INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

University Microfilms International
300 North Zeeb Road
Ann Arbor, Michigan 48106 USA
St. John's Road, Tyler's Green
High Wycombe, Bucks, England HP10 8HR

DELTA MODULATION ENCODING OF VIDEO SIGNALS

by

TSU-LEUN R. LEI

A dissertation submitted to the
Graduate Faculty in Engineering
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy,
The City University of New York.

1978

This manuscript has been read and accepted for the Graduate Faculty in Engineering in satisfaction of the dissertation requirement for the degree of Doctor of Philosophy

6/7/78
date

Chairman of
Examining Committee

6/2/78
date

Executive Officer

Dr. Joseph Garodnick

Prof. Se Jeung Oh

Prof. Donald L. Schilling (Chairman)

Prof. Herbert Taub

Prof. Frederick E. Thau

Supervisory Committee

The City University of New York

To my parents and Sophie

Abstract

DELTA MODULATION ENCODING OF VIDEO SIGNALS

by

Tsu-Leun R. Lei

Advisor: Professor Donald L. Schilling

The application of delta modulators (DM) on the encoding of video signals is investigated. Various kinds of delta modulation algorithms were tested by computer simulated systems. The output signal-to-noise ratio and the signal to frame difference ratio are measured. DM step size statistics are studied. Channel error responses are also studied. Delta modulator encoded pictures are shown for subjective comparison. A discussion of entropy encoding of the DM output bit stream is also included.

The Song mode one dimensional adaptive delta modulator (ADM) is studied in Chapter 1 for its application on encoding video signals. Two new algorithms, the A-mode ADM and the B-mode ADM, are proposed and studied for the

purpose of reducing edge busyness. The two dimensional delta modulators (2DDM) are simulated and studied. The advantages of the 2DDM are the higher signal-to-noise ratio and the elimination of edge busyness. Entropy encoding, discussed in Chapter 4, further supports the statement that the two dimensional DM encoding is superior to the one dimensional DM encoding. Channel error studies have shown that the one dimensional DM has a better channel error response. However, the biased 2DDM can also produce good quality pictures in the presence of channel errors.

The results of this research show that the delta modulator is indeed an effective and economical digital encoding system for video signals.

ACKNOWLEDGEMENTS

The author wishes to express his sincere gratitude to Professor Donald L. Schilling for his guidance. His broad knowledge and practical experiences have helped significantly in this research. His financial assistance during the course of this work is greatly appreciated.

Appreciation is also extended to Dr. Joseph Garodnick for his valuable comments and suggestions. The constant discussions with Mr. Norman Scheinberg have helped immeasurably in the understanding of many fine details of experimental procedures. The interest shown by Dr. Richard Marsten in listening to new ideas has encouraged and inspired many of the works. Thanks are also expressed to the colleagues at the communications laboratory of the City College of New York for their assistance and friendship.

A special appreciation is given to my wife Sophie for her love, understanding, encouragement, and support.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

Chapter I

INTRODUCTION

1.1 The Digital Communications System

The ultimate objective of a communications system is to accurately reproduce the input signals in a most efficient and economical way. Many communications systems convert analog signals into digital signals for processing and transmission [1-3]. There are several advantages for this:

- Digital signals can be regenerated. With the aid of repeaters, a digital communications system has the capability of transmitting information over long distances without deterioration.
- Digital computers can be easily linked to a digital transmission system as a powerful means of control and processing device.
- The common phenomenon of cross talk distortion in an analog channel can be avoided if digital communication system is used.
- It is easier to implement a time-division multiplexing switching system for a digital signal than to implement a frequency-division multiplexing system for an analog signal.
- In a large scale communications network, when certain

facilities are shared by many subscribers, the system can be operated efficiently only if digital transmission systems are employed.

The shortcomings of digital transmission are its wide bandwidth requirement and its complex circuitry. The problems in hardware complexity are gradually being solved by the recent development of the advanced large-scale-integrated circuit technology. The narrow band digital transmission techniques, however, still require further research.

A straight forward pulse-code modulation (PCM) encoding for video signals with 8-bit resolution sampling at the Nyquist rate requires a digital channel bandwidth of 16 times the original analog bandwidth. Many techniques aiming at reducing the bandwidth requirement have been studied. There has been some degree of success on many different approaches [4,5]. There are two approaches which are commonly used to efficiently encode video analog signals. One is the Transformation encoding and the other is the Feedback Predictive encoding.

1.2 Transformation Encoding

In a PCM video encoding system, the quantized value of each picture element (pixel) is transmitted sequentially. However, pictorial data is not a random variable, it contains a significant two dimensional structure [6,7].

This particular structure provides the redundancy in video signals. Efficient coding techniques can be employed to remove the redundancy contained in the analog or in the PCM encoded video signals. One of the coding techniques is the linear transformation [8-16] which transforms a set of statistically dependent pixels to a set of coefficients which is less dependent. These transformed coefficients are then filtered by discarding some of the insignificant coefficients.The remaining coefficients represent the magnitude of the essential basis pictures in the transformed domain. They are coded and transmitted through the channel with less channel bandwidth, because part of the redundancy has been removed by the filtering process in the transformed domain. The ideal transformation is the one that produces independent coefficients. However, the most efficient transformation algorithm known today can only produce uncorrelated coefficients. It is known as the Karhunen-Loève transformation in which the resulting coefficients are uncorrelated but not necessarily independent.

Other transformations currently being used are the Fourier transformation, the Hadamard transformation, the Slant transformation, and the Haar transformation, etc. [8,9,16]. One-dimensional and two-dimensional transformations are both employed in the transformation coding techniques. The two dimensional transformation can achieve higher

bandwidth compression than the one-dimensional transformation as indicated by Wintz [10]. The two-dimensional adaptive transformation coding appears to be able to achieve the highest bandwidth compression. However, practical applications of the transformation encoding techniques especially when real time process is required, are limited by their extremely complicated circuitry and long computation time. In this dissertation , we are interested in searching for a simple, effective and economical means of digital transmission system capable of real time video processing. Therefore the transformation picture coding techniques were not studied in detail.

1.3 Feedback Predictive Encoding

It is known that sampled values of a video signal have strong correlations. A coding scheme is called a feedback predictive coding if the correlation properties between the adjacent pixels are utilized. Feedback predictive coding is a procedure in which the transmitter transmits not the actual sample value but the difference between the actual sample value and the predicted value. At the receiver, the received difference value is added to the receiver predicted value to reproduce the actual message. In a noiseless communication system, the predicted value for each pixel at both the transmitter and the receiver are the same.

Therefore, truthful communication is obtainable when the differences between the actual samples and the predicted values are transmitted. The prediction of the future pixels is based upon the information of the previous pixels. There are many different approaches to utilize the pixel correlations and to predict the future samples. Figure 1.1 shows a feedback predictive coding system which consists of three functional elements, a quantizer, a predictor, and a channel bit encoder/decoder. The quantizer quantizes the difference signal between the input sample X_k and the predictive value $\hat{X}_k$. The channel bit encoder codes it into binary digits for transmission. The predictor predicts the next sample value $\hat{X}_{k+1}$ based on the previous predicted sample value $\hat{X}_k$ and the present quantized difference Q_k.

1.3.1 The Quantizer

After the difference between the actual sample value and the predicted value of a pixel is obtained, it is quantized, coded, and transmitted. Generally, a feedback predictive coding system which employes a two-level quantizer is called a delta modulator (DM). When a multi-level quantizer [17-19] is used, the coding system is called a differential-pulse-code modulator (DPCM). A delta modulator quantizes the difference value into two levels, +S or -S, where S denotes the step size of a DM. The channel

bit encoder then, encodes it into a binary state '1' or '0'. Hence, a delta modulator transmits one output bit for each sample. A DPCM quantizes the difference value into n levels. Generally, n is an integer of the m-th power of 2. A DPCM is called a m-bit DPCM if

$$n = 2^m .$$

If the distances between quantization levels are fixed and uniformly spaced as shown in Fig. 1.2, it is a linear quantizer. For a DM, it is called a linear DM. For a DPCM, it is called a linear quantized DPCM. A non-linear quantizer, as shown in Fig. 1.3, has non-uniform space distribution between quantization levels. A DPCM with non-linear quantizer is called a companded DPCM. Since a DM has only two quantization levels in the quantizer, there is no companded DM and the channel bit encoder is placed parallel to the quantizer to encode the difference signal D_k directly in actual hardware implementation.

If the quantization levels of a quantizer remain fixed during the entire period of coding process, it is a nonadaptive quantizer. If the space between the quantization levels of a quantizer varies according to the statistical properties of the input signal, knowing which quantizer slots were occupied by the previous samples, it is an adaptive time-varying quantizer [20-23]. For a DM, it is called an adaptive DM. For a DPCM, it is called an

adaptive DPCM.

Linear DM and linear DPCM have limited dynamic ranges in a sense that significant signal to noise ratio can be obtained only if the instantaneous input signal power does not vary beyond a very narrow range. Adaptive quantizer can extend the coding dynamic range by adapting the quantization level according to the input signal characteristics. Therefore, the feedback predictive encoders for video signals currently being researched are the companded DPCM, adaptive DPCM, and the adaptive delta modulators.

1.3.2 The Predictor

In a horizontally scanned video system, the predicted value of the present pixel, $P_{m,k}$, at the coordinate (m,k) as shown in Fig. 1.4 can be obtained by multiplying a weighting coefficient, α, to the predicted value, $P_{m,k-1}$, of the previous pixel,

$$P_{m,k} = \alpha P_{m,k-1} \tag{1.1}$$

This is the simplest kind of predictor. For a DPCM, α is the prediction coefficient calculated from the correlation coefficients of pixels [24,25]. For a DM, it is usually the value '1' unless a leaky integrator is desired where α is a value slightly less than 1.

When more than one pixels in a video frame are involved

in the prediction scheme, the predictor is called a high order predictor [26]. If these pixels are positioned along one scanning line, it is an one-dimensional predictor. If pixels of the previous lines of the same frame are utilized, it is called an intraframe two dimensional predictor. An interframe multi-dimensional predictor is the one when pixels of the present frame as well as that of the previous frames are used in the prediction scheme.

Intraframe DPCM [27-32] and interframe DPCM [33-41] have been extensively studied. However, researches for intraframe DM and interframe DM are only in their beginning stages. This dissertation investigates the practical applications of the one-dimensional and the two-dimensional adaptive delta modulation for the encoding of video signals.

1.4 The Video Applications Of Delta Modulation

The delta modulator started out as a very simple analog device [42]. As the step size adaptive algorithms became more sophisticated and the digital logic circuitry became capable of operating at higher speed, the delta modulator evolved into a complex and high speed digital system. The future of delta modulator in the video field is very bright. By far, delta modulator is still the least complex system for video applications. Although DPCM color

television system has been demonstrated successfully by NEC [40] and COMSAT Laboratories [41], the much simpler delta modulator system is going to get more attention because of its many attractive features:

- There is no need for complex computations. Some adders and shift registers are sufficient to generate all the adaptive step sizes.
- There is no need for multiplexing, buffering or parallel to serial conversion of the output digits.
- It does not require A/D converter and sample-and-hold device. At high speed operation this is a tremendous savings in component costs.
- A DM can be implemented on a single PC board. This means, simple structure, low cost, space saving, and power saving.

The delta modulation for colored video signals is currently being investigated by Dr. Schilling at The City College of New York. In this dissertation, only black-and-white video signals are considered. Three black-and-white slides are used as the video signal sources. They are named the "Girl", the "Boy", and the "Test Pattern".

