies a_i = d \quad .$$

It can be seen that the only values of a_i's in the range [c,d] that make $\frac{\partial a_i}{\partial y_i} = 0$ are the boundaries of this interval, and this can be prevented from causing any problem by choosing c and d, such that the interval [c,d] contains the limit values for the a_i's.

6.1.2 Example of A Function Possible with the Zero Order: Logarithmic Screen

In a logarithmic process we want the relationship

$$I_o(I_{in}, 0) = m\log(I_{in}/I_r) \qquad (6.16)$$

between the output and input intensities where m is a constant and I_r is shown in Fig. 3.3. The corresponding screen density profile for a binary recording medium shown in Fig. 3.3 is obtained from Eq. (3.19) as

$$f(x) = \frac{1}{m}[a+\frac{b-a}{L}x]^2, \quad 0 \le x \le L. \qquad (6.17)$$

Hence the initial values for the a_i's are

$$a_i = \frac{1}{m}[a+\frac{b-a}{L}x_i]^2, \quad 0 \le x_i \le L, \quad i = 1,N. \qquad (6.18)$$

From Eq. (6.9) the corresponding y_i's are

$$y_i = \sin^{-1}\sqrt{\frac{1}{d-c}\frac{1}{m}[a+\frac{b-a}{L}x_i]^2-c} . \qquad (6.19)$$

Hence to find the optimum screen giving a logarithmic relationship as in Eq. (6.16) with a recording medium characteristic curve g(logE), we must perform the minimization

$$\min_{y_1,y_2,\dots,y_N} E = \sum_{k=1}^{N}\{m\log(I_k/I_r)-[\frac{1}{N}\sum_{i=1}^{N}g(\log I_k-(d-c)\sin^2 y_i-c)]^2\}^2 \qquad (6.20)$$

with the initial values of Eq. (6.19).

A computer routine has been generated which performs the above minimization. A subroutine which is taken from the IMSL library [6-2] and uses a quasi-Newton algorithm

69

for finding the minimum of a function of N variables [6-3] is the main part of this routine. The results for a piecewise linear recording medium (Fig. 4.4 with a=0, b=1) with different slopes in its linear part (gamma) are shown in Figs. 6.2 through 6.9. It is assumed that there are 30 discrete points in the halftone screen density profile and the density values are between 0 and 2 in these figures. It is also assumed that the values of log I_r (shown in Fig. 3.3) lie in the middle of the values of log L_1, and log I_2 (shown in Fig. 4.4). Note that the plots of the input-output curves are semi-logarithmic and hence the ideal result is a straight line. The optimized output curve is seen to approximate the ideal result with a small error.

To examine how the discretization affects the degradation of the output, we have computed the degradation from Eq. (6.5) for a recording medium as above with a gamma equal to one and for different numbers of points in the halftone screen density profile. The result is shown in Fig. 6.10. Figure 4.6b is a similar result with a continuous density profile, and comparing it with Fig. 6.10 shows that for higher values of N the two graphs agree quite well.

6.2 Nonzero Order

The general formula relating the output intensity in
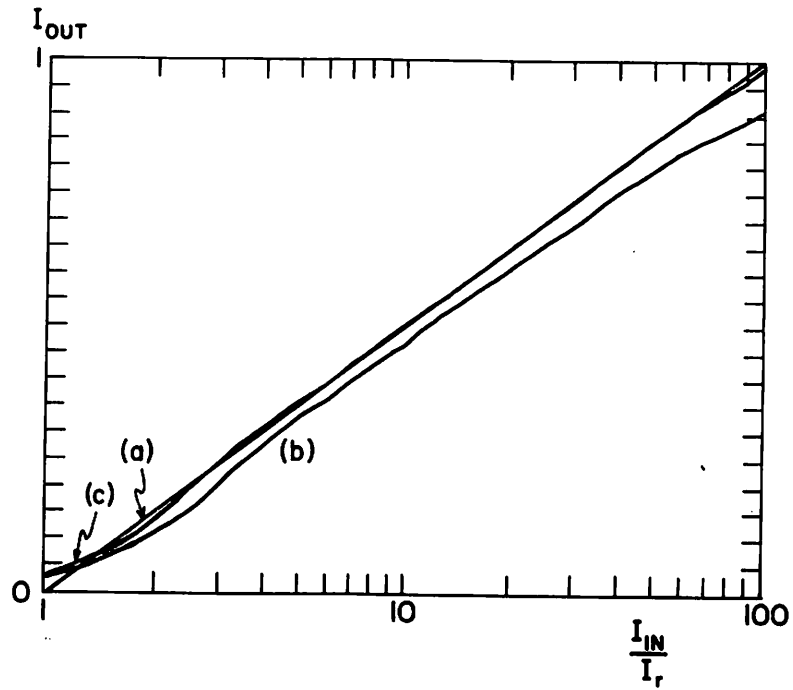
Figure 6.2 Logarithmic transfer function for a piece-
wise linear recording medium with gamma=1.0

(a) ideal; (b) degraded; (c) optimized

Figure 6.3 Halftone cell shape corresponding to Figure 6.2

(a) ideal; (b) optimized

Figure 6.4   Logarithmic transfer function for a piecewise
linear recording medium with gamma=3.0

(a) ideal;   (b) degraded;   (c) optimized

Figure 6.5   Halftone cell shape corresponding to Figure 6.4

(a) ideal;   (b) optimized
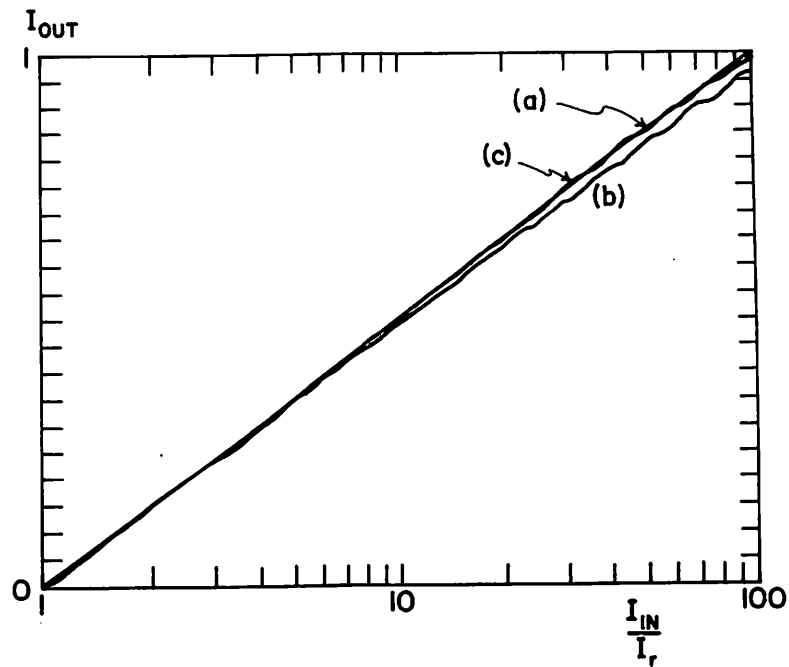
Figure 6.6 Logarithmic transfer function for a piecewise
linear recording medium with gamma=5.0

(a) ideal; (b) degraded; (c) optimized
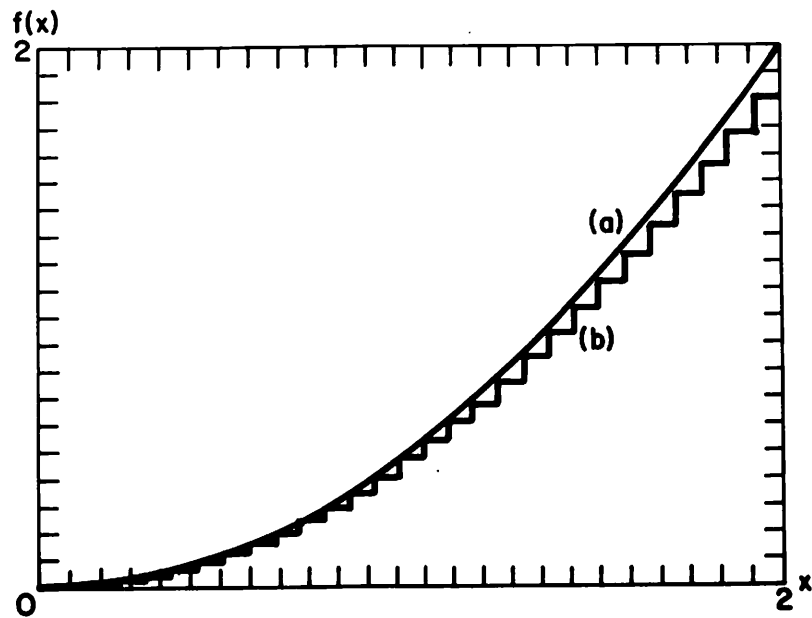
Figure 6.7 Halftone cell shape corresponding to Figure 6.6

(a) ideal; (b) optimized
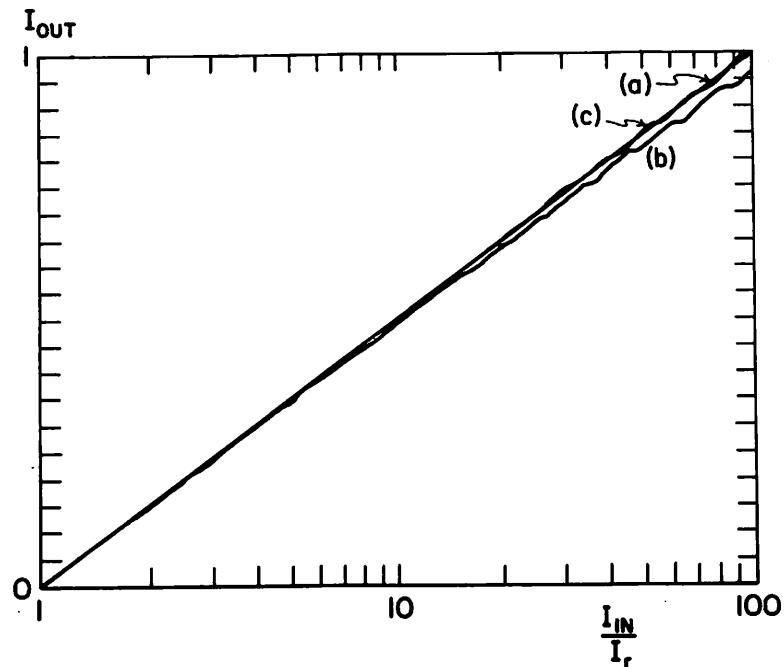
Figure 6.8 Logarithmic transfer function for a piecewise
linear recording medium with gamma=10.0

(a) ideal; (b) degraded; (c) optimized
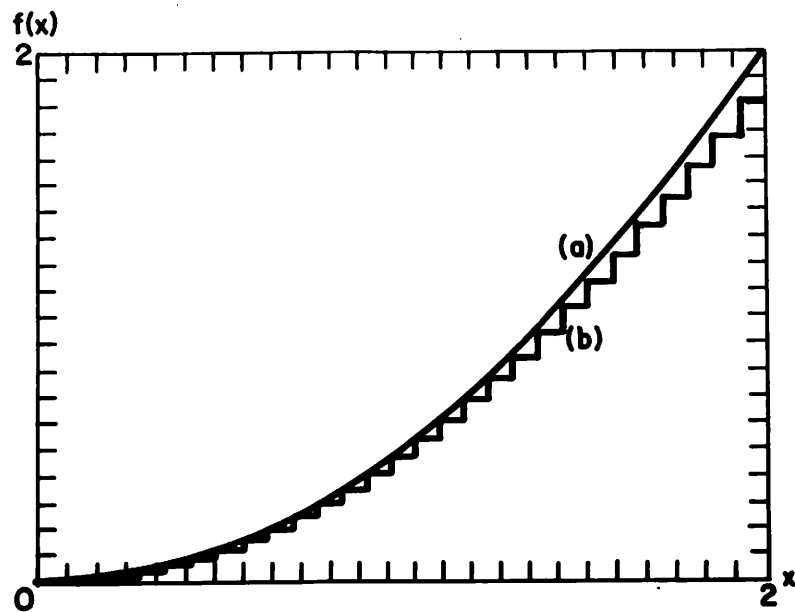
Figure 6.9 Halftone cell shape corresponding to Figure 6.8

(a) ideal; (b) optimized
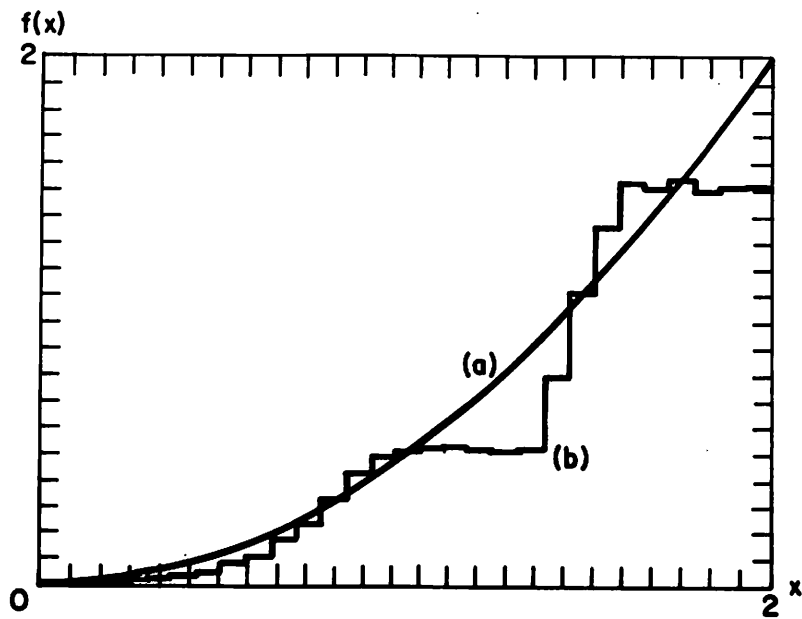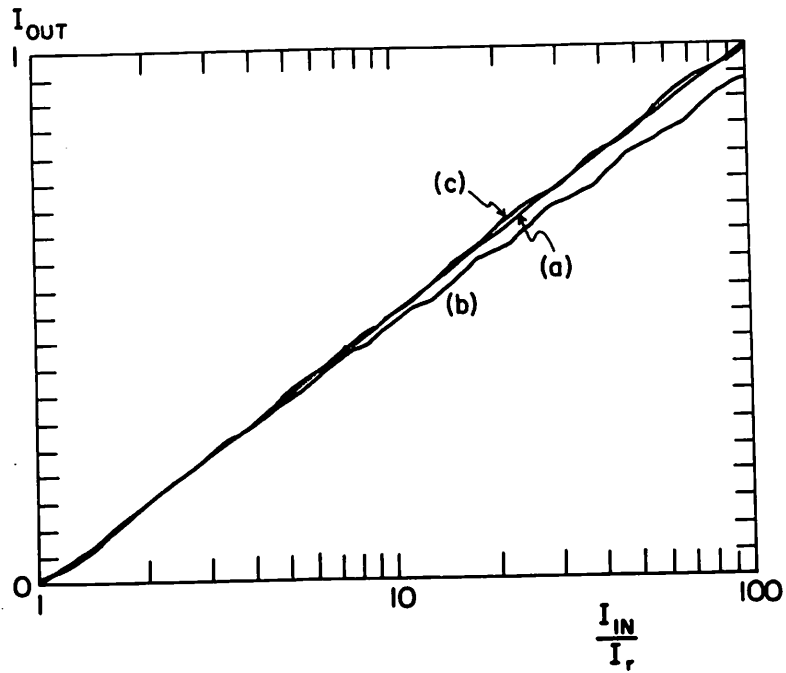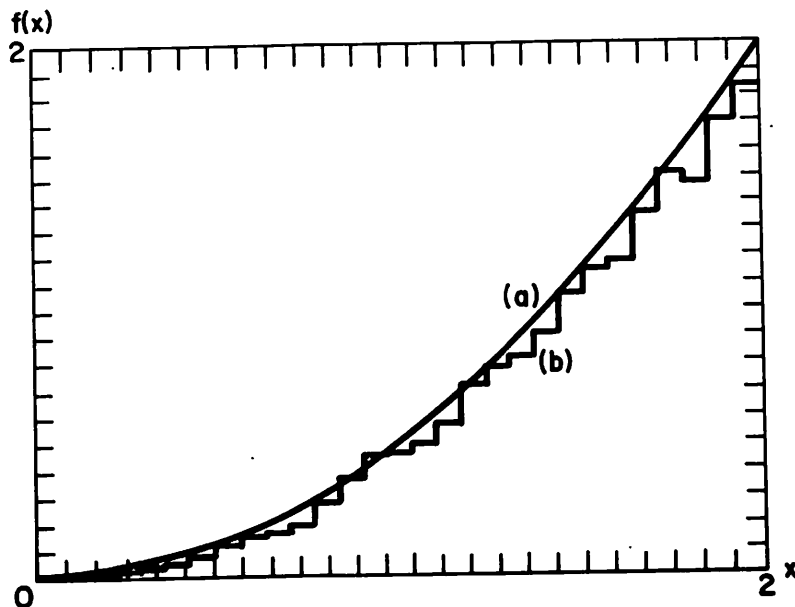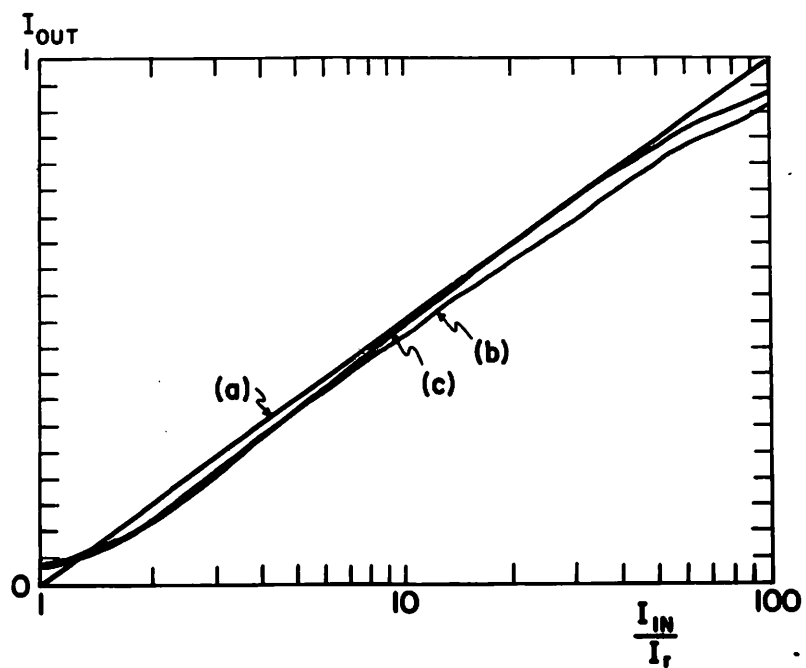
Figure 6.10 Logarithmic transfer function for a piecewise
linear recording medium with gamma=1.0

(a) ideal
(b) 16 points in the discrete density profile
(c) 200 points in the discrete density profile

the nonzero orders to the input intensity is

$$I_o(I_{in},k) = |\{\frac{1}{L} \int_0^L g[\log I_{in} - f(x)] e^{-j\frac{2\pi kx}{L}} dx\}|^2 \quad (6.26)$$

from Eq. (3.5a). Using a quantized approximation to $f(x)$, as in the zero order case we can write (from Fig. 6.1)

$$I_o(I_{in},k) = |\frac{1}{L}[\int_0^{L_1} g(\log I_{in} - a_1) e^{j\frac{2\pi kx}{L}} dx + \quad (6.27)$$

$$\int_{L_1}^{L_2} g(\log I_{in} - a_2) e^{j\frac{2\pi kx}{L}} dx + \ldots + \int_{L_{N-1}}^L g(\log I_{in} - a_n) e^{j\frac{2\pi kx}{L}} dx]|^2 .$$

For simplicity we let

$$g(\log I_{in} - a_i) = g_i, \qquad i = 1, N. \quad (6.28)$$

Then

$$I_o(I_{in},k) = \frac{1}{4\pi^2 k^2} |g_1(e^{j\frac{2\pi kL_1}{L}} - e^{j\frac{2\pi k \cdot 0}{L}} + \ldots + g_N(e^{j\frac{2\pi kL}{L}} -$$

$$e^{j\frac{2\pi kL_{N-1}}{L}})|^2$$

$$= \frac{1}{4\pi^2 k^2} |g_1 \cdot e^{j\frac{2\pi k \cdot 0}{L}} (e^{j\frac{2\pi k(L_1-0)}{L}} - 1) + \ldots + \quad (6.29)$$

$$g_N e^{j\frac{2\pi kL_{N-1}}{L}} (e^{j\frac{2\pi k(L-L_{N-1})}{L}} - 1)|^2 .$$

Now from Eq. (6.3)

$$L_1 - 0 = L_2 - L_1 = L_3 - L_2 = \ldots = L - L_{N-1} = \frac{L}{N}, \quad (6.30)$$

so when Eq. (6.30) is used in Eq. (6.29) we have

$$I_o(I_{in}, k) = \frac{1}{4\pi^2 k^2} \left| g_1 + g_2 e^{j\frac{2\pi k L_1}{L}} + \ldots + g_N e^{j\frac{2\pi k L_{N-1}}{L}} \right|^2 \cdot \left| e^{j\frac{2\pi k}{N}} - 1 \right|^2. \tag{6.31}$$

Note also that

$$L_1 = \frac{L}{N}, \quad L_2 = 2\frac{L}{N}, \quad \ldots, \quad L_i = i\frac{L}{N}, \quad i = 1, N, \tag{6.32}$$

consequently Eq. (5.31) can be written as

$$I_o(I_{in}, k) = \frac{1}{4\pi^2 k^2} \cdot \left| e^{j\frac{2\pi k}{N}} - 1 \right|^2 \cdot \left| \sum_{i=1}^{N} g(\log I_{in} - a_i) e^{j\frac{2\pi k(i-1)}{L}} \right|^2. \tag{6.33}$$

Equation (6.33) gives the output intensity in any nonzero order k in terms of the input intensity, the characteristic function of the recording medium and the discrete grey levels on the halftone screen.

6.2.1 Synthesis Problem (Nonzero Order)

The synthesis problem here, as in the case of the zero order, is to determine suitable a_i values from Eq. (6.33). The procedure that we take is the same as the one for the zero order. Namely, we minimize the expression

$$\min_{a_1, a_2, \ldots, a_N} E = \sum_{m=1}^{M} \left[I_o(I_{in}, k) - \frac{1}{4\pi^2 k^2} \left| e^{j\frac{2\pi k}{N}} - 1 \right|^2 \cdot \left| \sum_{i=1}^{N} g(\log I_m - a_i) e^{j\frac{2\pi k(i-1)}{L}} \right|^2 \right]^2. \tag{6.34}$$

To limit the resulting density values as in Eq. (6.8) we transform the problem to the minimization

$$\min_{y_1,y_2,\ldots,y_N} E = \sum_{m=1}^{M} \left\{ I_o(I_{in},k) - \frac{1}{4\pi^2 k^2} \left| e^{j\frac{2\pi k}{L}} - 1 \right|^2 \cdot \left| \sum_{i=1}^{N} g[\log I_m - (d-c)\sin^2 y_i - c] e^{j\frac{2\pi k(i-1)}{L}} \right|^2 \right\}^2.$$

(6.35)

The y_i's are related to the a_i's through Eq. (6.9). To initialize the values of the y_i's for this minimization procedure, we use the density values obtained for the screen when the recording medium is binary.

6.2.2 Example of a Function Possible in the Nonzero Order: Level Slice Screen

To obtain a level slice transformation the first order is a suitable choice. For this function we want

$$I_o(I_{in},1) = \begin{cases} 0 & \text{for } I_{in} < I_a \\ \frac{1}{\pi^2} & \text{for } I_a \le I_{in} < I_b \\ 0 & \text{for } I_{in} \ge I_b \end{cases}$$

(6.36)

and this function is shown in Fig. 4.8. The density profile giving such a relationship for the binary recording medium shown in Fig. 3.3 is

$$f(x) = \begin{cases} \log(I_a/I_r) & \text{for } 0 \le x < \frac{L}{2} \\ \log(I_b/I_r) & \text{for } \frac{L}{2} \le x \le L \end{cases}$$

(6.37)

from Eq. (4.10). The corresponding density levels in the

discrete screen are

$$a_i = \begin{cases} \log(I_a/I_r) & \text{for} \quad i < \frac{N}{2} \\ \\ \log(I_b/I_r) & \text{for} \quad i \geq \frac{N}{2}. \end{cases} \qquad (6.38)$$

Using Eq. (6.38) in Eq. (6.9), we have

$$y_i = \begin{cases} \sin^{-1}\sqrt{\dfrac{\log(I_a/I_r)-c}{d-c}} & \text{for } i < \dfrac{N}{2} \\ \\ \sin^{-1}\sqrt{\dfrac{\log(I_b/I_r)-c}{d-c}} & \text{for } i \geq \dfrac{N}{2}. \end{cases} \qquad (6.39)$$

With the initial values for y_i's as above we then want to minimize

$$\min_{y_1,\ldots,y_N} E = \sum_{m=1}^{M} \{I_o(I_m,1)-\frac{1}{4\pi^2}|e^{j\frac{2\pi}{N}}-1|^2 \cdot \\ \sum_{i=1}^{N} g[\log I_m - (d-c)\sin^2 y_i - c]|^2\}^2. \qquad (6.40)$$

A computer routine similar to the one for the logarithmic process has been generated to perform the above minimization. With the same assumptions about the recording medium and the halftone screen as with the logarithmic process (section 6.1.2), the result is shown in .Figs. 6.11 through 6.18. It is seen that the optimization procedure is successful in restoring the shape of the level slice function, but cannot do much for the edges with finite slope in the degraded response. As explained in section (4.2.1), the finite slope edges are a direct consequence of the characteristic curve of the recording

Figure 6.11   Level slice transfer function for a piecewise
              linear recording medium with gamma=1.0

              (a)   ideal;   (b)   degraded;   (c)   optimized
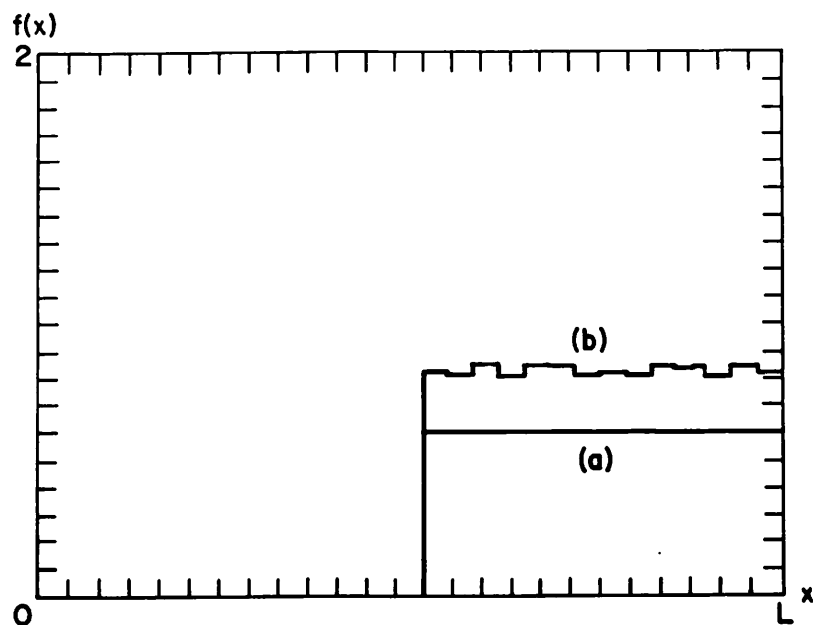
Figure 6.12   Halftone cell shape corresponding to
              Figure 6.11