The video system employed in our laboratory as described below is shown in Fig. 1.5. The video signal is generated from a commercial quality television camera band-limited to 4 MHz. The frame rate of the camera is

30 frames/sec. One complete frame is taken and stored in a scan converter that was employed to slow down the video signal so that it can be processed in a non-interlaced mode by a PDP-8 computer that was used to simulate the DM algorithms. The output from the computer is stored in a second scan converter. After a frame of video signal is processed, the second scan converter is switched to the video mode to display on a TV monitor at the video scanning rate the processed picture that was just stored. Subjective judgement of the quality of the encoded pictures is done by looking at the monitor screen directly or by comparing the photographs taken from the monitor screen.

Some degradations have occurred as the video signal passed through the video system, PCM encoded and displayed on the monitor for photographing. Figure 1.6 through Fig. 1.8 show the direct-from-camera pictures and the through-system 10-bit PCM encoded pictures. The PCM encoded pictures will serve as the "original" pictures to be compared by the DM encoded pictures for subjective evaluations. There is a diagonal flash across every picture, some are clearly noticeable, some are not. It is the photographic effects, and should be distinguished from other encoder generated distortions.

1.5 Summary For The Dissertation

Multi-dimensional DPCM algorithms have been studied extensively, especially at the Bell Telephone Laboratories for the application of Picturephone [28-30,32-38]. However, multi-dimensional delta modulators are believed to be first proposed by the author and his colleagues [31-33]. Part of the work which deals with two dimensional intraframe delta modulators are discussed in detail in this dissertation. In Chapter 2, one dimensional delta modulators are discussed, starting with the linear DM in section 2.2, followed by the adaptive delta modulators in section 2.3. The Abate [61] mode ADM and the Song [62] mode ADM are chosen to demonstrate the ability and the shortcomings of the one dimensional DM in video signal encoding. Values of the output signal-to-noise ratio are shown for the objective measurement. The encoded pictures are also shown for the purpose of subjective comparison.

Two new types of ADM, the A-mode ADM and the B-mode ADM are introduced. They are discussed in sections 2.3.3 and 2.3.4 respectively. The A-mode ADM and the B-mode ADM algorithms concentrate in reducing slope overload and edge busyness for a more subjectively pleasing video image. As a result, their signal-to-noise ratios are also increased by up to 3 dB. Frame difference, phase shift, and image distortion are also discussed in Chapter 2.

In Chapter 3, two dimensional delta modulators are discussed, starting with an introduction describing the reasons for using two dimensional encoding, followed by several two dimensional DM algorithms; the Normal Mode 2DDM, the Look-Ahead 2DDM, the 3-state 2DDM, the Weighted-Average 2DDM and the Prediciton 2DDM. The advantages and disadvantages of each algorithm are discussed and compared. Values of SNR and encoded pictures are shown. A conclusion is given in the last section which compares all the DM algorithms described in Chapter 2 and Chapter 3.

In Chapter 4, we discussed the entropy encoding techniques for coding the DM output bit stream. The Huffman code and the run length code are the two most applicable code for the DM output bit stream. Biased 2DDM is discussed; in which long runs of directional bits can be obtained for more efficient run length coding.

The channel error responses of the DMs are discussed in Chapter 5. Two dimensional delta modulators have different error response as that of the one dimensional delta modulators. Biased 2DDM has the best error response in the two dim nsional DM category and the Weighted-Average 2DDM is the worst. The resulting decoded pictures are shown for different DM algorithms with and without leaky integrators.

This dissertation is concluded in Chapter 6 which summarizes the research work. Appendix and references are

included after Chapter 6 to complete this dissertation.

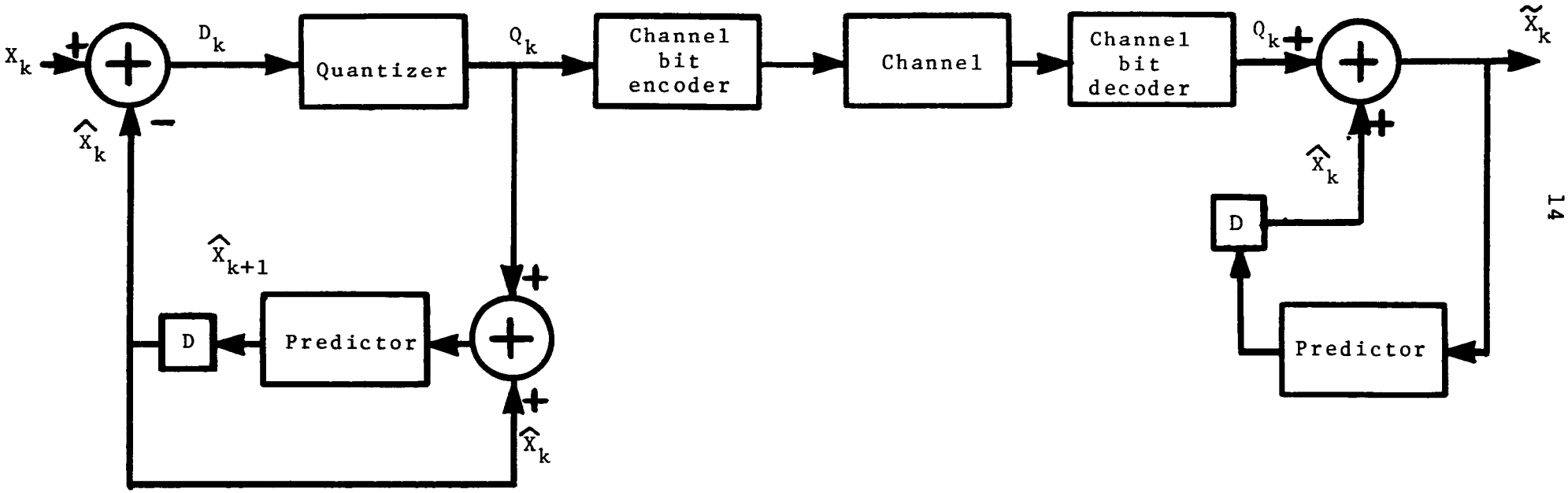


Fig. 1.1 Block diagram of feedback predictive coding system.

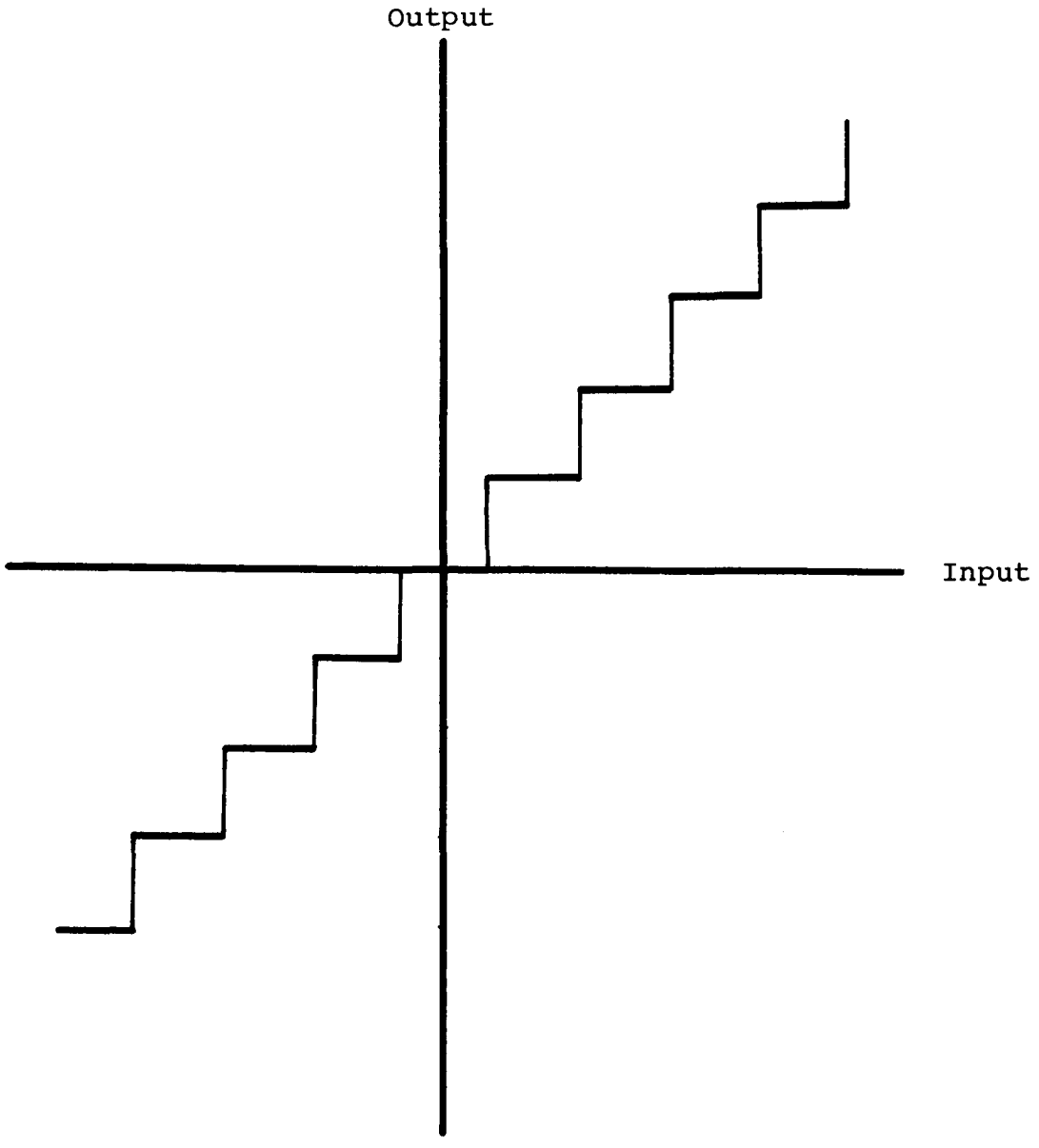


Fig. 1.2 A linear quantizer.