              (a)   ideal;   (b)   optimized
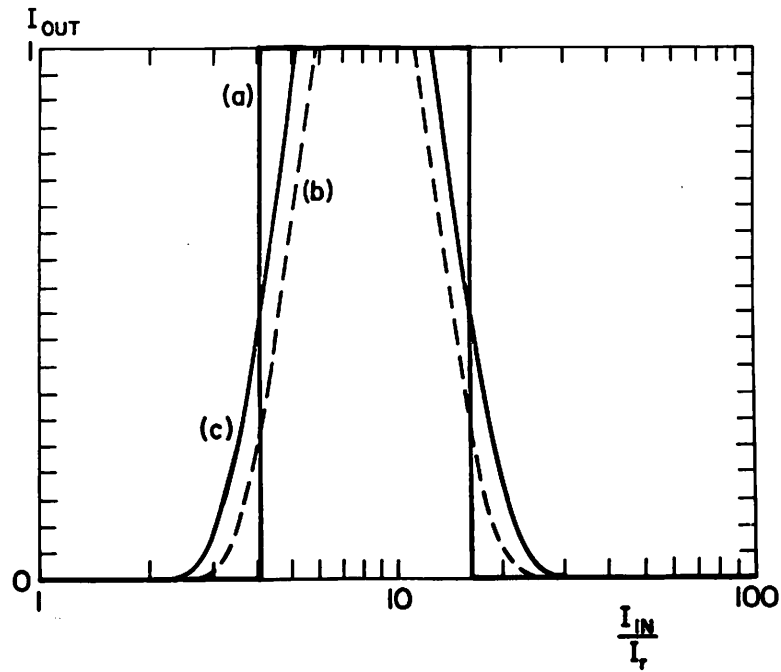
Figure 6.13   Level slice transfer function for a piecewise
linear recording medium with gamma=3.0

(a) ideal;   (b) degraded;   (c) optimized
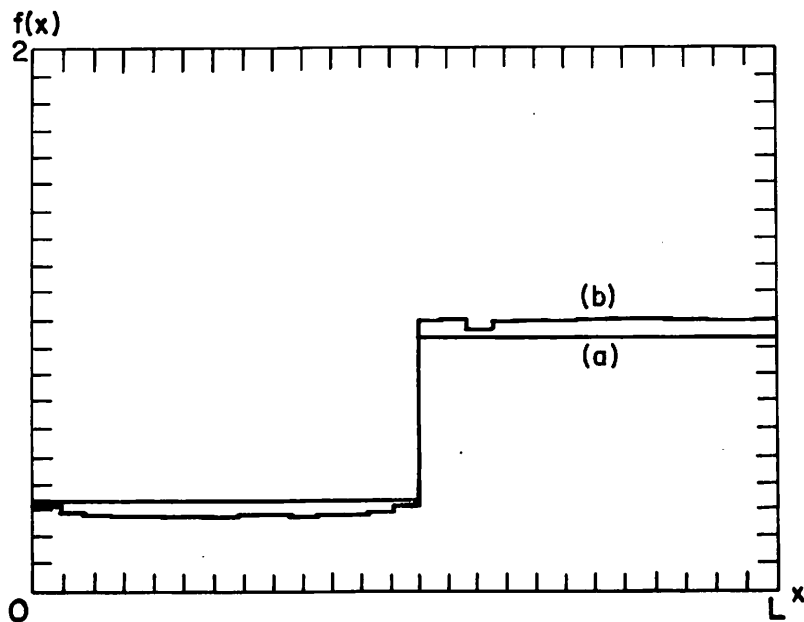
Figure 6.14   Halftone cell shape corresponding to
Figure 6.13

(a) ideal;   (b) optimized
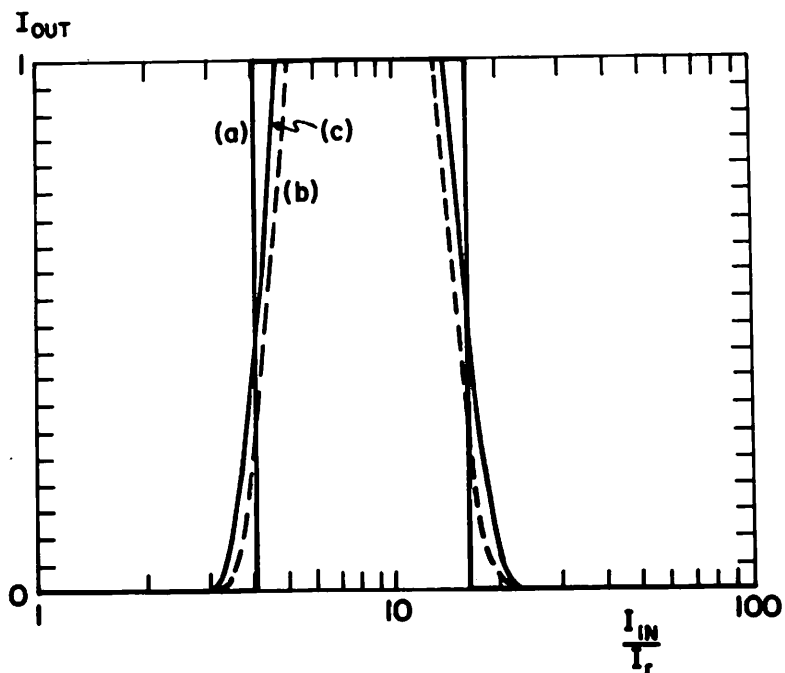
Figure 6.15 Level slice transfer function for a piecewise
linear recording medium with gamma=5.0

(a) ideal;   (b) degraded;   (c) optimized
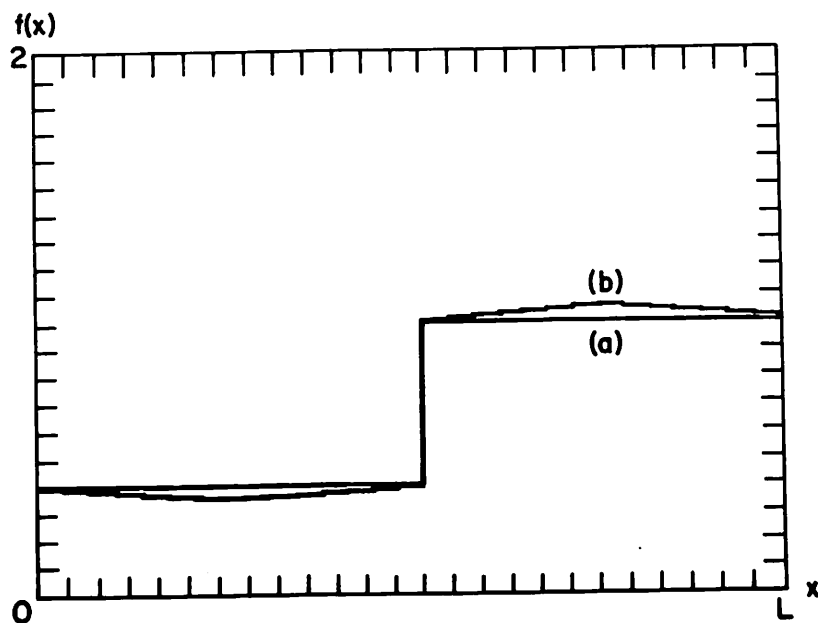
Figure 6.16 Halftone cell shape corresponding to
Figure 6.15

(a) ideal;   (b) optimized
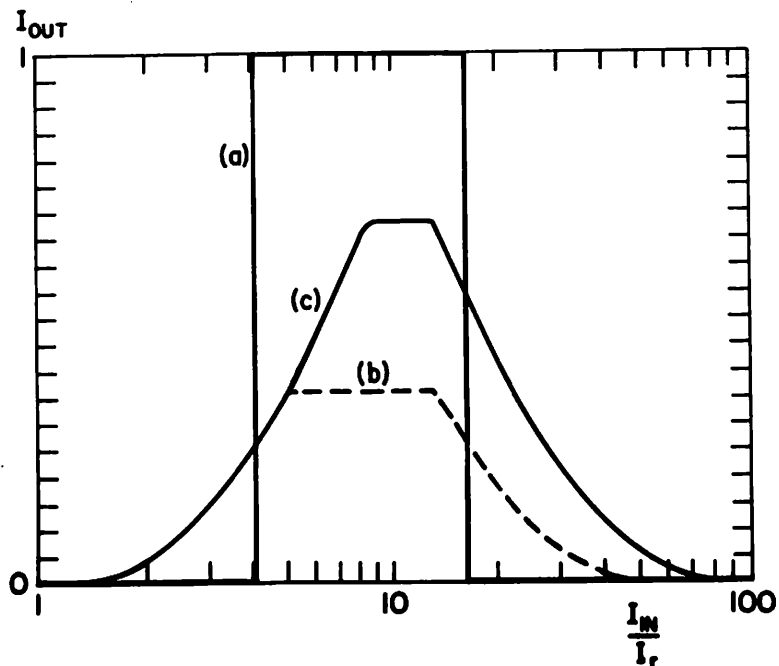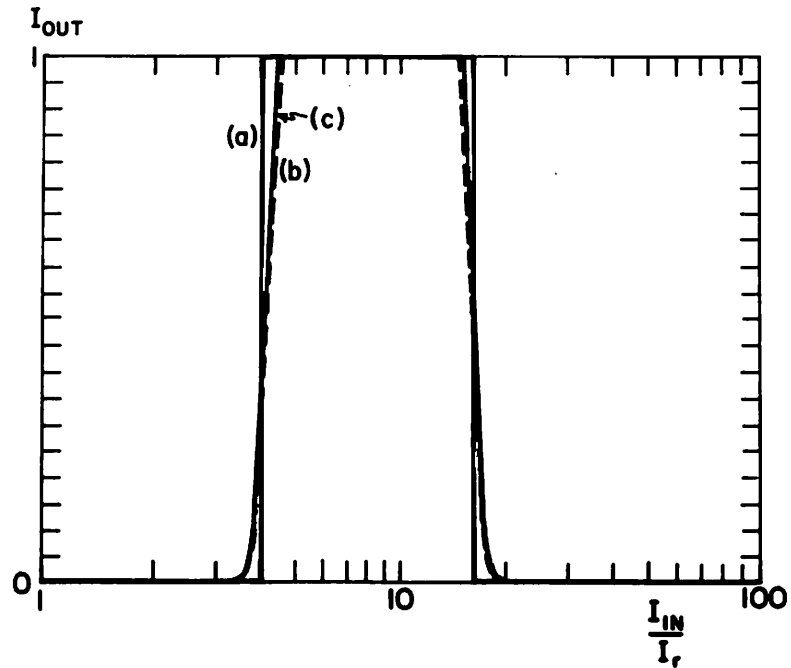
Figure 6.17   Level slice transfer function for a piecewise
linear recording medium with gamma=10.0

(a)   ideal;   (b)   degraded;   (c)   optimized
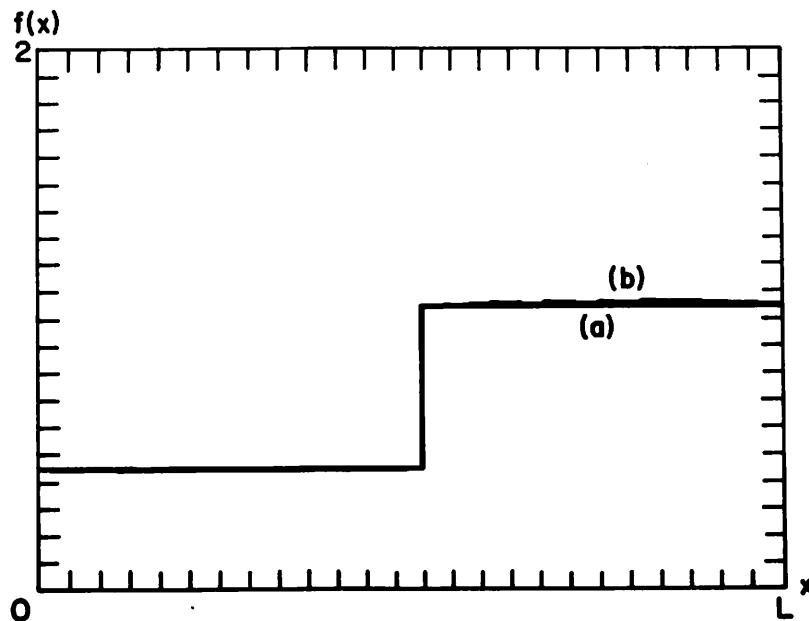
Figure 6.18   Halftone cell shape corresponding to
Figure 6.17

(a)   ideal;   (b)   optimized

medium and cannot be improved.

CHAPTER 7

EXPERIMENTS WITH HALFTONE SCREENS

7.1 Halftone Screen Fabrication

To generate a halftone screen we must plot a given density profile (such as those obtained in Chapter 6) on photographic film. To do so, various methods have been tried in previous work. Some more recent methods are: optical filtering of binary gratings by Strand [7-1], use of a plotting microdensitometer by Dashiell and Sawchuk [2-1], and moving a slit grating on top of the photographic film with repeated exposures as done by Liu, Goodman and Chan [7-2].

We have used a computer controlled Dicomed Digital Image Recorder to generate our screens. This machine can plot up to 4096x4096 discrete points in a 4"x4" area. There are 256 different intensity levels for each exposure setting. To plot any pattern on the film with this machine, an input matrix of numbers between zero and 255 representing the digitized pattern is read into the computer. Before doing so, we have to find out what is the relationship between those numbers and the final resulting density on the film. To find out this relationship a grey

level test pattern was plotted by the Dicomed on Kodak SO-115 film (the same film used for the screens). The exposed film was then developed for the recommended time in D-19 developer. The densities on the test pattern were then measured on a Macbeth transmission densitometer. This procedure was repeated for different exposure settings on the Dicomed. That results in a series of curves which represents the final density on SO-115 film versus the input numbers to the computer for each exposure setting on the Dicomed. From these curves one that covered the density range of our halftone screen was selected. This curve was then used as the basis for fabrication of that screen.

As a check on the calibration, a test pattern was generarated on a separate film with the same exposure setting of the Dicomed each time a new screen was made. If any deviations were observed, the exposure setting was re-adjusted and the experiment was repeated.

7.2 Experiments with the Logarithmic Screen

To make a logarithmic screen experimental data for the characteristic curve of the standard LCLV (45° twisted nematic hybrid field effect) was gathered. This data was then used in the optimization routines for the logarithmic process (sec. 6.1.2) to obtain the density profile of the optimized screen. The corresponding halftone screen was

later generated using the method described in section 7.1. The screen was then used in a real-time experiment to test its effectiveness.

7.2.1 Logarithmic Screen Fabrication

The experimental set up to determine the LCLV characteristic curve is shown in Fig. 7.1. With no input, the control surface of the light valve is illuminated directly by an incoherent arc lamp source. Readout was performed coherently at 632.8 nm with a simple optical filering system. In this case, no spatial filter was used in the filter plane. The resulting output intensity transmittance versus input intentisy is shown in Fig. 7.2. By taking the square root of the output intensity transmittance valves, and the logarithm of the input intensity values a plot of the amplitude transmittance vs. log exposure curve of the device was obtained. This is shown in Fig. 7.3 after normalization.

A discrete version of this graph was then fed to the computer routine for generation of the optimized density profile. The result is shown in Fig. 7.4. There are 16 discrete points in this graph. Using these values in the procedure of section 7.1, the corresponding halftone screen was made. Recalling that we can plot 4096 points over 86 mm with this method, the halftone screen has
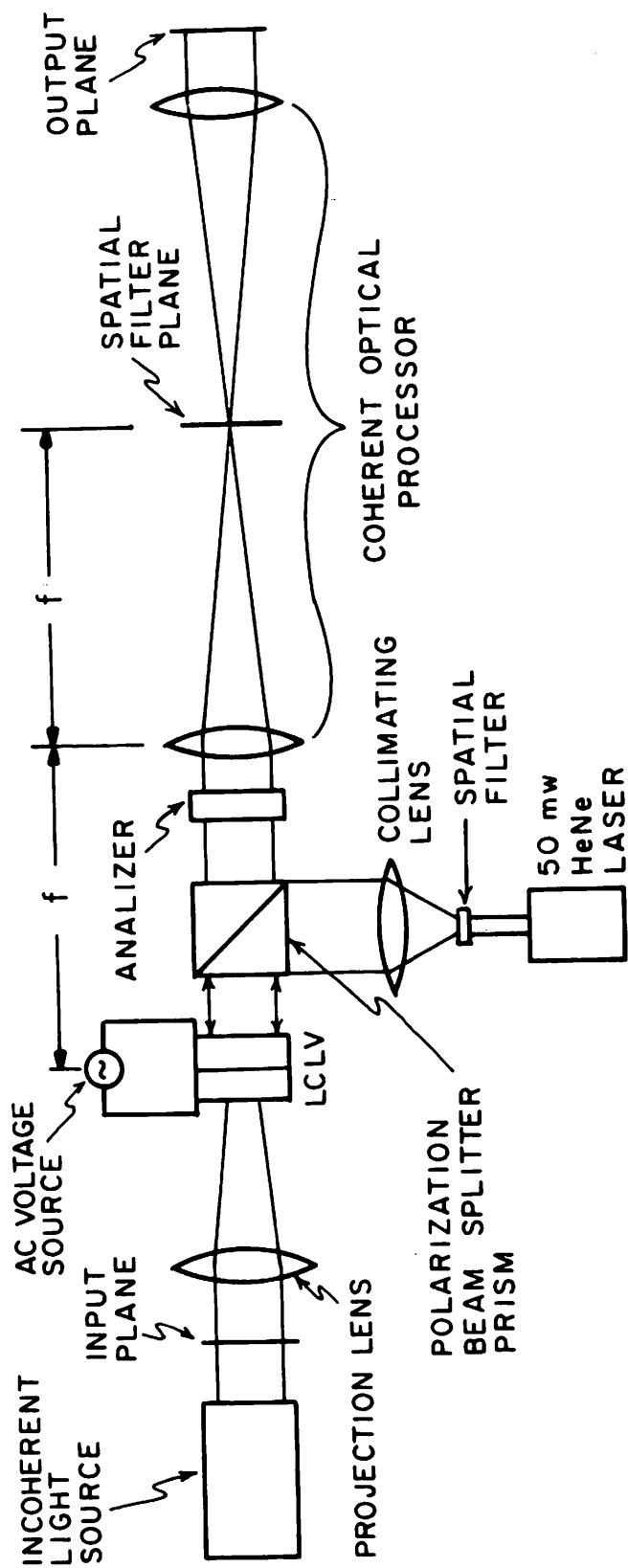
Figure 7.1 Experimental set-up for real-time parallel logarithmic
filtering
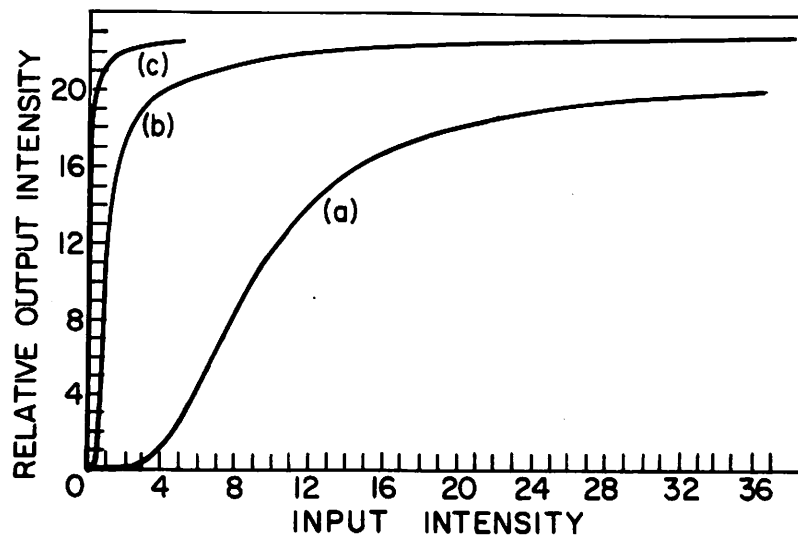
Figure 7.2   LCLV input-output transfer characteristic

    (a)   input axis x 1
    (b)   input axis x 10
    (c)   input axis x 100
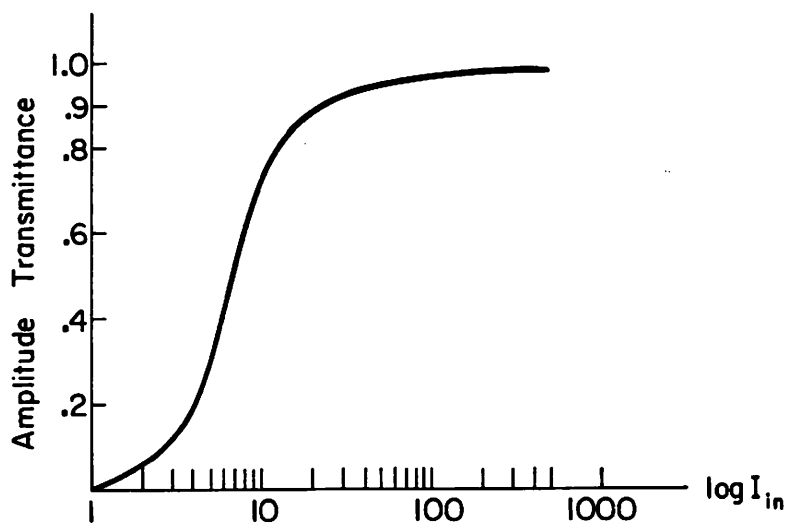
Figure 7.3   Amplitude transmittance vs log exposure
              characteristic for LCLV of Figure 7.2
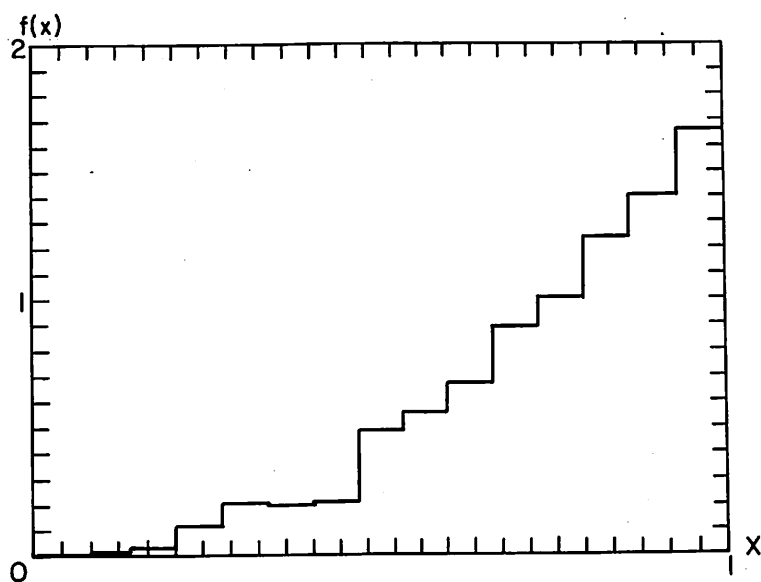
Figure 7.4   Density profile of the halftone
              screen optimized for Figure 7.3

$$\frac{4096}{16} \times \frac{1}{86 \text{ mm}} \cong 3 \text{ lines/mm}$$

This halftone screen was then placed in the input plane of the experimental setup shown in Fig. 7.1. Note that the logarithmic screen was designed on the assumption of using the zero-order of the diffraction pattern of the halftoned picture. A simple pinhole spatial filter at the position of the zero order in the filter plane was used to pick out this order. The resulting transfer function is shown in Fig. 7.5. It approximates a logarithmic transfer function with less than 5% error over one decade and less than 10% error over another decade.

7.2.2 Logarithmic Screen Performance

To test the effectiveness of the logarithmic screens in another experiment, two crossed multiplicatively combined Ronchi rulings were used as an input picture for the experimental setup of Fig. 7.1. The period of these rulings was approximately 3 mm, much higher than the halftone screen period of 0.33 mm. The spectrum in the filter plane is shown in Fig. 7.6a. Next, the logarithmic halftone screen was placed in contact with the rulings. The filter plane spectrum is shown in Fig. 7.6b. The difference in Fourier spectra between multiplicatively combined gratings and additively combined gratings obtained by real-time logarithmic filering is as follows. The additive spectrum components lie only on the x and y axis
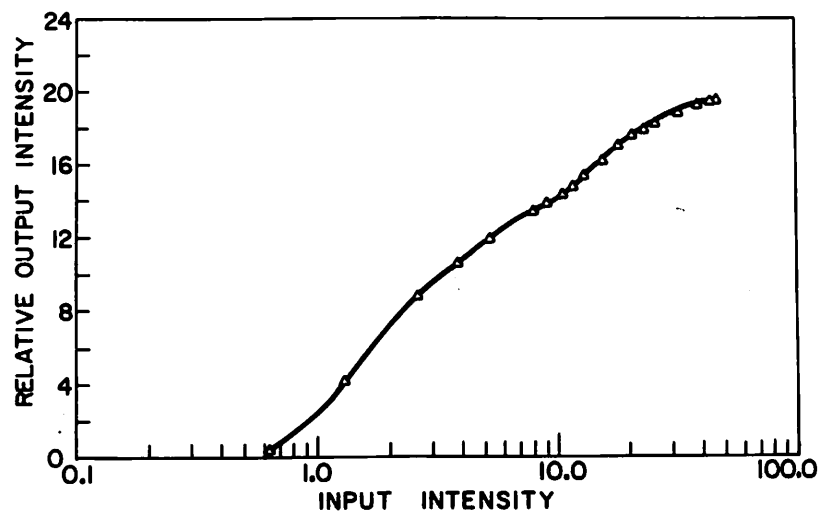
Figure 7.5   Transfer characteristic of the LCLV
with the optimized screen

(a)

(b)

Oth   1st   2nd
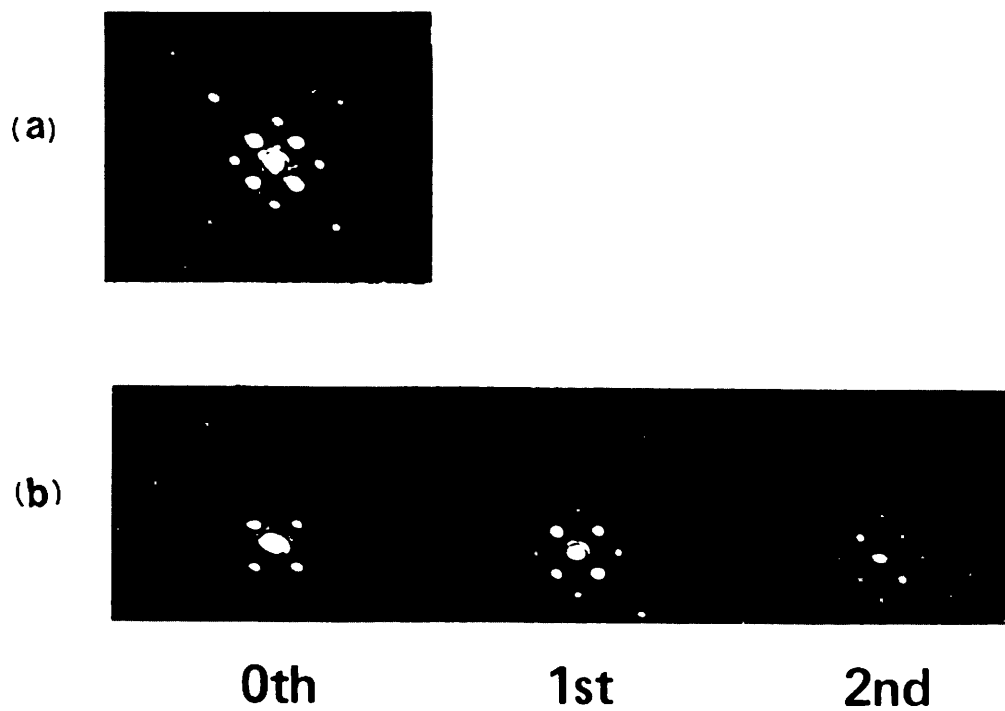
Figure 7.6   Fourier spectra of crossed gratings

      (a)   multiplicative spectrum
      (b)   additive spectrum (result in the
            Oth order)

around the zeroth diffraction order in the frequency domain while the multiplicative spectrum contains cross-term off-axis components. Figure 7.6 also shows the higher diffraction orders that arise due to halftone screen. For simple logarithmic processing, these higher orders would be eliminated by spatial filtering.