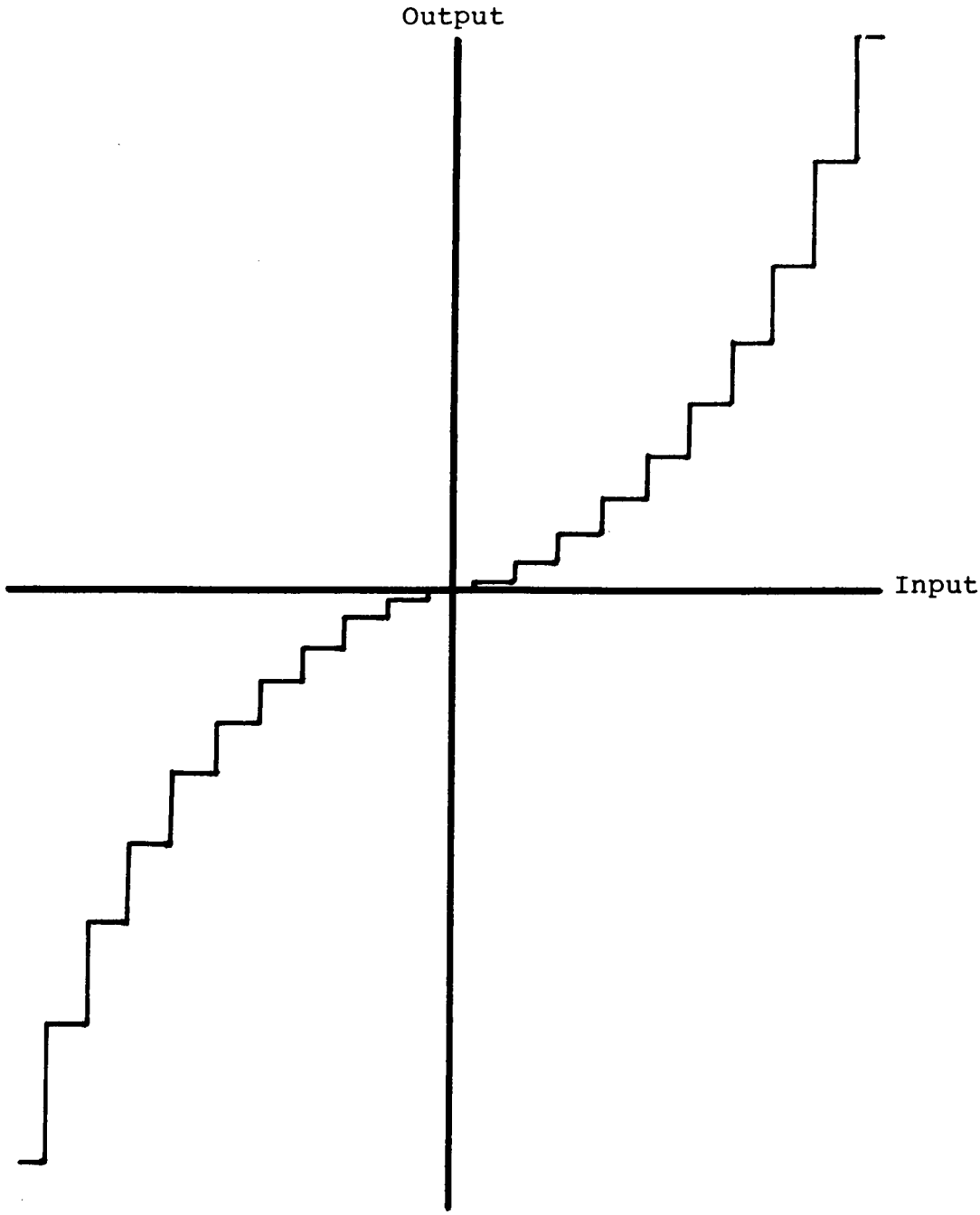


Fig. 1.3 A nonlinear quantizer.

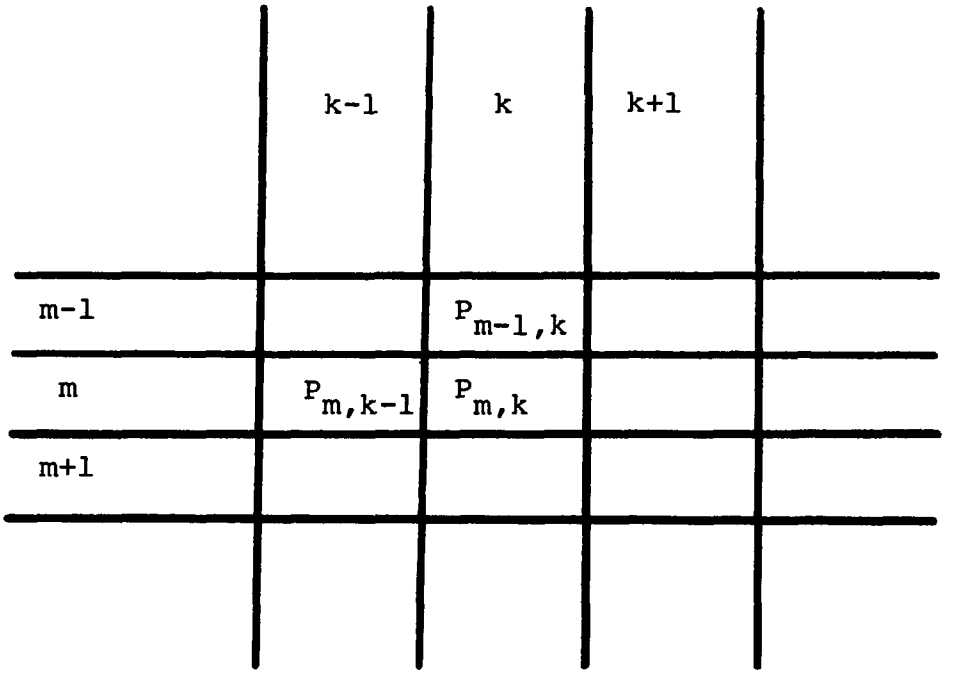


Fig. 1.4 The relative pixel positions.

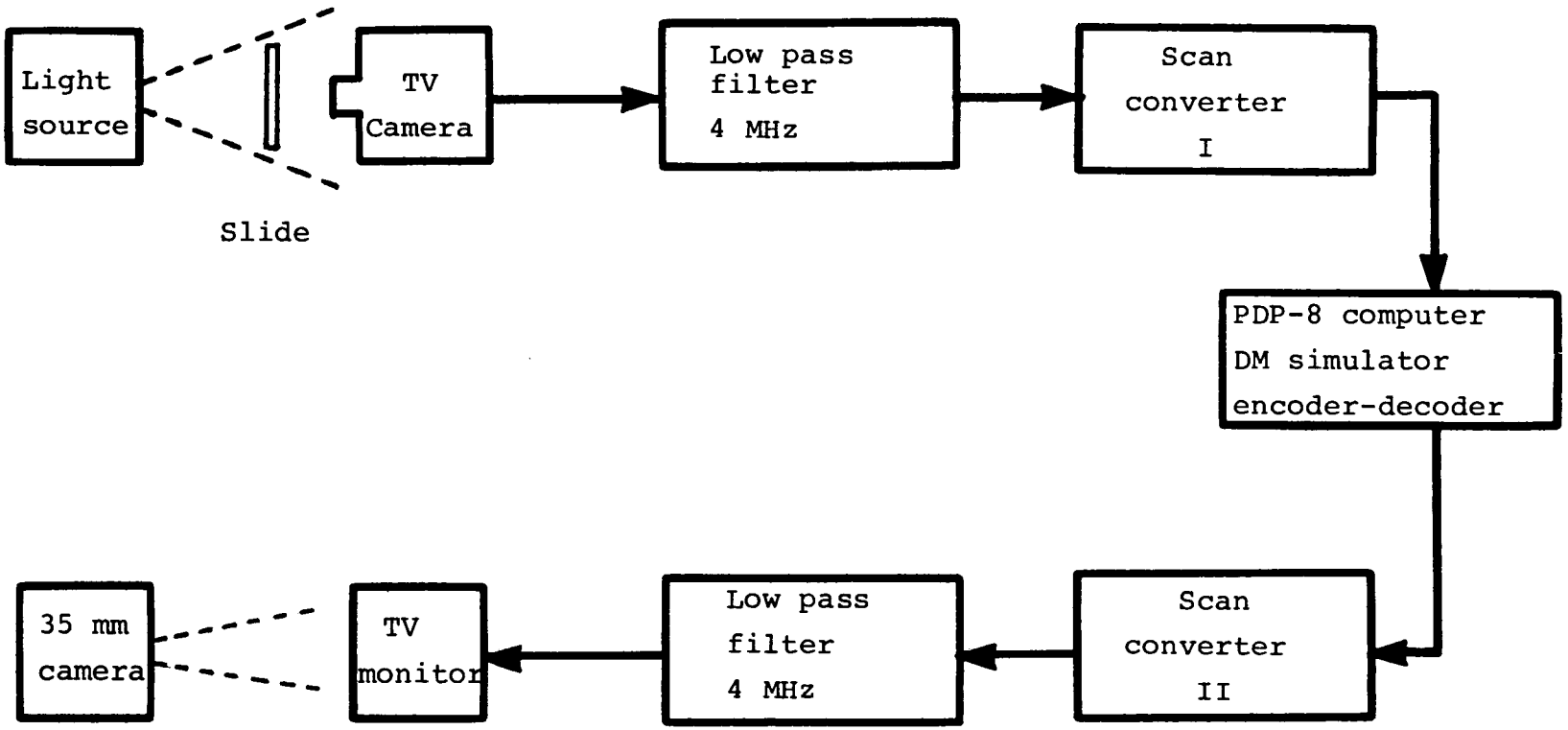


Fig.1.5 The video system employed to test DM algorithms.

(a)

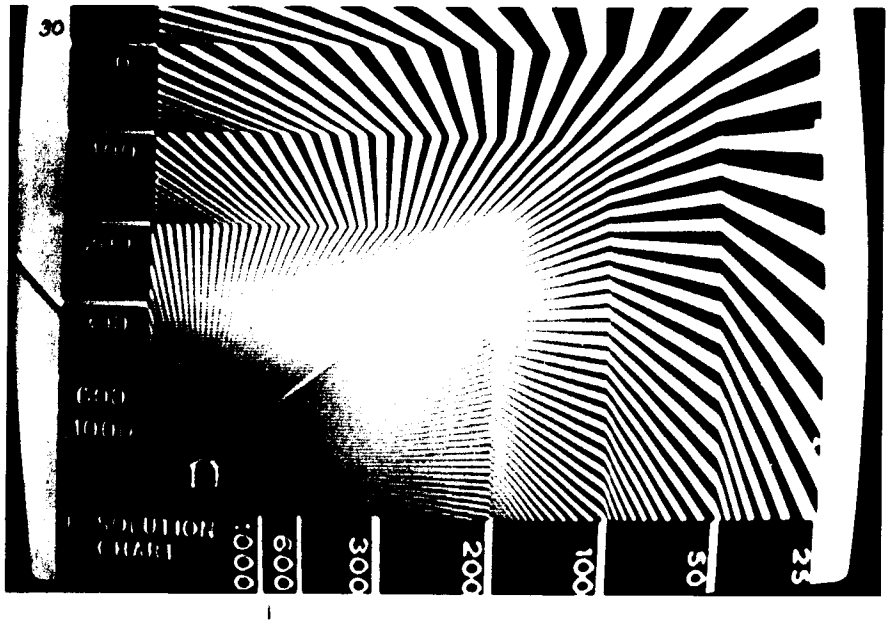
(b)

Fig. 1.6 (a) The analog picture, (b) the through-system, 10-bits PCM encoded picture.

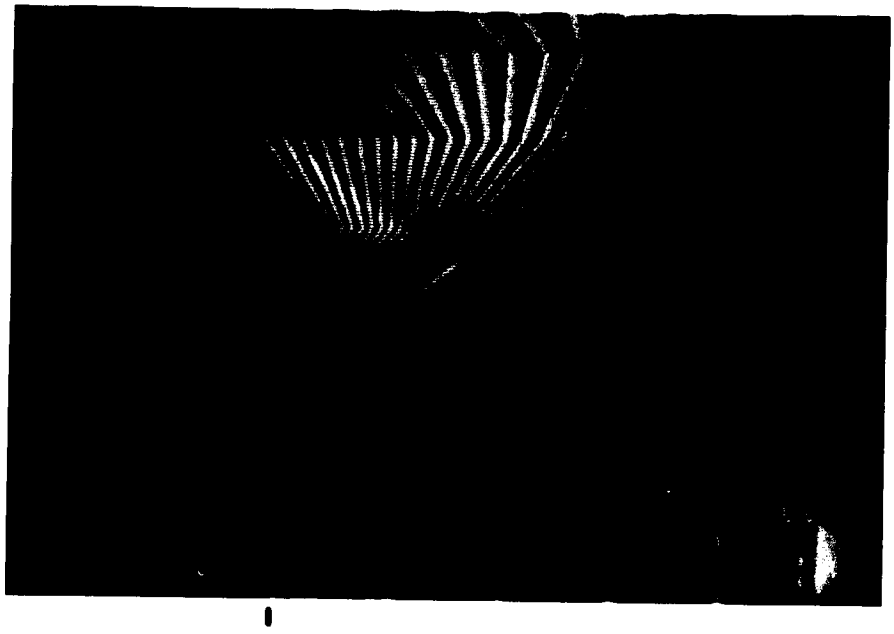
(a)

(b)

Fig. 1.7 (a) The analog picture, (b) the through-system, PCM encoded picture.