The two-dimensional point logarithm obtained here in real time is particularly useful in a signal processing technique called homomorphic filtering [7-3,7-4]. Homomorphic filtering is useful for processing multiplied or convolved signals, with specific applications in radiography [7-4,7-5], speech processing, and image processing [7-3].

CHAPTER 8

DIRECT METHOD FOR OBTAINING NONLINEAR FUNCTIONS

IN REAL TIME

Although the halftone screen method is a flexible way to obtain different nonlinear functions, it is often desirable to bypass the halftoning step and obtain the nonlinearities directly. By eliminating the halftone screen we rely on the nonlinear characteristic of the recording medium to generate an overall nonlinearity. Tai, Cheng and Yu [8-1] have used the nonlinear portion of the photographic film characteristic curve to generate a logarithmic function. Because their method uses photographic film, it cannot be implemented in real time.

In this chapter, we discuss how the nonlinear characteristic of a special birefringent liquid crystal device can be used to produce nonlinear transformations in real time. In particular, we consider implementation of the analog-to-digital conversion nonlinearity by this method.

8.1 Birefringent Liquid Crystal Device

The liquid crystal light valve described in chapter two can be used as a pure birefringent device. In this case no twist is present in the alignment of the liquid crystal with its substrates. As the input light changes the birefringence of the liquid crystal layer through the photoconductor, the output beam goes through a series of phase changes. This, in turn, is transformed into output intensity variations by the crossed polarizers. The output intensity variaion with respect to the applied voltage to the liquid crystal (no input light) is shown in Fig. 8.1, and its variation with respect to the input intensity is shown in Fig. 8.2. It is seen that the optically controlled device characteristic curve has fewer intensity peaks than the electrically controlled one. This is due to the saturation of the photoconductor material at high input light intensities.

The direct nonmonotonic nature of the device is useful for some limited approximations to obtain nonmonotonic functions. Different parts of the curve shown in Fig. 8.2 can approximate different polynomials. Each period of this curve also approximates a sine-squared function. A particularly useful function is the optical parallel analog-to-digital conversion function which is discussed in more detail in the following.
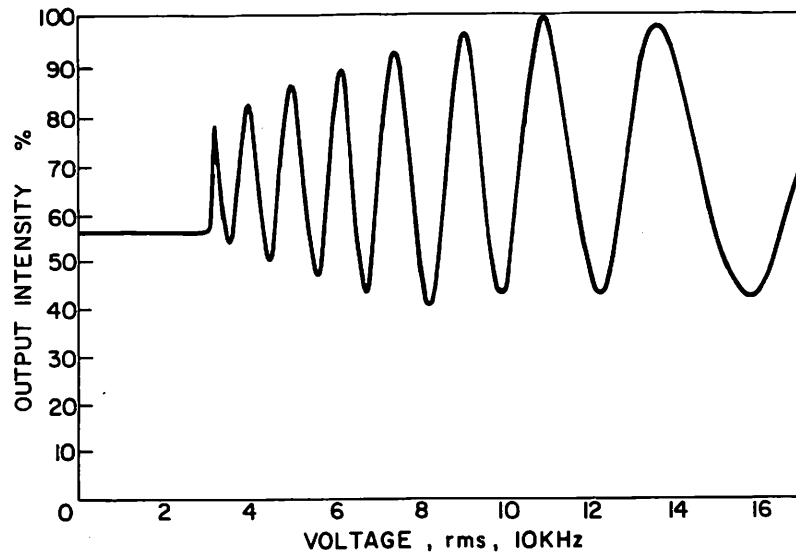
Figure 8.1   Electrical response of the pure
birefringent LC device
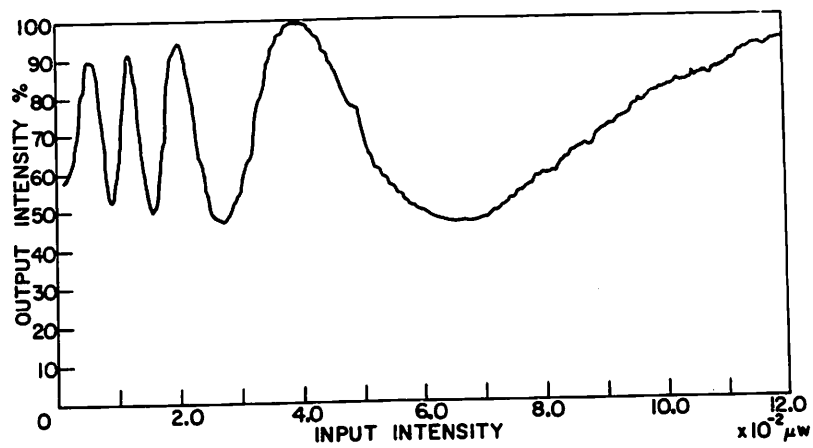
Figure 8.2   Optical response of the pure
birefringent LC device

8.2 Real-Time Parallel Analog-to-Digital Conversion

The process of analog-to-digital (A/D) conversion is the representation of analog (continuous) information in digital (binary) form. For one-dimensional signals any coding procedure which assigns to each signal value a group of digits (bits) performs such an operation. For two-dimensional signals (images) we must find a method which assigns to every continuous-tone image a number of binary images (bit planes). Each binary image records the value of one particular bit of the quantized intensity at each image point. Fig. 8.3 shows how the A/D conversion bit planes may be obtained from the device characteristic. For simplicity, it is assumed that the device has a uniform sine squared characteristic over two periods as shown in Fig. 8.3a. By thresholding the output at 1/2 we get a binary function wich can serve as one bit of a three bit binary code. The other two bits are then obtained by attenuating the input intensity to effectively get the first cycle and first half-cycle of the characteristic curve. These curves are then thresholded at the 1/2 level as shown in Figs. 8.3b and c. Note that the binary functions in Fig. 8.3 are the three bits of a reflected binary or grey code corresponding to any continuous input intensity value between 0 and 8. To obtain all these three bits in parallel we can place an array of 3 periodically
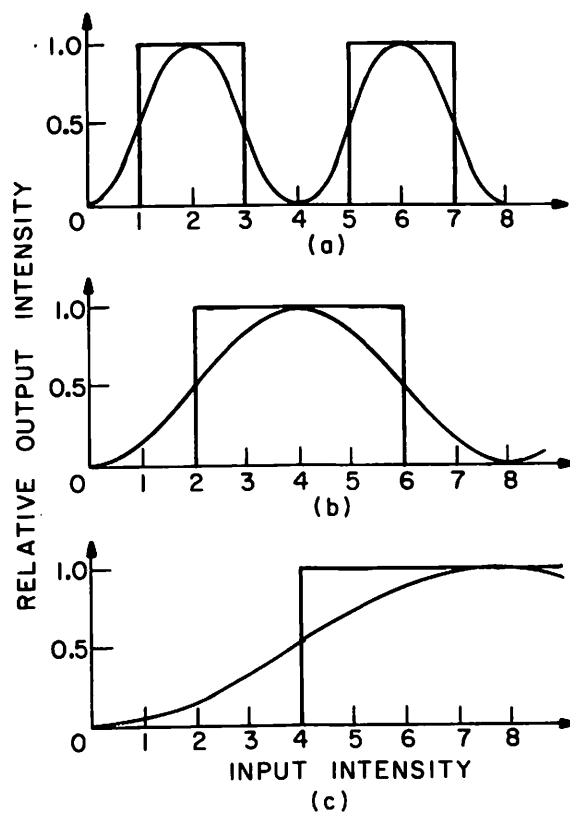
Figure 8.3   Analog-to-digital converter bit
             planes for three bit Gray code

repeated attenuating strips, which can attenuate by factors of 1, 1/2, and 1/4 over the write surface of the LC device. If the period of this array is much smaller than the inverse of the maximum spatial frequency of the input picture, then corresponding to each input intensity value we have three outputs. These outputs are the three sinusoidal curves of Fig. 8.3. By placing an array of electronic thresholding devices in the output plane, we can sense the three bits at the same time all in parallel as shown in Fig. 8.4. Similar ideas have been used for .electro-optic A/D conversion [8-2,8-3], but none have been achieved in an optically controllable device.

8.3 Analog-to-Digital Conversion Experiment

In this experiment we want to obtain the bit planes of an input picture in real time as explained above. The first step is determination of the response curve of the liquid crystal device. From this curve we obtain suitable quantization and thresholding levels for the actual A/D experiments.

8.3.1 Device Transfer Function

The experimental set up to obtain the input output characteristic of the pure birefringent LC device is shown in Fig. 8.5. The input light source is a mercury arc lamp. A pair of fixed and rotatable polarizers in the input light
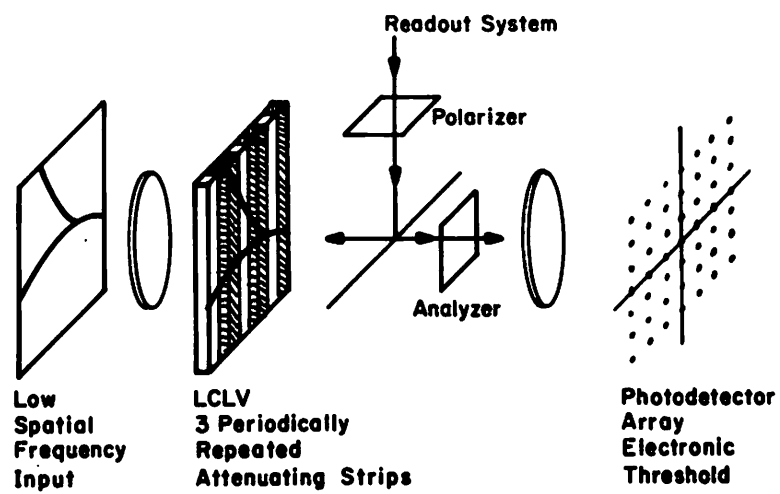
Figure 8.4 System for parallel A/D conversion
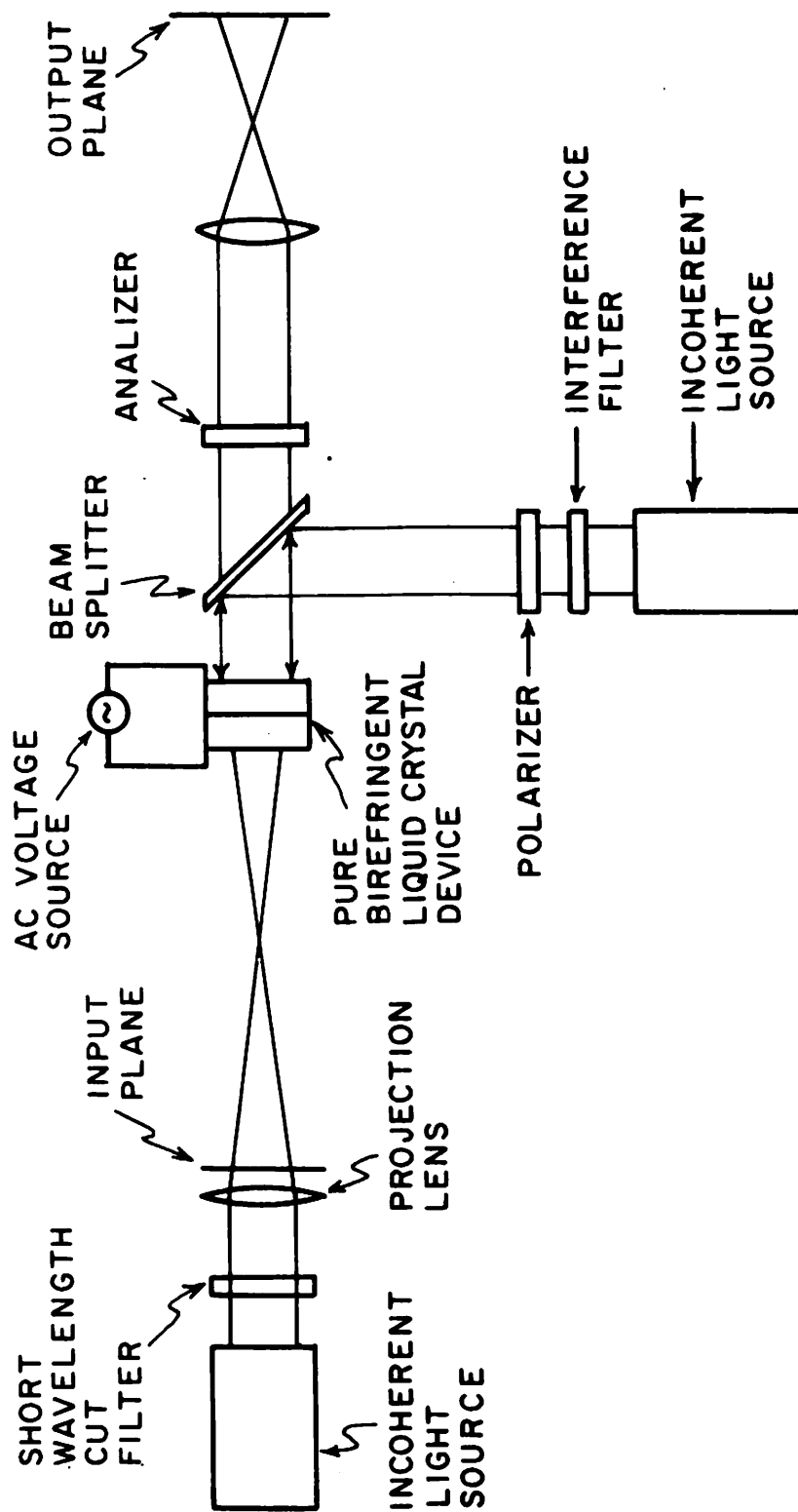
Figure 8.5 Experimental set-up for real-time parallel
analog-to-digital conversion

beam provide a means to vary the input light intensity. The Corning short wavelength cut filter eliminates wavelengths shorter than 493 nm to make sure that the write beam wavelength is within the sensitivity range of the CdS photoconductor. The read light source is another mercury arc lamp. Due to the dispersion of birefringence in the liquid crystal, it should have a narrow spectral bandwidth. A Corion interference filter with peak wavelength 434.7 nm and bandwidth of 18.4 nm was used to meet this requirement for the read light. With no picture in the input plane the output intensity varies in a quasi-sinusoidal fashion with increasing input illumination due to the changing birefringence. If the amount of birefringence varied linearly as a function of the write beam intensity, one would expect a strictly sinusoidal variation of the output intensity. However, a number of factors, including the optical nature of the liquid crystal and the photoconductor characteristic properties affect the output characteristic and produce an approximately sinusoidal output whose frequency varies (monotonically) with input intensity. The response curve obtained is shown in Fig. 8.6.

8.3.2 A/D Experiment

Although the theory behind the A/D conversion assumes a strictly periodic response characteristic, it is in general possible to produce the desired bit planes using

103
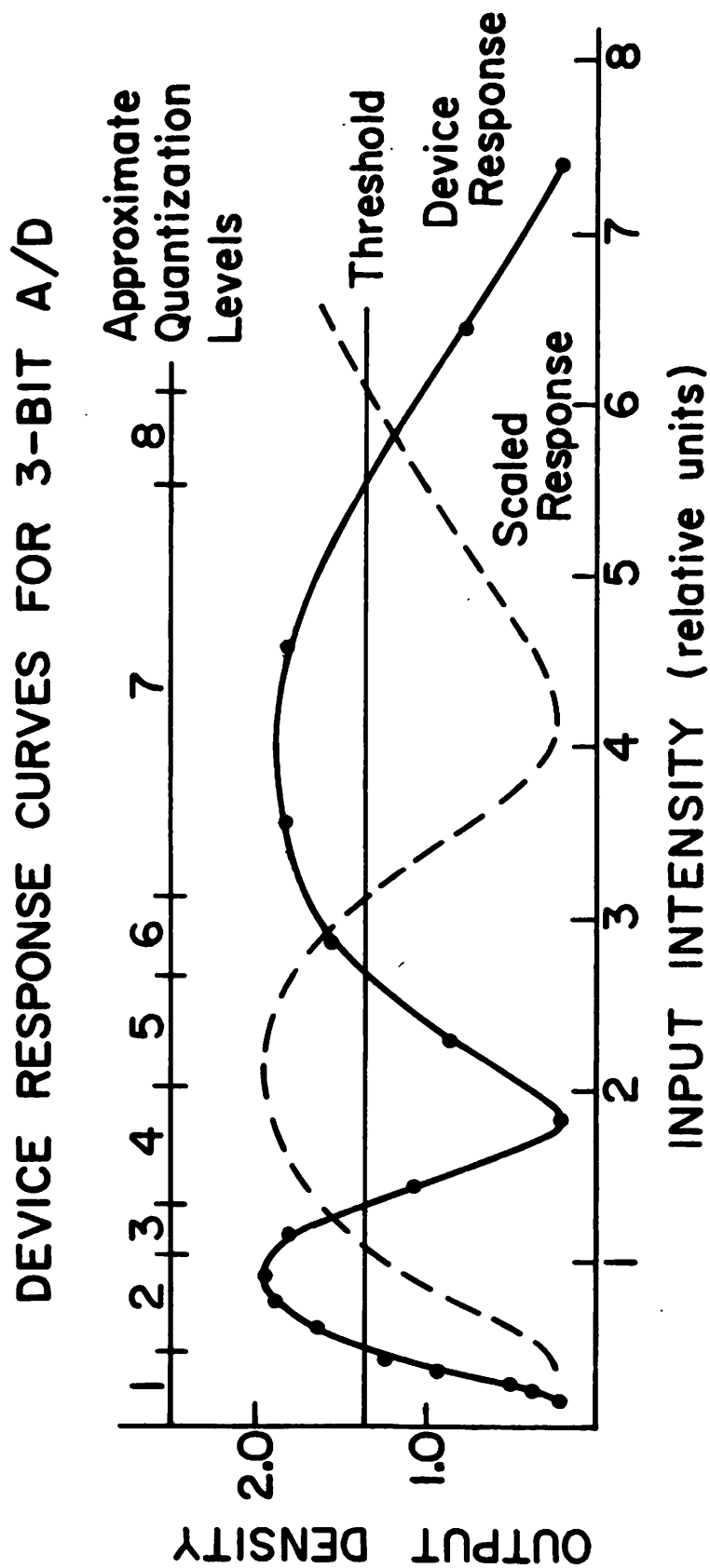
Figure 8.6 Response curve of the LC device used for the three bit
A/D conversion. The solid curve is the measured response.
The dotted curve represents the same response with a
fixed attenuation

the quasi-periodic response curves of the actual devices. The trade-off that must be made is that one must resort to a non-uniform quantization. The quantization levels used in this expriment are indicated in Fig. 8.6.

There are no attenuating strips on the LC device used in this experiment. Instead, the bit planes were generated serially. Also, the output was recorded on hardclipping film rather than a thresholding detector array. A test target was generated consisting of an eight gray level step tablet. The gray levels were chosen to match the quantization levels shown in Fig. 8.6. This test object was imaged onto the LC device and the output was photographically hardclipped. This produced the least significant bit plane of a three-bit A/D conversion. Next the write illumination intensity was decreased, effectively rescaling the device response curve to generate the next bit-plane. The last bit-plane (most significant bit) was obtained by attenuating the write intensity again and photographing the output. The input and three bit planes thus generated are shown in Fig. 8.7.

Although the output contains some noise, the experiment illustrates the principle of real-time parallel incoherent optical A/D conversion. It was found later that the computer generated gray scale was somewhat noisy due to the grain of the high contrast film used. It is possible
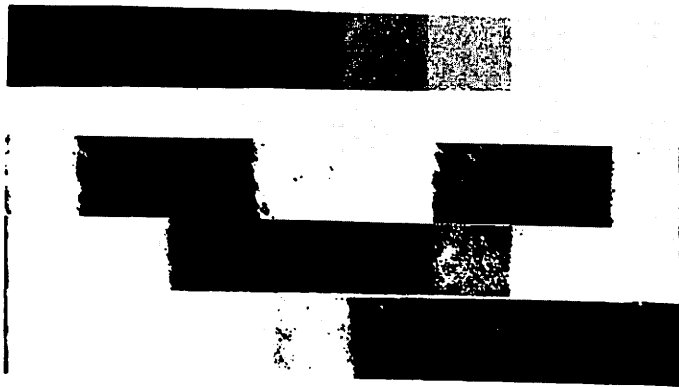
Figure 8.7 Direct analog-to-digital conversion.
The eight level analog input is shown
at the top. Below is the binary coded
output in the form of three bit-planes
of the Gray code

that future experiments with cleaner inputs and improved periodic light valves should produce better experimental results and more bits of quantization.

These results represent the first real-time parallel A/D conversion to be performed on two-dimensional input. The A/D conversion bit rate potential can be estimated from typical parameters of devices currently available. The important parameters are device resolution which is typically 40 line pairs/mm, device size which is on the order of 50mmx50mm, and speed which is generally designed for 30 frames/second. Multiplying all the parameters together implies an A/D conversion rate of 1.2×10^{8} points processed per second.

CHAPTER 9

NONLINEAR PROCESSING USING INTENSITY-TO-SPATIAL

FREQUENCY CONVERSION

In the direct method of nonlinear optical processing, described in the previous chapter, the halftone screen is eliminated at the expense of losing versatility. In this chapter we discuss a method of real-time nonlinear optical processing which does not need any halftone screen. The method operates by transforming intensity to spatial frequency before performing the nonlinear operation. Although we still need a coherent optical processor for this method, its advantages of eliminating the halftone screen and flexibility makes it one of the most promising methods for real-time nonlinear optical processing.

In the following sections we describe a new liquid crystal device capable of intensity-to-spatial frequency transformation and describe how this property is used to obtain an overall nonlinear transformation of intensity. We also consider the processing limitations of this method and present experimental results showing a real-time level slice nonlinearity.

9.1 Variable Grating Mode (VGM) Liquid Crystal Device

In chapter two it was mentioned that conduction effects will predominate if the conductivity of the liquid crystal is sufficient. In this situation, with certain conditions of thickness, voltage, frequency, etc., stable domains are formed [9-1 to 9-3]. These domains are like a volume phase grating structure. Figure 9.1 shows a picture of such domains taken through a polarizing microscope. The spatial period of these domains may be extremely fine. The period of the grating shown in Fig. 9.1 is approximately 18 μm. The interesting property of these domains is that their period is a function of the applied voltage. It has been observed that the relationship between the spatial frequency of these domains and the bias voltage is linear over a voltage range of 20 to 60 volts and spatial frequency range of 100 to 700 line pairs per milimeter [9-3]. By adding a photoconductor layer to the liquid crystal the period of the domains should be a function of the light intensity falling on the photoconductor. The liquid crystal and the photoconductor layers can be arranged in a form similar to the standard LCLV described in Chapter 2 to read out the phase grating variation in reflection. Another option is to simply sandwich the photoconductor and the liquid crystal layers together and use different wavelengths for the write and read beams. The latter configuration has been employed in the device
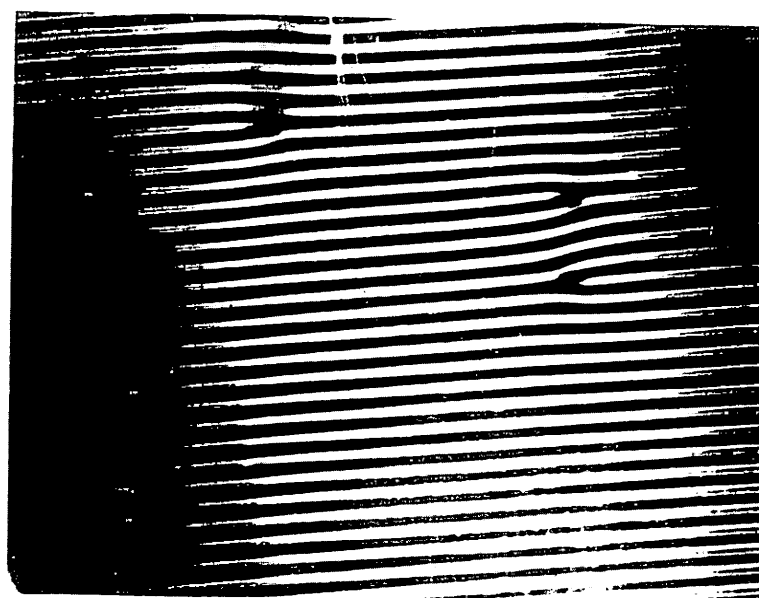
Figure 9.1   VGM domain structure

used in our experiment as shown in Fig. 9.2.

9.2 Nonlinear Processing with the Liquid Crystal Variable Grating Mode

The VGM liquid crystal device can be considered to be an intensity-to-spatial frequency converter capable of operating on two-dimensional images. When an input image is imaged onto the photoconductor surface of this device the intensity variations of the input image change the local grating frequency. If we read out in coherent light and Fourier transform the read-out image, then different spatial frequency components of the encoded image would appear at different locations of the frequency plane as shown in Fig. 9.3a. This means that different locations in the frequency plane correspond to different intensity values in the input image. Thus, by placing appropriate spatial filters in this plane we should be able to obtain different transformations of the input intensity in the output plane as depicted in Fig. 9.3b. This figure describes the variable grating mode nonlinear processing graphically. The input intensity variation is converted to spatial frequency variation by the characteristic function of the VGM device (upper right-hand quadrant). This spectrum is modified by a filter in the Fourier plane (upper left-hand quadrant). Finally, the intensity is observed in the output plane (lower left-hand quadrant).

APPLIED
VOLTAGE

GLASS
PLATES

CONDUCTIVE
NESA-COATINGS

LIQUID CRYSTAL

IMAGING
LIGHT

PHOTOCONDUCTIVE COATING
OVER THE CONDUCTIVE
NESA COATING
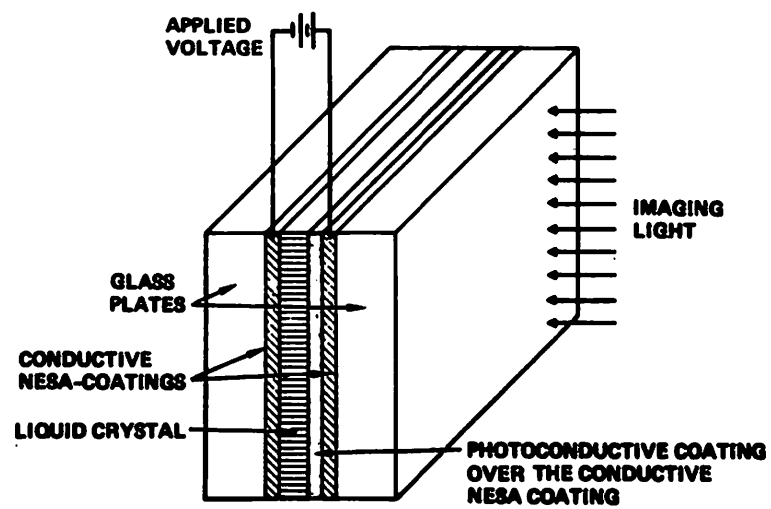
Figure 9.2   Construction of an optically
             activated VGM real-time device
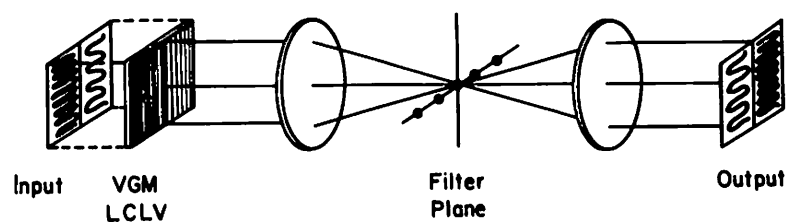
(a)
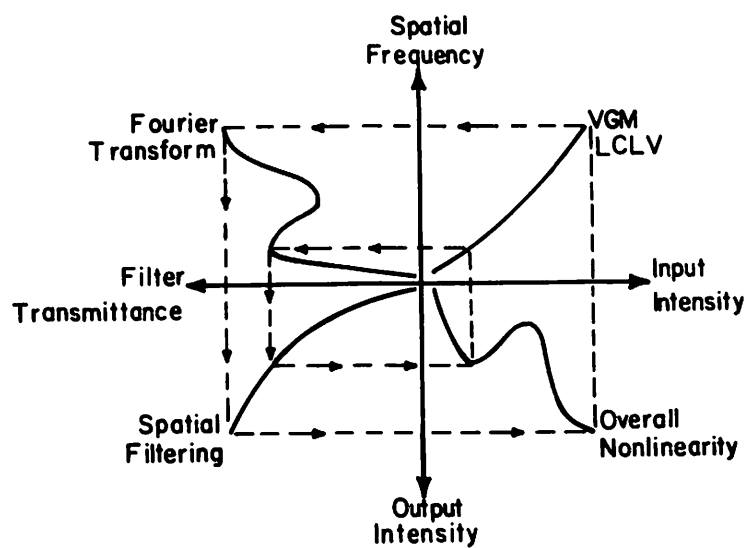
(b)

Figure 9.3   Implementation of a nonlinearity
             utilizing the variable grating effect

             (a)   optical system
             (b)   components of the nonlinear
                   transformation

Taken together these yield the overall nonlinearity (lower right-hand quadrant). Design of a proper spatial filter for a desired transformation is a relatively easy task compared to the design of a halftone screen. The only parameter available in the design of the spatial filter is the location in the frequency plane. For example, for a level slice transformation a simple slit that passes certain frequency bands will be adequate. For a better understanding of the nonlinear processing with VGM, we now formulate this process in mathematical terms.