(a)



(b)

Fig. 1.8 (a) The analog picture, (b) the through-system, 10-bits PCM encoded picture.

Chapter II

ONE DIMENSIONAL DELTA MODULATORS

2.1 Introduction

One dimensional delta modulation is a very simple and effective method to encode video signals. Depending on the algorithm used, a one dimensional delta modulator (1DDM) is capable of transmitting high quality images through a digital channel occupying a narrow bandwidth. Low cost, simple circuit construction and good channel error response are other attractive features of the one dimensional delta modulator. A laboratory model has been built at The City College of New York. It operates at up to 22 MHz sampling rate using the Song mode adaptive step size algorithm. The laboratory model has successfully demonstrated that under limited bandwidth and limited power conditions, one dimensional delta modulators can achieve satisfactory results.

In this chapter a brief discussion on linear DM, Abate mode DM, and Song mode DM will be given, followed by a detailed description on two new types of 1DDM. Tables and graphs of output signal-to-noise ratio (SNR) and the dynamic range of various kinds of 1DDM are shown. The encoded pictures, error pictures and frame difference

pictures are also included to provide a comparison for a subjective judgement.

The laboratory simulated DM for this dissertation is assumed to have 128 distinct quantization levels. Under normal condition of 0 dB relative input power, the input waveform is adjusted to just fill up the 128 quantization levels without saturation. At -6 dB relative input power, the signal voltage only fills up to 64 quantization levels. At +6 dB relative input power, the signal voltage swings within a range of 256 quantization levels. In this case, the simulated DM opens the window to allow 256 quantization levels to avoid saturation. The quantization step is always kept at a constant and is called the minimum step size, S_0. The sampling rate is 1024 samples per line, which corresponds to twice the Nyquist rate for a NTSC video formatted picture bandlimited to 4 MHz. It is described as sampling at 2 bits per picture element (pixel). The scanning lines are not interlaced. All storage registers are reset to zero at the beginning of each frame. Unless otherwise specified, all the values of SNR and the encoded pictures are taken under the conditions of 2 bits/pixel, 128 quantization scale and noninterlaced scanning system.

2.2 Linear Delta Modulator

The simplest kind of delta modulator is a linear DM.

Linear DM first appeared in a French patent issued in 1946. Further studies on linear DM have been done by deJager and many others [42-48]. Figure 2.1 is the block diagram of a digital linear DM. Figure 2.2 shows how a linear DM is operated to encode an incoming signal. Let E_k denote the k-th DM error sign bit and X_{k+1} denote the next DM estimate to be generated after the k-th clock pulse, the linear DM algorithm can be represented as

$$E_k = \mathrm{Sgn}[M_k - X_k] \tag{2.1}$$

$$X_{k+1} = X_k + SE_k \tag{2.2}$$

where M_k is the k-th input sample,

X_k is the present DM estimate at the k-th clock pulse,

S is the step size of the linear DM.

The advantage of a linear DM is simple and low cost. The disadvantage of a linear DM is its narrow dynamic range. The computer simulation of a linear DM shows that at its optimum step size, relative to the input signal power, the linear DM is able to achieve good output signal to noise ratio, (Table 2.1). However, a slight variation in the step size or a slight change in the input signal power will result in a sharp decrease in SNR as shown in Table 2.2, Fig. 2.3, and Fig. 2.4. A step size could be too small for

a particular picture, thus, causing a severe slope overload situation in the encoded picture. However, the same step size could be too large for the next picture, thus, causing a large amount of quantization noise.

It can be seen from Fig.2.4 and Table 2.3 that the optimum step size of a linear DM for the "Boy" at 0 dB input power is 6 S_0. The output SNR is 23.7 dB as shown in Table 2.1. It is comparable to the SNR of any other adaptive lDDM schemes shown in the same table. Each of the linear DM encoded pictures shown in Fig. 2.5 and Fig. 2.6 was encoded with its optimum step size. Except some visible granular noise and certain degree of smearing along the edges, the optimized linear DM can encode video signals quite well. However, the lack of dynamic range makes the linear DM not practical for television applications. Figure 2.7 shows the extreme case when the "Test Pattern" is encoded by a linear DM with a step size of 3 S_0. Severe distortion and smearing effects are clearly visible.

2.3 Adaptive Delta Modulators

Many different step size adaptive algorithms have been proposed and implemented to increase the dynamic range of a delta modulator since Winkler [49] first presented his version of adaptive algorithm in 1963. Among these algorithms [50-65] only two will be discussed

in this dissertation, they are the Abate mode and the Song mode algorithms. Both of them are very effective and can be easily implemented with digital circuits.

2.3.1 Abate Mode Adaptive Delta Modulator

Abate [61] first presented his idea of discrete instantaneous step size adaptive algorithm in 1967. Later studies by Song, Garodnick and Schilling [62] showed that video signal with a first order Markov Gaussian characteristics should be encoded by an adaptive DM whose step size increases and decreases in a nonlinear manner. The Abate mode adaptive delta modulator (ADM) changes the step size linearly and, therefore, is not suitable for video encodings. However, experimental studies are still carried out and included in this chapter for comparison purposes.

Let E_k be the k-th error sign bit, S_k be the k-th step size and X_{k+1} be the new DM estimate to be generated after the k-th clock pulse, the Abate mode algorithm can be represented as follows:

$$E_k = \text{Sgn}[M_k - X_k] \tag{2.3}$$

$$S_k = |S_{k-1}| \; E_k + S_0 E_{k-1} \tag{2.4}$$

$$X_{k+1} = X_k + S_k \tag{2.5}$$

where M_k is the k-th input sample,

X_k is the present DM estimate at the k-th clock pulse,

S_{k-1} is the (k-1)th step size,

S_0 is the minimum step size,

E_{k-1} is the (k-1)th error sign bit.

Abate mode ADM is good for speech applications [61-64]. However, its slow rising and slow settling characteristics limit its ability for encoding video signals. The excessive slope overload noise and edge ringing noise make the SNR of Abate mode ADM consistantly worse than that of the Song mode ADM (see Table 2.1). The SNR as a function of input signal power is listed in Table 2.4 and is plotted in Fig. 2.3. Figure 2.3 shows that the DM encoding dynamic range is expanded by using adaptive step size algorithms. The computer simulation shows that the Abate mode ADM can produce acceptable quality pictures as shown in Fig. 2.8. The slope overload noise cannot be easily seen from the encoded pictures, so the difference picture between the original picture and the encoded picture is taken as shown in Fig. 2.9 with an amplification factor of 2. It can be seen that the encoding noise is mainly the slope overload noise concentrated along the edges.

2.3.2 Song Mode Adaptive Delta Modulator

Song et al. [62,63] have shown that assumming the sampled video signal is a first order Markov sequence with a Gaussian amplitude distribution, an exponentially adapted step size is the optimum algorithm in the mean square error sense for a one dimensional delta modulator to encode video signals. Similar result was also obtained by Jayant [65]. They both showed that wide dynamic range and high SNR can be achieved by adapting the step size of a delta modulator exponentially. The exponential characteristics of the Song mode ADM is shown in Appendix 2.

Since the Song mode ADM is the known existing highly efficient one dimensional DM, the results shown in this section will be used as a guide-line to be compared by other one dimensional and two dimensional algorithms discussed in this dissertation.

In the simplified version of the Song mode ADM [66], the step size, S_k, is defined as

$$S_k = |S_{k-1}| \ (E_k + \frac{1}{2} E_{k-1}). \qquad (2.6)$$

The step size varies by a factor of 1.5 or 0.5 depending on the two most recent error sign bits, E_k and E_{k-1}. The functional block diagram and the timing diagram of the Song mode ADM are shown in Fig. 2.10 and Fig. 2.11 respectively.

The step size of the Song mode ADM grows exponentially. If no maximum allowable value is imposed upon the step size, it could be so large that a large amount of overshoot would occur. Restrictions on the maximum allowable step size will slow down the response time of the Song mode ADM and, thus, increase the slope overload noise. However, the resulting smaller overshoot makes the overall SNR better than that of the nonrestricted ADM. Table 2.5 shows that the SNR is highest when the maximum step size, S_{max}, is 16 S_0. In the case of a smooth picture, such as the "Girl", even a S_{max} of 8 S_0 will do as good as that of 16 S_0. The resulting visual effect is that with a smaller S_{max}, the picture appears softer with a larger S_{max}, the vertical edges are sharper but with more ringing and edge busyness.

In order to understand more about the effect of the limitation on the step size, the statistics of the step size are shown in Table 2.6 and Table 2.7. They are also depicted in Fig. 2.12. One can see that most of the step size values are smaller than 16 S_0 with the exception of the "Test Pattern". The busy black and white stripes of the "Test Pattern" requires the step size to grow beyond 32 S_0 to catch the steep input signals.

In order to be more realistic and practical, the step size algorithm used in this simulation study does not

vary exactly according to Eq.(2.6). Because of the finite length of shift registers in a DM, truncation is necessary in the step size calculations. The actual growth of the step size from S_0 is a sequence,

$$S_0,\ 2S_0,\ 3S_0,\ 4S_0,\ 6S_0,\ 9S_0,\ 13S_0,\ 19S_0,\ 28S_0,\ldots$$

Therefore, certain integer values will never be assigned to the step size, as can be seen from Table 2.6 and Table 2.7, no matter how many times the multiplication of 1.5 or 0.5 in Eq.(2.6) has been carried out.

In order to gain more insight into the DM operation, the quantized value of the input video signal is studied. The video signal voltage distribution of the three source slides, the "Girl", the "Boy" and the "Test Pattern" are shown in Fig. 2.13 in a scale of 128 quantization levels. The high concentration of the voltage distribution at the lower end of the "Girl" is due to the broad dark background in the picture. The two humps at both end of the voltage distribution of the "Test Pattern" is because the picture consists of mainly black and white stripes. The "Boy" has a relatively even distribution and is considered a typical picture. It agrees with the data given by Kretzmer [6].

Since the "Boy" gives a typical video signal, the measurement of SNR in this dissertation will be based on this slide. The SNR of the other slides will sometimes be

shown if these data present some differences and are worth mentioning. However, the pictures shown will mainly be the "Girl". This is because that the "Boy" and the "Test Pattern" have too much details, thus, are difficult to distinguish among pictures encoded by different algorithms. Moreover, in order to observe the effects of slope overload and edge busyness, long and clearly defined vertical edges are required. The hat and the hand of the "Girl" are the perfect objects for this purpose.

The Song mode ADM encoded pictures are shown in Fig. 2.14 and Fig. 2.15. Comparing Fig. 2.14(a) and Fig. 2.8, one can see that because of the faster rising and settling characteristics of the Song mode ADM, the edge busyness along the edges are shorter than that of the Abate mode ADM.

The intrinsic slope overload characteristics of a DM produces phase shift in the encoded waveforms. When the video signal is sampled at low rate, the input samples can vary in amplitude by a large amount. The resulting encoding phase shift can be several pixels in distance as shown in Fig. 2.16. This creates picture distortion. The gate-like input signal in Fig. 2.16 has three gates, A, B, and C, separated by distances a and b as indicated. After encoding, the gates are separated by distances a' and b', which are not the same as a and b. In other words, the waveform is distorted not only in shape

but also in relative positions. The entire encoded picture is compressed or stretched at different areas depending on the signal amplitude of the neighboring pixels. The noise measurement based on the pixel amplitude difference between the original and the encoded picture consists of distortion noise, overshoot noise, phase shift noise and any other discrepancies between the original and the encoded pictures. However, human eyes are not sensitive to the overall phase shift of the encoded pictures, so the overall phase shift noise should be excluded from the total noise power. The shaded areas in Fig. 2.16 represent the total encoding noise. The overall phase shift noise can be removed by sliding the encoded waveform to the left to a proper position as shown in Fig. 2.17 before the SNR calculation is taken. However, the distortion noise cannot be removed by this sliding. In Fig. 2.17, the phase shift noise is minimized at gate A and gate C by shifting to a proper position, but the noise power is increased at gate B. This is due to the encoding distortion which causes a' to be less than a and b' to be greater than b. Table 2.8 lists the SNR values with the encoded waveform shifted to different positions. They are also illustrated in Fig. 2.18. The SNR used throughout this dissertation will be the one which gives the highest value. It is usually the one with a one pixel position shift toward the

left. We can conclude, by drawing a smooth curve connecting the discrete points in Fig. 2.18, that the highest SNR can be obtained for the "Girl" if the encoded waveform if shifted to the left by one half of a pixel distance before the noise power measurement is taken. Similarly, the encoded waveform should be shifted to the left by one pixel distance for the "Boy", and one and one half pixel distance for the "Test Pattern" in order to obtain the highest and the most meaningful SNR. Unfortunately, only integer number of pixel distance can be shifted in the simulation set-up. Therefore, the maximum points of the curves in Fig. 2.18 cannot be verified.

In order to filter the out-of-band noise, a simple 2-bit average subroutine is added to the main simulation program. The results are listed in Table 2.8. This 2-bit average scheme has another effect, that is, the encoded waveform if automatically shifted to the left by one half of a pixel distance. Table 2.8 shows that the SNR's of the ADM with the 2-bit average low pass filter are better than those without low pass filter, but not better than those with the proper shift. Hence, we can conclude that the 2-bit average scheme cannot improve all of the SNR values, and that the proper shift in position has a greater effect in SNR than filtering. The SNR values shown in this dissertation are, therefore, not filtered but

shifted into proper position for the best results.

Figure 2.19(a) is the difference picture of the Song mode ADM magnified by a factor of 2 without shifting the position of the encoded picture. One can see the slope overload noise along the edges clearly. Figure 2.19(b) is the same difference picture but the encoded picture is shifted to the left by one pixel distance. The slope overload noise is reduced but the distortion noise causes the contour of the girl to be clearly visible. All the difference pictures shown in the subsequent sections are magnified by a factor of 2 for the reason of clarity and are shifted to the left by one pixel distance unless otherwise specified.

Figure 2.20(a) is the difference picture of the "Boy" without shifting in positions. Figure 2.20(b) is the same difference picture with a phase shift of one pixel to the left. Again, as mentioned before, a diagonal flash across the picture is caused by some unknown photographic effects.

The dynamic range of the Song mode ADM has been appeared in literatures [62,63]. It is also shown in Table 2.4 and Fig. 2.3 based on our laboratory simulations.

Look-Ahead Song Mode ADM

At every clock pulse, the ADM will make a decision to select one of the two possible values to be the next

DM estimate. The decision is based on the comparison of the present input sample and the previous DM estimate. The lack of ability to look ahead for the future events makes the ADM overshoot very often. A little amount of overshoot is desirable because it enhances the presence of edges. But large overshoots should be avoided [67,68]. The Look-Ahead scheme [69-71,73,74] as shown in Fig. 2.21 will eliminate the overshoot problem. A Look-Ahead ADM generates the two possible next DM estimates, X^{+}_{k+1} and X^{-}_{k+1}, before the k-th clock pulse arrives. At the clock pulse the input sample M_k compares with X^{+}_{k+1} and X^{-}_{k+1}, instead of the present DM estimate X_k, to decide whether X^{+}_{k+1} or X^{-}_{k+1} is the better representation of M_k. Fig. 2.21 shows how a large overshoot is avoided at the k-th sample by using the Look-Ahead scheme. The SNR is consequently improved. Table 2.9 shows the SNR's of the Song mode ADM with and without the Look-Ahead scheme.

Although the Look-Ahead scheme solves the overshoot problem of the ADM, other problems associated with the ADM such as slope overload and edge busyness are still hampering the video applications of adaptive delta modulator. In the following sections we will present different ADM algorithms to solve or to ease these problems.

2.3.3 A-Mode Adaptive Delta Modulator

The exponentially rising characteristics of the Song mode ADM in response to a step input will cause large overshoot and edge busyness if the step size is not limited. There are two ways to limit the step size. One is to confine the step size to be less than or equal to a predetermined threshold value as was described in the last section. The other is to control the growth of the step size, S_k, so that it asymptotically approaches a predetermined maximum value, S_{max}, as described by

$$S_k = [\ |S_{k-1}| + \frac{S_{max} - |S_{k-1}|}{c}\ ]E_k \qquad (2.7)$$

where c is a constant to be determined. When a slope overload condition is detected, that is, when three consecutive DM error sign bits are of the same sign, the step size will increase according to (2.7) by the amount $(S_{max} - |S_{k-1}|\ )/c$ at every sample. Equation (2.7) shows that the step size increment at each sample is decreased from the previous increment until the maximum step size is reached. For instance, when the previous step size is the minimum value, S_0, i.e.,

$$|S_{k-1}| = S_0$$

$|S_k|$ equals

$$S_0 + \frac{S_{max} - S_0}{c} \quad .$$

The increment is equal to

$$\frac{S_{max} - S_0}{c}$$

which is the maximum possible increment in step size. When the previous step size is already very large, i.e.

$$|S_{k-1}| \cong S_{max} \quad .$$

$|S_k|$, from Eq.(2.7), will approximately equal to S_{max}. The increment in step size is very small. In other words, the step size is self-limited to the value, S_{max}.

The constant, c, is determined experimentally with the consideration that the largest increment of step size, which occurs when the video signal is not actively varying, should be small enough not to cause too much granular noise and contour noise, and yet be large enough to provide a fast response in case an edge is detected. The DM response, to a step input, of the Song mode ADM with and without S_{max} and the A-mode ADM are illustrated in Fig. 2.22.

The complete step size algorithm for the A-mode ADM is

$$S_k = [|S_{k-1}| + \frac{S_{max} - |S_{k-1}|}{c}] E_k, \text{ if } E_k=E_{k-1}=E_{k-2} \quad (2.8(a))$$

$$S_k = \begin{cases} \frac{1}{2}|S_{k-1}|E_k & \text{for } S_{k-1} \neq S_0 \\ S_0 & \text{for } S_{k-1} = S_0 \end{cases} \quad \text{if } E_k \neq E_{k-1} \qquad (2.8(b))$$

$$S_k = [\,|S_{k-1}| + S_0]\, E_k \qquad \text{if } E_k = E_{k-1} \neq E_{k-2} \qquad (2.8(c))$$

where S_0 is the minimum step size. Eq.(2.8(b)) describes the step size adaptation when the DM error sign bits are alternating. The step size is reduced to one half of the previous step size value except when the previous step size is already the minimum step size, S_0. In this case, the step size is unchanged and kept at S_0. Eq.(2.8(c)) applies when an overshoot has just occured. Eq.(2.8(c)) will ensure the DM estimate to have a quick recovery from the overshoot and a fast settling toward a steady state. The purpose of adding a small value, S_0, in Eq. (2.8(c)) is to avoid a possible large amplitude oscillation right after an overshoot has occurred due to the truncation effect when an odd number is divided by 2 in Eq.(2.8(b)). This is illustrated in Fig. 2.23.

The relation between the initial step size, S_i, and the step size k samples later, S_{i+k}, of the A-mode ADM in response to a step input can be described by

$$S_{i+k} \cong S_{max} - (S_{max} - S_i)\, e^{-k/c} . \qquad (2.9)$$

Equation (2.9) is only an approximation based on the first order approximation of a sequence to an exponential function which will be discussed later.

The derivation of Eq.(2.9) is shown below:

From Eq.(2.7),

$$S_{i+1} = S_i + \frac{S_{max} - S_i}{c}$$

$$= \frac{S_{max}}{c} + (\frac{c-1}{c})\ S_i \ .$$

The next step size, S_{i+2}, is

$$S_{i+2} = S_{i+1} + \frac{S_{max} - S_{i+1}}{c}$$

$$= \frac{S_{max}}{c} + (\frac{c-1}{c})\ [\ \frac{S_{max}}{c} + (\frac{c-1}{c}) S_i\ ]$$

$$= \frac{1}{c}\ (1 + \frac{c-1}{c})\ S_{max} + (\frac{c-1}{c})^2\ S_i \ .$$

The step size at k samples later is

$$S_{i+k} = \frac{1}{c}\ [\ 1 + \frac{c-1}{c} + (\frac{c-1}{c})^2 + \ldots + (\frac{c-1}{c})^k\ ]\ S_{max}$$

$$+ (\frac{c-1}{c})^k\ S_i$$

$$= S_{max} - (\frac{c-1}{c})^k\ (S_{max} - S_i) \ . \qquad (2.10)$$

An exponential function can be written as

$$e^{-1/c} = 1 - 1/c + 1/2c^2 - 1/6c^3 + \ldots \qquad (2.11)$$

If $c \gg 1$, using first order approximation, we obtain

$$e^{-1/c} \cong 1 - 1/c \ . \tag{2.12}$$

Substituting (2.12) into (2.10) gives

$$S_{i+k} = S_{max} - (S_{max} - S_i)e^{-k/c} \ . \tag{2.9}$$

The first order approximation of (2.12) is valid because in our DM simulation, the value of c is found to be far greater than 1 (Table 2.10).

Equation (2.9) is a bised exponential curve, starting at S_i gradually approaching S_{max} with a time constant, c. It is plotted in Fig. 2.24.

The DM estimate at k samples later starting with a initial estimate, X_i, can also be approximated by

$$X_{i+k} \cong X_i + [kS_{max} - c(S_{max} - S_i)(1 - e^{-k/c})] . \tag{2.13}$$

It is readily derived as follows,

$$X_{i+1} = X_i + S_i$$

$$X_{i+2} = X_{i+1} + S_{i+1}$$

$$= X_i + S_i + S_{i+1}$$

$$\cdots$$

$$X_{i+k} = X_i + \sum_{j=0}^{k-1} S_{i+j} \ .$$

From Eq.(2.10) we get,

$$X_{i+k} = X_i + \sum_{j=0}^{k-1} [S_{max} - (\frac{c-1}{c})^j (S_{max} - S_i)]$$

$$= X_i + kS_{max} - (S_{max} - S_i)\, c[1-(\frac{c-1}{c})^k].$$

using the first order approximation of Eq.(2.12), we obtain

$$X_{i+k} \cong X_i + kS_{max} - c(S_{max} - S_i)(1 - e^{-k/c}). \qquad (2.13)$$

The first two terms on the right hand side of (2.13) represent a straight line with an initial value, X_i, and a slope, S_{max}. The rest of the right hand side of (2.13) is an exponential curve. Equation (2.13) is plotted in Fig. 2.25. It can be called a modified exponential curve, starting with an exponential characteristics and gradually modifying itself to approach a straight line with slope, S_{max}. This kind of response of DM estimate is faster than that of the Song's true exponential response in the beginning of the curve as illustrated in Fig. 2.22. This gives the advantage of faster rise-up, such that for a small to medium height edges it has less slope overload noise. However, for a large step input signal, this modified exponential curve gives a slower response, thus, avoids excessive overshoot noise. This is a trade-off between the slope overload noise and the overshoot noise.