9.3 Variable Phase Grating Analysis

Assume that we have a phase grating (periodic array of phase shifting elements) extending to infinity in both the x and y directions. Let $\underline{g}(x,y)$ be the complex amplitude transmittance of the phase grating. Since the grating extends to infinity we can assume that $\underline{g}(x,y)$ is a <u>periodic</u> function of x and write it as a Fourier series expansion

$$\underline{g}(x,y) = \sum_{n=-\infty}^{\infty} \underline{c}_n \exp(j2\pi nx/L) \qquad (9.1)$$

where

$$\underline{c}_n = \frac{1}{L} \int_0^L \underline{g}(x,y)\exp(-j2\pi nx/L)\,dx \qquad (9.2)$$

and L is the period of the grating. Because we assume that the grating extends to infinity in the y-direction, $\underline{g}(x,y)$ is not an explicit function of y.

114

Now we assume that there is an aperture $a(x,y)$ in the input $(x-y)$ plane where

$$a(x,y) = \begin{cases} 1, & \text{for } (x,y) \text{ inside the aperture} \\ 0, & \text{otherwise} \end{cases} \tag{9.3}$$

Then the transmittance function is

$$\underline{t}(x,y) = \underline{g}(x,y) \cdot a(x,y) = a(x,y) \cdot \sum_{n=-\infty}^{+\infty} \underline{c}_n \exp(j2\pi nx/L). \tag{9.4}$$

If this aperture is placed in the input plane of a coherent optial processor, then in the Fourier plane we get

$$\underline{T}(f_X, f_Y) = \underline{A}(f_X, f_Y) * \sum_{n=-\infty}^{+\infty} \underline{c}_n \delta(f_X - \tfrac{n}{L}, f_Y) \tag{9.5}$$

where

$$\underline{T}(f_X, f_Y) = \mathscr{F}\{\underline{t}(x,y)\}$$

$$\underline{A}(f_X, f_Y) = \mathscr{F}\{a(x,y)\}$$

or

$$\underline{T}(f_X, f_Y) = \sum_{n=-\infty}^{+\infty} \underline{c}_n \underline{A}(f_X - \tfrac{n}{L}, f_Y). \tag{9.6}$$

For simplicity, we assume that the aperture is square with dimension bxb. Then

$$a(x,y) = \text{rect}(\tfrac{x}{b}) \text{rect}(\tfrac{x}{b}) \tag{9.7}$$

and

$$\underline{A}(f_X, f_Y) = b^2 \text{sinc}(bf_X) \text{sinc}(bf_Y) . \tag{9.8}$$

Hence

$$\underline{T}(f_X, f_Y) = \sum_{n=-\infty}^{+\infty} \underline{c}_n b^2 \text{sinc}[b(f_X - \frac{n}{L})] \text{sinc}(bf_Y) \quad (9.9)$$

and in the Fourier plane we have different diffraction orders each of which is a two-dimensional sinc function.

. Note that if b is large [9-4]

$$\lim_{b \to \infty} b \text{ sinc } bt = \delta(t) \quad (9.10)$$

and we have

$$\underline{T}(f_X, f_Y) = \sum_{n=-\infty}^{+\infty} \underline{c}_n \delta(f_X - \frac{n}{L}) \delta(f_Y)$$

$$= \sum_{n=-\infty}^{+\infty} \underline{c}_n \delta(f_X - \frac{n}{L}, f_Y) \quad . \quad (9.11)$$

Assume that we use a filter $\underline{H}$ in the Fourier plane that passes the first order only. Then the amplitude in the output plane is

$$\underline{V}(x, y) = \mathscr{F}^{-1}\{\underline{c}_1 \delta(f_X - \frac{1}{L}, f_Y) \underline{H}(f_X, f_Y)\}$$

$$= \underline{c}_1 \underline{H}(\frac{1}{L}, 0) \quad . \quad (9.12)$$

The output intensity is then

$$I_o(x_0, y_0) = |\underline{c}_1 \underline{H}(\frac{1}{L}, 0)|^2 \quad (9.13)$$

and the value of the intensity at any point in the output plane is related to the value of the filter function at the frequency $(\frac{1}{L}, 0)$ in the Fourier plane. Now if we omit the assumption of $b \to \infty$ the output amplitude is

$$\underline{V}(x_0,y_0) = \mathscr{F}^{-1}\{\underline{c}_1 b^2 \text{sinc}[b(f_X-\tfrac{1}{L})]\cdot\text{sinc}(bf_Y) \qquad (9.14)$$
$$\cdot\underline{H}(f_X,f_Y)\}.$$

This assumes that the orders are nonoverlapping, a condition which occurs if b>>2L. For an aperture of general shape

$$\underline{V}(x_0,y_0) = \mathscr{F}^{-1}\{\underline{c}_1\underline{A}(f_X-\tfrac{1}{L},f_Y)\underline{H}(f_X,f_Y)\}. \qquad (9.15)$$

If b is the average dimension of the aperture, then $\underline{A}$ will have a width on the order of 2/b as in the case of the square aperture. The output intensity in this case is

$$I_0(x_0,y_0) = |\mathscr{F}^{-1}[\underline{c}_1\underline{A}(f_X-\tfrac{1}{L},f_Y)\underline{H}(f_X,f_Y)]|^2. \qquad (9.16)$$

Now, if $\underline{H}(f_X,f_Y)$ is slowly varying, we can consider it to be essentially constant over the width of $\underline{A}$. Thus we can write

$$I_0(x_0,y_0) \cong |\mathscr{F}^{-1}[c_1\underline{A}(f_X-\tfrac{1}{L},f_Y)\underline{H}(\tfrac{1}{L},0)]|^2$$

$$\cong |\underline{H}(\tfrac{1}{L},0)\mathscr{F}^{-1}[c_1\underline{A}(f_X-\tfrac{1}{L},f_Y)]|^2$$

$$\cong |\underline{H}(\tfrac{1}{L},0)c_1 a(x_0,y_0)\exp(-j2\pi x_0/L)|^2 \qquad (9.17)$$

$$\cong c_1|a(x_0,y_0)\cdot\underline{H}(\tfrac{1}{L},0)|^2 .$$

Going back to the definition of a(x,y) in Eq. (6.3) we have

$$I_0(x_0,y_0) = \begin{cases} c_1|\underline{H}(\tfrac{1}{L},0)|^2, & \text{for } (x_0,y_0) \text{ within the aperture} \\ 0 & \text{, otherwise.} \end{cases} \qquad (9.18)$$

Thus, this filtering process produces an aperture in the output plane corresponding to an aperture $a(x,y)$ in the input plane modulated by the phase grating with period L. The output intensity is related to the grating period through the filter function given by Eq. (9.18).

The above analysis can be used to estimate the processing limitations of the VGM device. The fundamental question to be answered is how large the smallest picture element, or pixel, must be with respect to the VGM grating frequency.

Suppose that the usable spatial frequency range over which the VGM device can operate is $\Delta \nu_o$. To avoid crosstalk from higher orders the VGM device can only be operated between the first and second diffraction orders of the lowest frequency used. If this lowest frequency is denoted by ν_o then we have

$$\nu_o = 2\nu_o - \nu_o = \nu_o \tag{9.19}$$

as shown in Fig. 9.4. Now if the number of intensity levels that we wish to distinguish is N, then the Fourier transform of any one aperture represented by $2\Delta\nu$ in Fig. 9.4 must be contained within a region of width ν_o/N in the Fourier plane. So we must have

$$2\Delta\nu \leq \nu_o/N \ . \tag{9.20}$$

OPTICAL
FOURIER
TRANSFORM

VGM
IMAGE

$2\Delta\nu$

ν_0

$2\nu_0$

LOWEST
VGM
FREQUENCY

OBJECT
SPECTRUM
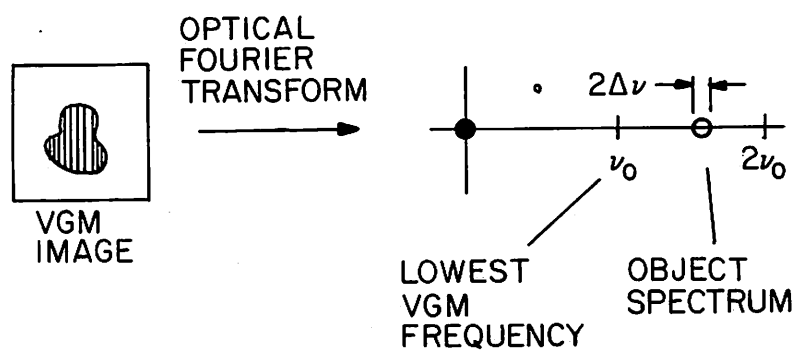
Figure 9.4   Gray level resolution with VGM

If the pixel width is b then the width of its Fourier tranform is 2/b which means

$$2\Delta\nu = \frac{2}{b} \quad . \tag{9.21}$$

Combining Eq. (9.20) and (9.21) we have

$$\frac{2}{b} \leq \nu_o/N \tag{9.22}$$

or

$$b\nu_o \geq 2N. \tag{9.23}$$

This requires that the pixel size spans 2N periods of the lowest grating frequency if N grey levels are to be processed.

As an example if ν_o =200 cycles/mm and N=50, then each pixel must have a size b=0.5 mm. Thus if the device has a 50 mm square area a 100x100 pixel image could be processed.

9.4 Experiment with the VGM Device

In this experiment we show the ability of the VGM device to generate a level-slice nonlinearity. The experimental setup is shown in Fig. 9.5. The input picture is illuminated by an arc lamp source and imaged onto the photoconductor surface of a VGM device which has an existing phase grating structure due to a bias voltage. The grating period is locally modulated by the input
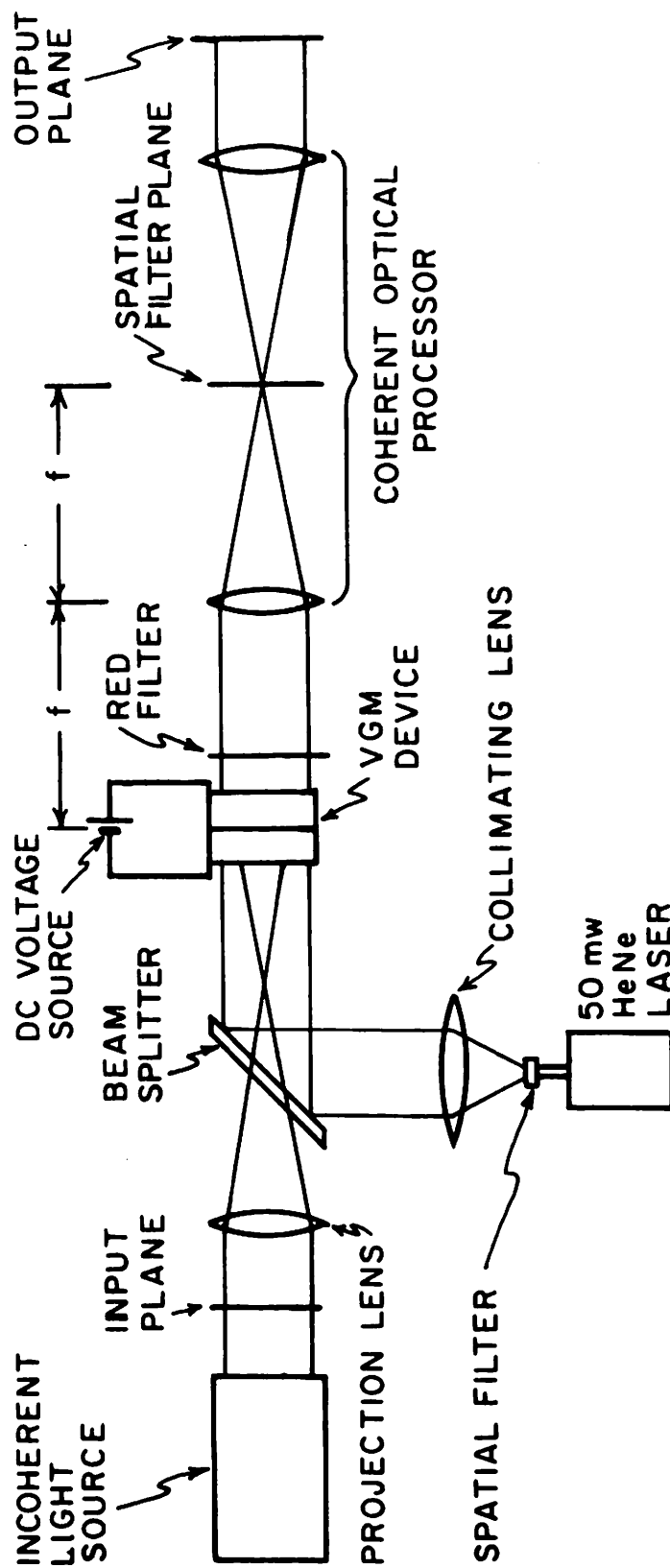
Figure 9.5 Experimental set-up for real-time parallel level-slicing

picture intensity, and this modulation is mapped into a position along a line in the spatial filter plane. A red filter ensures that only the laser beam enters the coherent optical processor. Small circular annuli of varying radii are used to pass certain spatial frequency bands. This in effect allows certain input intensity ranges to appear in the output. Figure 9.6 shows the input and level sliced output pictures. Figure 9.6a shows a positive print of the original image as photographed on the imaging screen. A negative of the original was used in the experiments. Figure 9.6b shows a low intensity level slice corresponding to a VGM periodicity of 120 lines/mm with approximately 3% bandwidth. In Fig. 9.6c another level, corresponding to 153 lines/mm, is shown. Figure 9.6d at 236 lines/mm, illustrates the interference from second harmonics. Weak second harmonics of the low-intensity image slice corresponding to 118 lines/mm can appear in the 236 lines/mm level slice. In Fig. 9.6e, a broader slice of approximately 11% bandwidth was taken centered about the level corresponding to 140 lines/mm. This picture may be compared with the previous slices and particularly with the slice shown in Fig. 9.6c. Finally, Fig. 9.6f shows a very high intensity slice at 440 lines/mm of 10% bandwidth. Three grey levels may be seen simultaneously; these correspond to the superposition of three wide intensity slices.