The increment of the step size at the (i+k)th sample can easily obtained by subtracting the two values of the

adjacent step sizes,

$$\Delta S_{i+k} = S_{i+k} - S_{i+k-1}$$

$$= [S_{max} - (S_{max} - S_i)(\frac{c-1}{c})^k]$$

$$- [S_{max} - (S_{max} - S_i)(\frac{c-1}{c})^{k-1}]$$

$$= (\frac{S_{max} - S_i}{c})(\frac{c-1}{c})^{k-1}$$

$$\cong (\frac{S_{max} - S_i}{c})\, e^{-(k-1)/c} \quad . \qquad (2.14)$$

Equation (2.14) is an exponentially decaying curve as shown in Fig.2.26. The constant, c, is the time constant of the exponential function of Eqs. (2.9), (2.13) and (2.14). The selection of the value of c is, therefore, based on the requirement of the speed of the step size to approach its maximum value.

The values of SNR of the A-mode ADM are shown in Table 2.10. Other values of c have also been studied. The results are very close to those shown in Table 2.10. However, any value of c other than those of the power of 2 will present difficulties in hardware realization. Therefore, they are not favorable choices and are not listed in Table 2.10. The choice of $S_{max} = 16$ and $c = 4$ seems to be better for all classes of video signals. The encoded picture and the difference picture are shown in Fig. 2.27(a) and Fig.2.27(b),

respectively. Although the measured SNR is in favor of the A-mode ADM over the Song mode ADM, the subjective improvement of the encoded pictures are too small to be noticed.

We found out later in our research that the rising characteristics of the A-mode ADM in response to a step input is the same as the Continuous Variable Slope DM (CVSD) algorithm. It is described in detail in Appendix 3.

2.3.4 B-Mode Adaptive Delta Modulator

The quantization noise of a delta modulator encoded video signal consists of two noise categories. One is the granular noise which distributes over the entire encoded picture. Its amplitude is low and so is the local noise power. But the total granular noise power in a frame of video signal contributes a significant amount toward the total quantization noise power. The other type of noise is the slope overload and overshoot noise. It appears only along clearly defined edges. Its local noise power is high. The high concentration of noise power in a small area together with the effect of "edge busyness" make this type of noise particularly annoying to human eyes. The adaptive delta modulation algorithms proposed by Abate, Song, and Jayant etc. can all encode video signals with high performance. However, they all have one thing in common,

that is, their adaptive schemes employed to minimize the total noise power did not treat the two noise categories separately. However, human psycho-visual effects suggested that the subjective quality of an encoded picture can be increased if special effort is incorperated into the adaptive algorithm to reduce the slope overload noise, thereby reduces edge busyness.

In this section, a new type of step size adaptive algorithm is studied which reduces the amount of slope overload and overshoot noise at the expense of possible increase of granular noise.

When responding to a step input signal, the desirable response of a delta modulator should have the following properties:

(1) The initial rise-up is fast so that the delay between the input signal and the encoded estimate is minimized, thus, less slope overload distortion is produced.

(2) The amount of overshoot is small so that the DM estimate can quickly settle down to a steady state.

(3) The response curves of all scanning lines are synchronized so that the effect of edge busyness is minimized.

The only curve which will fit the above description is a negative power exponential curve shown in Fig. 2.28

as compares to the positive power exponential characteristics of the Song mode ADM. This kind of response has a fast start at the beginning, a smooth approach at the end and a synchronous effect,because, as we will see later, that the step size of this type of ADM is not a function of the previous step sizes.

From the above analysis, a new algorithm is derived. We call it the B-mode ADM algorithm. In this algorithm the DM step size is a function of the present DM estimate. The reason is analyzed below.

For a video signal of 128 quantization levels, the sampled signal voltages are within the maximum amplitude M_p^+ and the minimum amplitude M_p^- as shown in Fig. 2.29. If a positive slope is detected and if the present DM estimate is at position "a" as shown in Fig.2.29, the step size should be large,because position "a" is closer to M_p^- and the estimate has a long way to move to M_p^+. If the present DM estimate is at position "b", the step size should be smaller than that of at "a". It is because that at position "b" the chance of having a contrasty edge, which needs a large step size, is smaller than that of at position "a". Similarly, at position "c", the step size should be very small, because position "c" is very close to M_p^+, the input signal should not be much higher than "c". Therefore, a small step size is a reasonable prediction. The net effect

is a DM estimate response curve having the desirable asymptotic exponential characteristics. Still, there is another problem to be considered, the detection of edges. If there is no means of detecting edges, and if this same step size algorithm is used at every sample, the granular noise will be so large that it will compensate any gain in SNR obtained by the reduction of slope overload. In order to solve this problem, four error sign bits are needed for the generation of the present step size,

$$S_k = \frac{1}{2} |S_{k-1}| E_k \qquad \text{if } E_k \neq E_{k-1} \qquad (2.15(a))$$

$$S_k = (|S_{k-1}| + S_0)E_k \qquad \text{if } E_k = E_{k-1} \neq E_{k-2} \qquad (2.15(b))$$

$$S_k = C_3 S_0 E_k \qquad \text{if } E_k = E_{k-1} = E_{k-2} \neq E_{k-3} \qquad (2.15(c))$$

$$S_k = f(X_k, E_k) \qquad \text{if } E_k = E_{k-1} = E_{k-2} = E_{k-3} \qquad (2.15(d))$$

where C_3 is a constant to be determined,

$f(X_k, E_k)$ is a function of the present DM estimate, X_k, and the present error sign bit, E_k.

Eq. (2.15(a)) describes the adaptation of step size when the DM error sign alternates. A parameter of 0.5 is used as is in the Song mode. When there are two identical error signs in a row but the one before them is different, this is an overshoot condition, and the step size is increased only by a small amount to ensure a quick settling

effect as shown in Eq. (2.15(b)). When there are three error bits with the same sign in a row, it is recognized as the beginning of an edge. A step size of an intermediate value C_3S_0 is used as indicated by Eq.(2.15(c)). When there are four identical error sign bits in a row, it is decided that a contrasty edge must be present. The step size is, thus, largely increased as described by the function $f(X_k,E_k)$ to ensure a fast rise-up.

The rest of the DM encoder's equations are the same as those of the Song mode. The function $f(X_k,E_k)$ will be derived next.

The Mathematical Derivation of $f(X_k,E_k)$

For the purpose of simplicity, we will assume that the amplitude of the video signal is within the range from 0 to M_p^+. Refer to Fig. 2.30, the DM estimate response should fit the following asymptotic exponential equations:

$$X_k = U - (U - X_0)e^{-k/C_2} \qquad \text{if } E_k\text{'s} = +1 \qquad (2.16)$$

$$X_k = -L + (L + X_0)e^{-k/C_2} \qquad \text{if } E_k\text{'s} = -1 \qquad (2.17)$$

where X_0 is the initial DM estimate,

U is the upper limit of Eq.(2.16). It should be greater than M_p^+. The DM estimate response follows the curve of Eq.(2.16) from X_0 up to M_p^+,

C_2 is the time constant to be determined,

-L is the lower bound of Eq.(2.17). It should be less than zero. The DM estimate response follows the curve of Eq.(2.17) from X_0 down to zero.

From Eqs. (2.16) and (2.17), we can derive the step size equation, $f(X_k,E_k)$. From Eq.(2.12), Eq.(2.16) can be approximated as

$$X_k \cong U - (U - X_0)(1 - \frac{1}{C_2})^k \quad . \tag{2.18}$$

From Eq. (2.18)

$$\begin{aligned} S_k &= X_{k+1} - X_k \\ &= [U - (U - X_0)(1 - \frac{1}{C_2})^{k+1}] - [U - (U - X_0)(1 - \frac{1}{C_2})^k] \\ &= (U - X_0)(1 - \frac{1}{C_2})^k[1 - (1 - \frac{1}{C_2})] \\ &= \frac{U - X_0}{C_2}(1 - \frac{1}{C_2})^k \\ &= \frac{U - X_k}{C_2} \end{aligned} \tag{2.19}$$

Similarly, Eq.(2.17) can be approximated as

$$X_k = -L + (L + X_0)(1 - \frac{1}{C_2})^k \tag{2.20}$$

From Eq. (2.20)

$$S_k = X_{k+1} - X_k$$

$$= [-L + (L + X_0)(1 - \frac{1}{C_2})^{k+1}] - [-L + (L + X_0)(1 - \frac{1}{C_2})^{k}]$$

$$= -\frac{(L + X_0)}{C_2} (1 - \frac{1}{C_2})^{k}$$

$$= -\frac{L + X_k}{C_2} \qquad (2.21)$$

In order to shorten computation time and to simplify hardware implimentation, several assumptions are made here. First, it is assumed that the voltage distribution of the input signal sample is symmetrical about its mean value, $\bar{M}$. Second, the input voltage swings from 0 to M_p^+ , and it is assumed that, $M_p^+ = 2\bar{M}$. Based on the above assumptions, the magnitude of U and L must be symmetrical about the mean, $\bar{M}$, and can be written as

$$U = M_p^+ + C_1\bar{M} = 2\bar{M} + C_1\bar{M}$$

$$= (C_1 + 2)\bar{M} \qquad (2.22)$$

and

$$-L = -C_1\bar{M} \qquad (2.23)$$

where C_1 is a constant to be experimentally determined. Equations (2.15(d)), (2.19), and (2.21) then can be written as

$$S_k = f(X_k, E_k)$$

$$= \frac{U - X_k}{C_2}$$

$$= \frac{(C_1 + 2)\bar{M} - X_k}{C_2} \qquad \text{if } E_k\text{'s} = +1 \qquad (2.24)$$

and

$$S_k = f(X_k, E_k)$$

$$= -\frac{L + X_k}{C_2}$$

$$= -\frac{C_1\bar{M} + X_k}{C_2} \qquad \text{if } E_k\text{'s} = -1 \qquad (2.25)$$

The step size response equations can be derived as follows:

$$S_k = X_{k+1} - X_k$$

and from Eq.(2.18)

$$S_k = \frac{(U - X_0)(1 - \frac{1}{C_2})^k}{C_2} \qquad \text{if } E_k\text{'s} = +1$$

$$\cong \frac{(U - X_0)}{C_2} e^{-k/C_2}$$

$$= \frac{(C_1 + 2)\bar{M} - X_0}{C_2} e^{-k/C_2} \quad . \qquad (2.26)$$

When tracking a negative slope we use Eq.(2.20),

$$S_k = X_{k+1} - X_k$$

$$= - \frac{(L + X_0)(1 - \frac{1}{C_2})^k}{C_2} \qquad \text{if } E_k\text{'s} = -1$$

$$\simeq - \frac{(L + X_0)}{C_2} e^{-k/C}$$

$$= - \frac{C_1\bar{M} + X_0}{C_2} e^{-k/C_2} . \qquad (2.27)$$

Equation (2.26) and Eq. (2.27) are plotted in Fig. 2.31.

The asymptotic exponential type of response will not only reduce the slop overload noise but it will also reduce the distortion of the encoded picture. Figure 2.16 is reproduced in Fig. 2.32 with the B-mode ADM. It is clear that faster rising response will reduce picture distortions.

The values of SNR of the B-mode ADM with various values of C_1 and C_2 are shown in Table 2.11. Many different sets of C_1 and C_2 have been studied. However, the four different combinations shown in Table 2.11 give the highest SNR value. Among them the combination of $C_1 = 1$ and $C_2 = 11$ shows the best overall performance over a general class of video signals.

The encoded picture and difference picture are shown

in Fig. 2.33(a) and Fig. 2.33(b), respectively. Although not much of a difference can be seen from these pictures, the measurements of SNR indicates consistantly that the B-mode ADM has higher SNR than any other one dimensional adaptive DM schemes.

2.4 The Comparison of One Dimensional Delta Modulators

Signals are always accompanied by noise, and noise will perturb the DM decision on the error sign, E_k, therefore no two DM encoded video frames are exactly the same. This kind of frame-to-frame discrepancy is annoying to human observers. A good ADM algorithm should generate only a little amount of frame-to-frame noise. The ratio of the encoded signal power to the frame difference power is shown in Table 2.12 for all different one dimensional DM algorithms. The linear DM, as expected, has the least amount of frame difference, or the highest signal to frame difference ratio. It is because that in linear DM the step size is fixed to a small value, and the greatest difference at each pixel is only 2 S_0. A-mode ADM, however, has the highest SNR or the lowest frame discrepancy among the ADM category. This is obviously because of its smooth step size limitation characteristics. The B-mode ADM does not produce good results because of its abrupt jump whenever an edge is detected. The Song mode algorithm, however, consistantly

gives the highest amount of frame difference. Although the differences among algorithms are small, the measurement of frame difference does provide us with another angle to observe the merits of the different algorithms.

The merits of the A-mode and the B-mode ADM can be seen from another unusual case, namely the 1 b/p sampling rate. At the sampling rate of 1 b/p, the encoded picture of any one dimensional DM is so bad that 1 b/p is seldom seriously considered in any applications. However, both the A-mode and B-mode ADM showed some improvement in the SNR and the encoded picture quality that shorter edge busyness and less amount of ringing are seen. **Figure 2.34** (a) and (b) are the encoded pictures of the Song mode ADM and the B-mode ADM at 1 b/p, respectively. The values of SNR are shown in Table 2.13. A 3 dB improvement of the B-mode ADM algorithm is significant and it deserves more investigations.

2.5 Other One Dimensional Adaptive Delta Modulators

The Song mode ADM algorithm encoded picture has little granular noise but a large amount of slope overload and overshoot noise as a result of its positive power exponentially rising and falling characteristics. The modified exponential ADM algorithm, the A-mode algorithm, and the negative power exponential asymptotic ADM algorithm,

the B-mode algorithm,will reduce the amount of overshoot and slope overload noise, but will increase the granular noise as a result of larger initial step size. It is desirable that two different

combined into one DM system. One mode is used to encode the flat area of a picture to yield low granular noise. One mode is used to encode image edges to quickly respond to the input signal in order to reduce slope overload and overshoot noise. A scheme that combines the Song mode algorithm and the B-mode algorithm was studied. The Song mode algorithm is used for the encoding of low activity areas, whereas, the B-mode algorithm is switched into operation whenever a big jump is required. One additional output bit is transmitted to inform the decoder the mode of the encoded signal. Because two output bits are needed for each sample, the channel bandwidth is twice as wide as compared to a normal 1-bit DM.

A threshold value, T, is first determined. The ADM operates with the Song mode if the difference between the input sample and the previous DM estimate is less than T. It will switch to the B-mode if the difference exceeds T. In order to ensure a smooth transition between these two modes, a detailed set of equations are used.

	CONDITIONS			
	E_k	E_{k-1}	previous mode	$M_k - X_k$
$S_k = \lvert S_{k-1} \rvert [E_k + \frac{1}{2}E_{k-1}]$	ϕ	ϕ	Song	<T
$S_k = S_{max}\ E_k$	ϕ	ϕ	B-mode	<T
$S_k = \frac{(C_1+2)\overline{M} - X_k}{C_2}$	1	ϕ	Song	<T
$S_k = -\frac{C_1\overline{M} + X_k}{C_2}$	0	ϕ	Song	<T
$S_k = \frac{(C_1 + 2)\overline{M} - X_k}{C_2}$	1	1	B-mode	>T
$S_k = -\frac{C_1\overline{M} + X_k}{C_2}$	0	0	B-mode	>T
$S_k = -\frac{1}{2}\lvert S_{k-1} \rvert$	0	1	B-mode	>T
$S_k = \frac{1}{2}\lvert S_{k-1} \rvert$	1	0	B-mode	>T

where S_{max} is the maximum step size of the Song mode algorithm which is considerably less than the maximum possible step size of the DM system.

The purpose of this study is to compare the performance of two DM systems. One system is sampled at a faster rate, but the information about the input signal carried through the channel to the receiver is less. The other system is sampled at half the rate of

the former system, but the information sent to the receiver at each sample is more accurate. Both systems have the same channel rate. The results are shown in Table 2.14.

It can be seen that the combination scheme does not work very well. It is mainly due to the lower sampling rate. The more accurate knowledge of the input sample at the receiver, by receiving two information bits for each sample, cannot compensate for the loss of the ability to track the less correlated input samples. The principal requirement of the operation of a DM is the close correlation of the input samples. Reducing the sampling rate will seriously impair the tracking ability of a delta modulator. Although this combination scheme can be entropy encoded to reduce the final channel rate, we did not pursue any further because of the intrinsic limitation of the low rate one dimensional DM. A better performance can easily be achieved by using a two dimensional delta modulator which will be discussed in the next chapter.

	Linear	Abate	Song	A-mode	B-mode
Girl	24.6	27.1	28.0	29.6	30.0
Boy	23.7	23.2	24.8	25.5	26.5
Test Pattern	20.1	19.2	19.4	21.8	22.6

Table 2.1 SNR (dB) of one dimensional DM.

Input signal power (dB)	+6	0	-6	-12
Boy	15.7	23.7	20.4	14.9

Table 2.2 SNR (dB) of Linear DM, step size is 6 S_o.

Step size	$2S_o$	$4S_o$	$5S_o$	$6S_o$	$7S_o$	$8S_o$	$10S_o$	$12S_o$
Boy	16.7	21.2	22.6	23.7	23.6	23.0	21.6	20.2

Table 2.3 SNR (dB) of Linear DM at different step sizes.

	Input signal power (dB)			
	+6	0	-6	-12
Abate	19.1	23.2	23.4	23.1
Song	20.7	24.8	24.0	23.3
A-mode	21.1	25.5	26.2	24.5
B-mode	/	26.5	26.3	25.8

Table 2.4 SNR (dB) of one dimensional DM with different input signal power.

	Maximun step size, S_{max}		
	$8S_o$	$16S_o$	$32S_o$
Girl	28.1	28.0	27.2
Boy	24.5	24.8	24.0
Test Pattern	18.2	19.4	18.3

Table 2.5 SNR (dB) of Song mode ADM with different maximun step sizes, S_{max} .

Step size	Girl		Boy	
	No. of occurrence	Probability	No. of occurrence	Probability
1	268565	0.5138	217502	0.4161
2	153155	0.2930	136394	0.2609
3	45122	0.0863	63752	0.1220
4	23848	0.0456	38620	0.0739
6	15602	0.0298	25578	0.0490
8	1800	0.0035	5402	0.0103
9	7530	0.0144	15072	0.0288
12	859	0.0016	3016	0.0058
13	3369	0.0064	7910	0.0151
16	2903	0.0056	9510	0.0182

Table 2.6 Step size statistics of Song mode ADM with S_{max}=16 S_o. Total number of pixels is 522,753.

Step size	Girl		Boy	
	No. of occurrence	Probability	No. of occurrence	Probability
1	268904	0.5144	217030	0.4152
2	152867	0.2924	136160	0.2605
3	44796	0.0569	61110	0.1169
4	23057	0.0441	36926	0.0706
5	531	0.0010	1108	0.0021
6	14750	0.0282	26306	0.0500
7	823	0.0016	2068	0.00396
8	17	0.00003	170	0.00033
9	8288	0.0159	17716	0.0340
10	901	0.00172	1884	0.0036
11	16	0.00003	144	0.00028
12	12	0.00002	354	0.0068
13	3745	0.00716	9550	0.0183
14	678	0.0013	1648	0.0032
15	393	0.0075	1008	0.0019
16	21	0.00004	518	0.0010
18	2	0	126	0.00024
19	1693	0.00324	4742	0.0090
21	521	0.001	1046	0.002
22	21	0.00004	212	0.004
24	4	0	348	0.00067
27	0	0	6	0.00001
28	679	0.0013	1896	0.0036
31	22	0.00004	222	0.0004
32	12	0	470	0.0009

Table 2.7 Step size statistics of Song mode ADM with S_{max}=32. Total number of pixel is 522,753.

	Pixels shifted when taking SNR measurement					
	n=-1	n=0	n=+1	n=+2	2-bit ave. n=0	2-bit ave. n=+1
Girl	24.8	27.5	28.0	25.3	29.2	
Boy	20.9	23.2	24.8	23.6	24.6	
Test Pattern	14.3	17.1	19.4	18.6	18.7	20.2

Table 2.8 SNR (dB) of Song mode ADM with phase shift but without filtering.

	non-Look-Ahead	Look-Ahead
Quantization error	24.8	25.5
Frame difference	25.5	27.4

Table 2.9 SNR (dB) of Song mode ADM encoded "Boy" at $S_{max}=16\ S_o$.

	S_{max}=16 C=4	S_{max}=32 C=4	S_{max}=32 C=8	S_{max}=16 C=8
Girl	29.6	28.0	29.3	30.3
Boy	25.5	25.2	25.5	25.3
Test Pattern	21.8	21.5	21.7	20.1

Table 2.10 SNR (dB) of A-mode ADM.

	C_1=1 C_2=11	C_1=1 C_2=13	C_1=1 C_2=16	C_1=2 C_2=16
Girl	30.0	30.2	29.5	30.0
Boy	26.5	26.4	25.6	26.5
Test Pattern	22.6	21.7	19.9	22.8

Table 2.11 SNR (dB) of B-mode ADM, $C_3=5S_o$.

	Abate	Song $S_{max}=16$	A-mode $C=4$ $S_{max}=16$	B-mode $C_1=1$ $C_2=11$ $C_3=5S_o$	Linear
Girl	27.7	26.7	28.9	27.3	29.7
Boy	26.4	25.5	27.9	26.5	30.7
Test Pattern	22.3	22.1	24.8	24.5	28.6

Table 2.12 Signal-to-frame difference ratio (dB) of one dimensional DM.

	Abate	Song	A-mode	B-mode
Girl	21.6	21.6	22.1	24.8
Boy	19.2	18.9	19.2	20.9
Test Pattern	15.6	15.0	15.4	15.6

Table 2.13 SNR (dB) of one dimensional DM sampled at 1 bit/pixel.