A ORIGINAL IMAGE

B

C

D

E

F

Figure 9.6   Level slicing with the VGM liquid crystal device

CHAPTER 10

CONCLUSIONS AND FUTURE RESEARCH TOPICS

In this dissertation we have presented three different techiques for performing nonlinear optical processing in real time. These methods specifically use the following ideas and devices to obtain nonlinear transformations: 1) halftone preprocessing with a real-time image transducer; 2) nonlinear characteristics of a real-time device; and 3) capability of a real-time device to produce intensity-to-spatial frequency conversion.

To implement halftone nonlinear processing in real time we have replaced the conventional photographic hard-clipping recording medium with a real-time image transducer. The nonbinary characteristic of the real-time device degrades the output of the process. A detailed analysis has been made and we have found exact methods for predicting such degradations. The analysis is general and works for any recording medium characteristic curve shape and any halftone screen cell shape. It has been shown that for some monotonic functions it is possible to obtain exact compensation for the nonideal characteristic curve of the recording medium. The compensation problem has been solved

in general by an approximate method which obtains the halftone screen density profile by minimizing the difference in a mean-square sense between desired and degraded outputs. The advantages of the halftone method are its flexibility to produce a variety of nonlinear transformations and that it is a well-explored technique. Its limitations are the soft threshold characteristic of the real-time device and necessity for accurate high space-bandwidth product halftone screens. A real-time logarithmic transformation via this method has been shown.

The inherent nonlinear characteristics of a special real-time birefringent liquid crystal device were used to obtain direct nonlinear transformations. This method is particularly interesting because it may be done with incoherent light and the space-bandwidth product of the image to be processed can be as large as the space-bandwidth product of the device. However, it is limited because the type of nonlinearity achievable depends on the nonlinear shape of the characteristic curve of the real-time device used. A real-time three-bit parallel incoherent analog-to-digital conversion nonlinearity has been obtained by this method.

The property of the liquid crystals to generate, under certain conditions, a volume phase grating whose period varies with the applied voltage was found suitable for

nonlinear optical processing. It was shown that a real-time device with such property is capable of intensity-to-spatial frequency conversion which can be used to generate nonlinear transformations. Like the halftone method this method is very flexible and can generate many nonlinearities. As in the halftone method, the carrier frequency modulation required limits the space-bandwidth product of the picture to be processed. A real-time level-slice nonlinearity was obtained by this method.

Several topics seem worthy of further effort. Our analysis for the halftone process shows that although we can achieve good approximations to the desired nonlinear transformations with nonideal devices, a device with sharp threshold characteristics can improve the results quite significantly. Research for real-time image transducers with such a characteristic should continue so that the full potential of the halftone method can be achieved. Finding new methods to design accurate, high space-bandwidth product halftone screens is still a problem.

The capabilities of direct nonlinear processing are not yet fully understood. New applications for the existing device characteristics remain to be found and new devices with specified characteristics are to be developed. One obvious extension is to achieve a parallel A/D conversion nonlinearity with more than three bits. In this

126

case we need a device which has four or more intensity peaks with very regular period in its transfer characteristic. Other parameters like wavelength and the bias voltage could be used to give extra dimensions to the whole process. For example, different nonlinear characteristics could be obtained at the same time for different wavelength. The immediate use of such a property is in A/D conversion where we then can obtain different bit planes in different colors.

In the area of the VGM device the theory presented in Chapter 9 should be extended to explictly consider entire images as a whole rather than just single apertures. This way the interaction of all apertures which comprise an image will be taken into account. Another theoretical question to be addressed concerns the fact that the VGM devices do not produce strictly periodic phase gratings as can be seen in Fig. . This results in a blurring of the diffraction orders. This in turn affects the number of gray levels that can be processed with a given device and a fixed image space-bandwidth product. The three dimensional nature of the phase grating structure observed in the VGM device is another factor that should be considered in a future analysis of this device.

APPENDIX A

To simplify

$$\int_{h(y-\varepsilon)}^{h(y)} f(x)\,dx \qquad\qquad (A.1)$$

where $h(\cdot)$ is the inverse function of $f(\cdot)$, let

$$u = f(x) \implies x = f^{-1}(u) = h(u) \qquad (A.2)$$

then

$$x = h(y-\varepsilon) \implies u = f(x) = f[h(y-\varepsilon)] = y-\varepsilon \qquad (A.3)$$

and

$$x = h(y) \implies u = f(x) = f[h(y)] = y \qquad (A.4)$$

also $du = f'(x)\,dx$ but since $x = h(f(x))$

$$f'(x)\cdot h'[f(x)] = 1 \quad\text{or}\quad f'(x) = \frac{1}{h'[f(x)]} = \frac{1}{h'(u)} \qquad (A.5)$$

$$\implies dx = h'(u)\,du$$

hence

$$\int_{h(y-\varepsilon)}^{h(y)} f(x)\,dx = \int_{y-\varepsilon}^{y} u\cdot h'(u)\,du. \qquad (A.6)$$

128

Using integration by parts in this integral we have

$$\int_{h(y-\epsilon)}^{h(y)} f(x)\,dx = uh(u)\Big|_{y-\epsilon}^{Y} - \int_{y-\epsilon}^{Y} h(u)\,du$$

$$= yh(y) - (y-\epsilon)h(y-\epsilon) - \int_{y-\epsilon}^{Y} h(u)\,du \quad .$$

(A.7)

REFERENCES

[1-1] Special issue on Optical Computing, <u>Proc. IEEE</u>, Vol. 65, January 1977.

[1-2] W.K. Pratt, <u>Digital Image Processing</u>, Wiley-Interscience, New York, 1978.

[1-3] J.W. Goodman, <u>Introduction to Fourier Optics</u>, McGraw-Hill, New York, 1968.

[1-4] M. Marquet and J. Tsujiuchi, "Implementation of Particular Aspects of Dehalftoned Images," <u>Optica Acta</u>, Vol. 8, pp. 267-277, 1961.

[1-5] H. Kato and J.W. Goodman, "Nonlinear Filtering in Coherent Optical Systems Through Halftone Screen Processes," <u>Applied Optics</u>, Vol. 14, pp. 1813-1824, 1975.

[1-6] S.R. Dashiell and A.A. Sawchuk, "Optical Synthesis of Nonlinear Nonmonotonic Functions," <u>Optics Communications</u>, Vol. 15, pp. 66-70, 1975.

[1-7] A.W. Lohmann and T.C. Strand, "Analog-to-Digital Conversion of Pictures with Optical Means," <u>Proceedings Electro-Optics/International Laser Conference 1975</u>,

Anaheim, Calif., pp. 16-21, 1975.

[1-8]   H-K. Liu and J.W. Goodman, "A New Coherent Optical Pseudo-Color Encoder," Nouv. Rev. Optique, Vol. 7, pp. 285-289, 1976.

[1-9]   B.J. Thompson, "Hybrid Processing Systems-An Assessment," Proc. IEEE, Vol. 65, pp. 62-76, 1977.

[1-10] D. Casasent, "A Hybrid Image Processor," Optical Engineering, Vol. 13, pp. 228-234, 1974.

[1-11] S. Iwasa and J. Feinleib, "The PROM Device in Optical Processing Systems," Optical Engineering, Vol. 13, pp. 235-242, 1974.

[1-12] T.D. Beard, et al., "AC Liquid Crystal Light Valve," Appl. Phys. Lett., Vol. 22, pp. 90-92, 1974.

[1-13] J. Grinberg, et al., "A New Real-Time Non-Coherent to Coherent Light Image Converter," Optical Engineering, Vol. 14, pp. 217-225, 1975.

[1-14] S.R. Dashiell and A.A. Sawchuk, "Nonlinear Optical Processing: Effect of Input Medium and Precompensation," Applied Optics, Vol. 16, pp. 2279-2287, 1977.

[1-15] A. Tai, T. Cheng and F.T.S. Yu, "Optical Logarithmic Filtering Using Inherent Film Nonlinearity," Applied Optics, Vol. 16, pp. 2559-2564, 1977.

[1-16] J.M. Pollack and J.B. Flannery, "A Low-Noise Image Amplifier," Society for Information Display 1976 International Symposium Digest, pp. 142-145, 1976.

[2-1] J.W. Goodman, "Operations Achievable with Coherent Optical Information Processing Systems," Proc. IEEE, Vol. 65, pp. 29-38, 1977.

[2-2] G.H. Heilmeier, et al., "Dynamic Scattering: A New Electrooptic Effect in Certain Classes of Nematic Liquid Crystals," Proceedings of the IEEE, Vol. 56, No. 7, pp. 1162-1171, July 1968.

[2-3] R. Williams, "Domains in Liquid Crystals," The Journal of Chemical Physics, Vol. 19, No. 2, 1963.

[2-4] J.D. Margerum and L.J. Miller, "Electro-Optical Application of Liquid Crystals," Journal of Colloid and Interface Science, (1977); Hughes Research Laboratories Research Report 499, June 1976.

[2-5] J. Grinberg, et al., "A New Real-Time Non-Coherent to Coherent Light Image Converter," Optical Engineering, Vol. 14, pp. 217-225, 1975.

[3-1] S.R. Dashiell and A.A. Sawchuk, "Nonlinear Optical Processing: Analysis and Synthesis," Applied Optics, Vol. 16, pp. 1009-1025, 1977.

[3-2]   S. Iwasa and J. Feinleib, "The PROM Device in Optical Processing Systems," Optical Engineering, Vol. 13, pp. 235-242, 1974.

[3-3]   J. Grinberg et al., "A New Real-Time Non-Coherent to Coherent Light Image Converter," Optical Engineering, Vol. 14, pp. 217-225, 1975.

[3-4]   S.R. Dashiell and A.A. Sawchuk, "Nonlinear Optical Processing: Effects of Input Medium and Precompensation," Applied Optics, Vol. 16, pp. 2279-2287, 1977.

[3-5]   J.W. Goodman, Introduction to Fourier Optics, McGraw-Hill, New York, 1968.

[4-1]   S.R. Dashiell and A.A. Sawchuk, "Nonlinear Optical Processing: Effects of Input Medium and Precompensation," Applied Optics, Vol. 16, pp. 2279-2287, 1977.

[5-1]   G.W. Batten, Jr., and R.L. Everett, "Control of Film Characteristics by Modulating Intensity and Space," J. Opt. Soc. Am., Vol. 68, pp. 1118-1124, 1978.

[5-2]   R.E. Bellman, and K.L. Cooke, Differential-Difference Equations, Academic Press, New York, pp. 63, 1963.

[6-1]   R.W. Hamming, Numerical Methods for Scientists and Engineers, McGraw-Hill, New York, pp. 673, 1973.

[6-2]   This subroutine was taken from the IMSL library 2 which is available from IMSL, 7500 Bellaire Boulevard, Houston, Texas 77036.

[6-3]   R. Fletcher, "Fortran Subroutines for Minimization by Quasi-Newton Methods," Report R7125 AERE, Harwell, England.

[7-1]   T.C. Strand, "Techniques and Applications of Nonlinear Processing with Halftones," Proc. SPIE, 20th Annual Technical Symposium, San Diego, Ca., August 1976.

[7-2]   H-K. Liu,   J.W. Goodman,   and   J. Chan, "Equidesitometry by Coherent Optical Filtering," Applied Optics, Vol. 15, pp. 2394-2399, 1976.

[7-3]   H.C. Andrews, A.G. Tescher, and R.P. Kruger, "Image Processing by Digital Computer," IEEE Spectrum, Vol. 9, No. 7, pp. 20-37, 1972.

[7-4]   A.V. Oppenheim, R.W. Schaefer and T.G. Stockham, Jr., "Nonlinear Filtering of Multiplied and Convolved Signals," Proc. IEEE, Vol. 56, pp. 1264-1291, 1968.

[7-5]   H. Kato and J.W. Goodman, "Nonlinear Filtering in Coherent Optical Systems Through Halftone Screen Processes," Applied Optics Vol. 14, pp. 1813-1824, 1975.

[8-1]   A. Tai, T. Cheng, F.T.S. Yu, "Optical Logarithmic Filtering Using Inherent Film Nonlinearity," Applied

Optics, Vol. 16, pp. 2559-2564, 1977.

[8-2]   H.F. Taylor, "An Electrooptic Analog-to-Digital Converter," Proc. IEEE, 63, pp. 1524-1525, 1975.

[8-3]   F.A. Ludewig, Jr., "Digital Transducer System," U.S. Patent No. 3087148, April 23, 1963.

[9-1]   G.H. Heilmeier, et al., "Dynamic Scattering: A New Electrooptic Effect in Certain Classes of Nematic Liquid Crystals," Proceedings of the IEEE, Vol. 56, No. 7, pp. 1162-1171, July 1968.

[9-2]   R. Williams, "Domains in Liquid Crystals," The Journal of Chemical Physics, Vol. 19, No. 2, 1963.

[9-3]   J.M. Pollack and J.B. Flannery, "A Low-Noise Image Amplifier," Society for Information Display 1976 International Symposium Digest, pp. 142-145, 1976.