Boy	Song	B-mode	Combination of Song mode & B-mode, 2 bits/pixel, $S_m=6S_o$, $C_1=2$, $C_2=16$, $T=8S_o$
SNR (dB)	24.8	26.5	23.6

Table 2.14 Comparison of Song mode, B-mode and combinational mode.

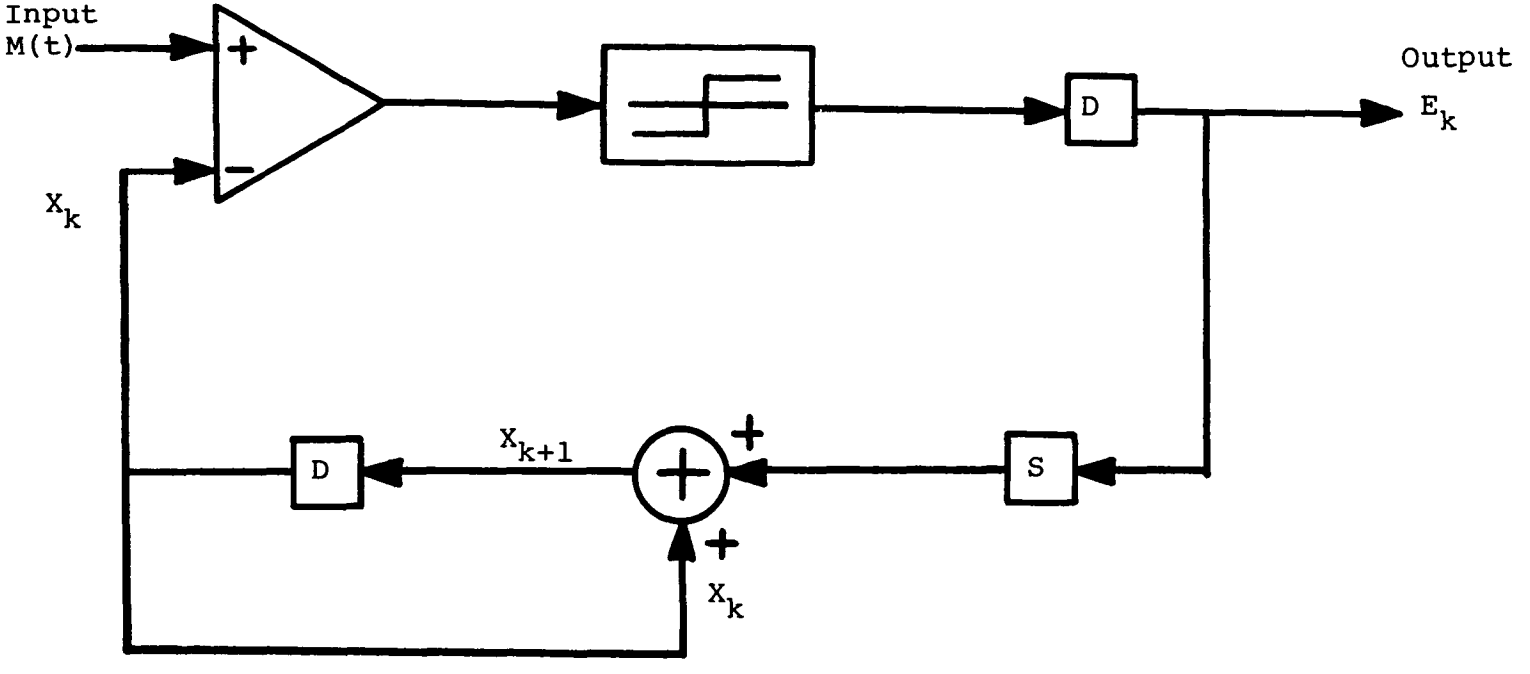


Fig. 2.1 Block diagram of digital Linear DM.

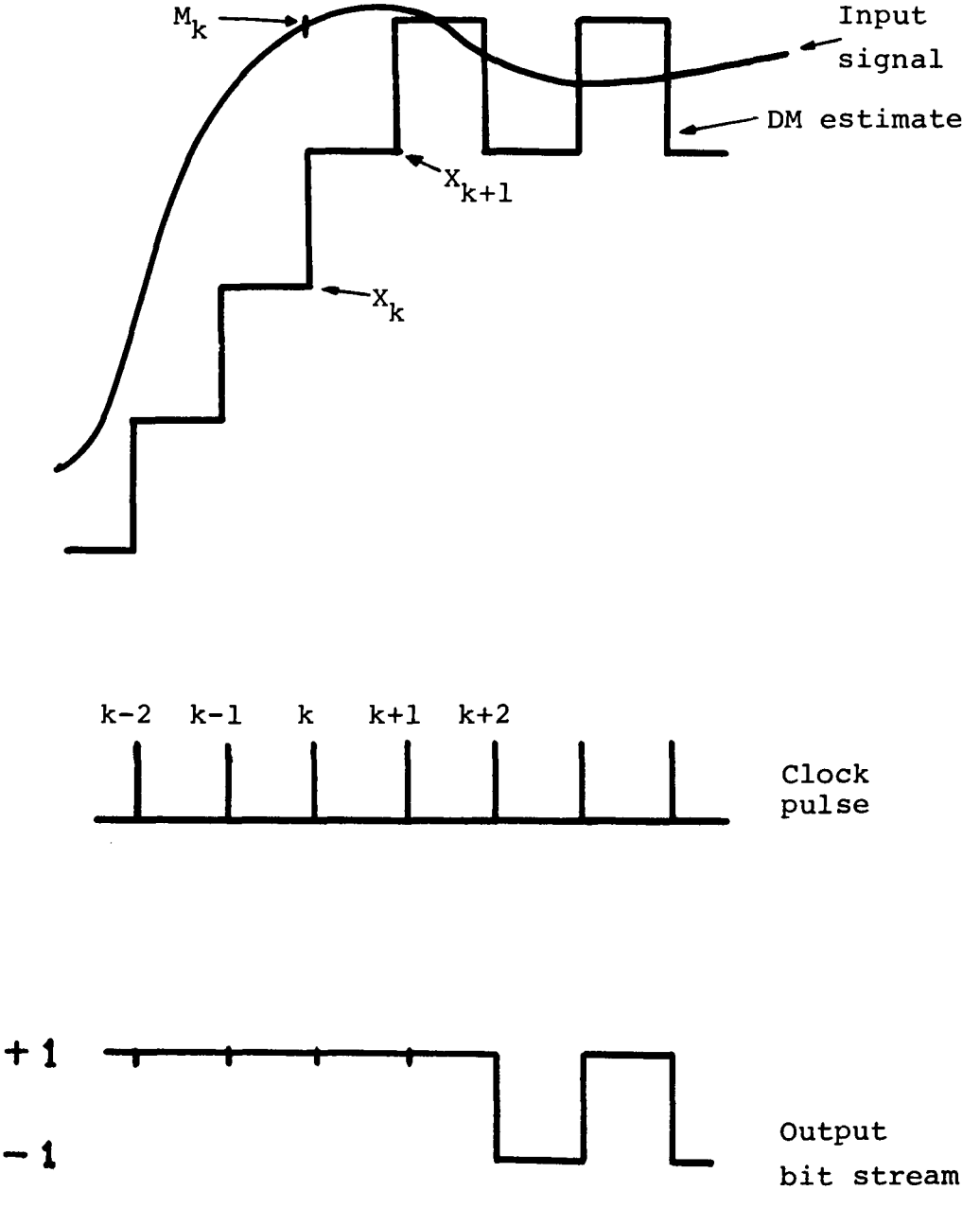


Fig. 2.2 **Timing** diagram of Linear DM.

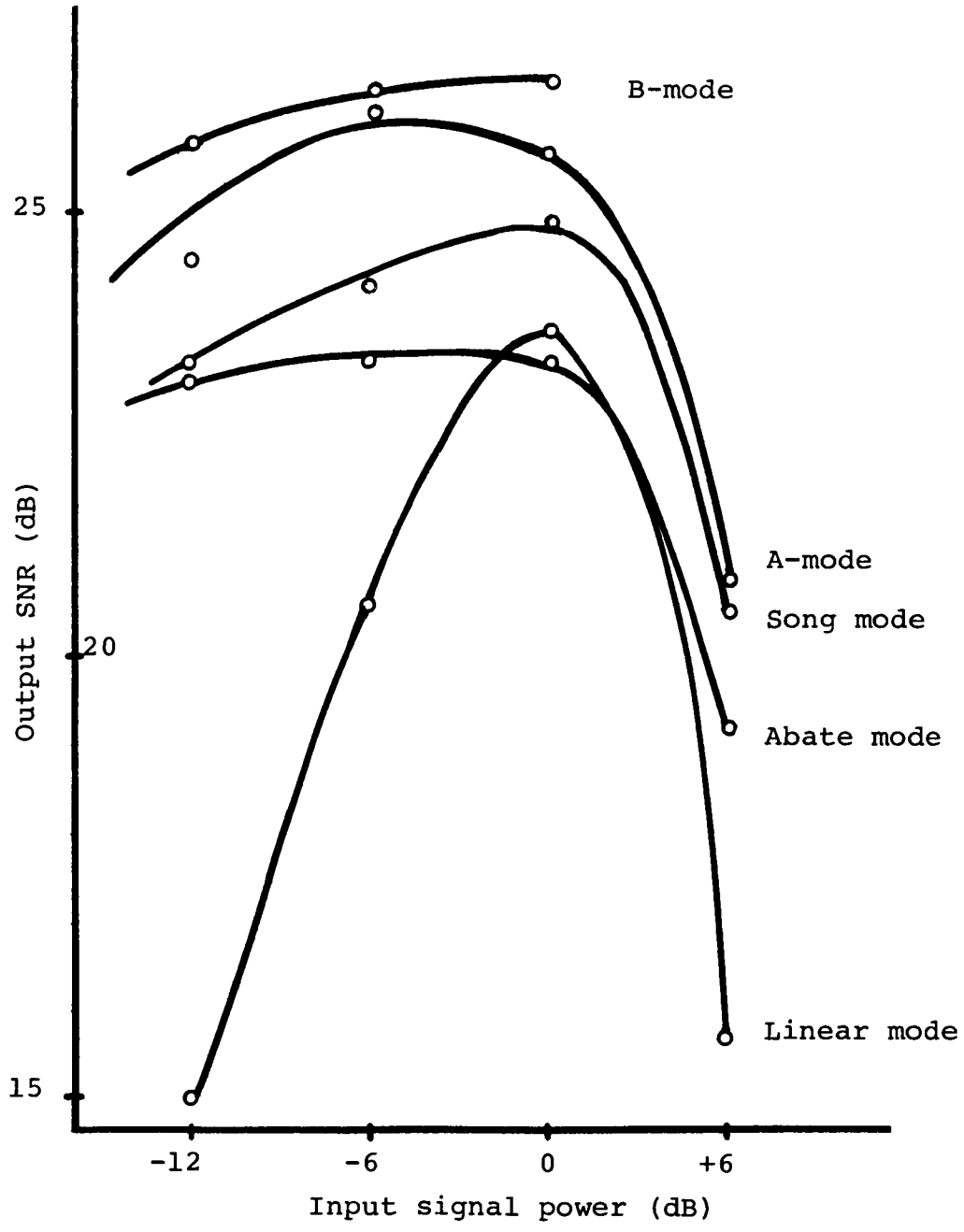


Fig. 2.3 Dynamic range of one dimensional DM.

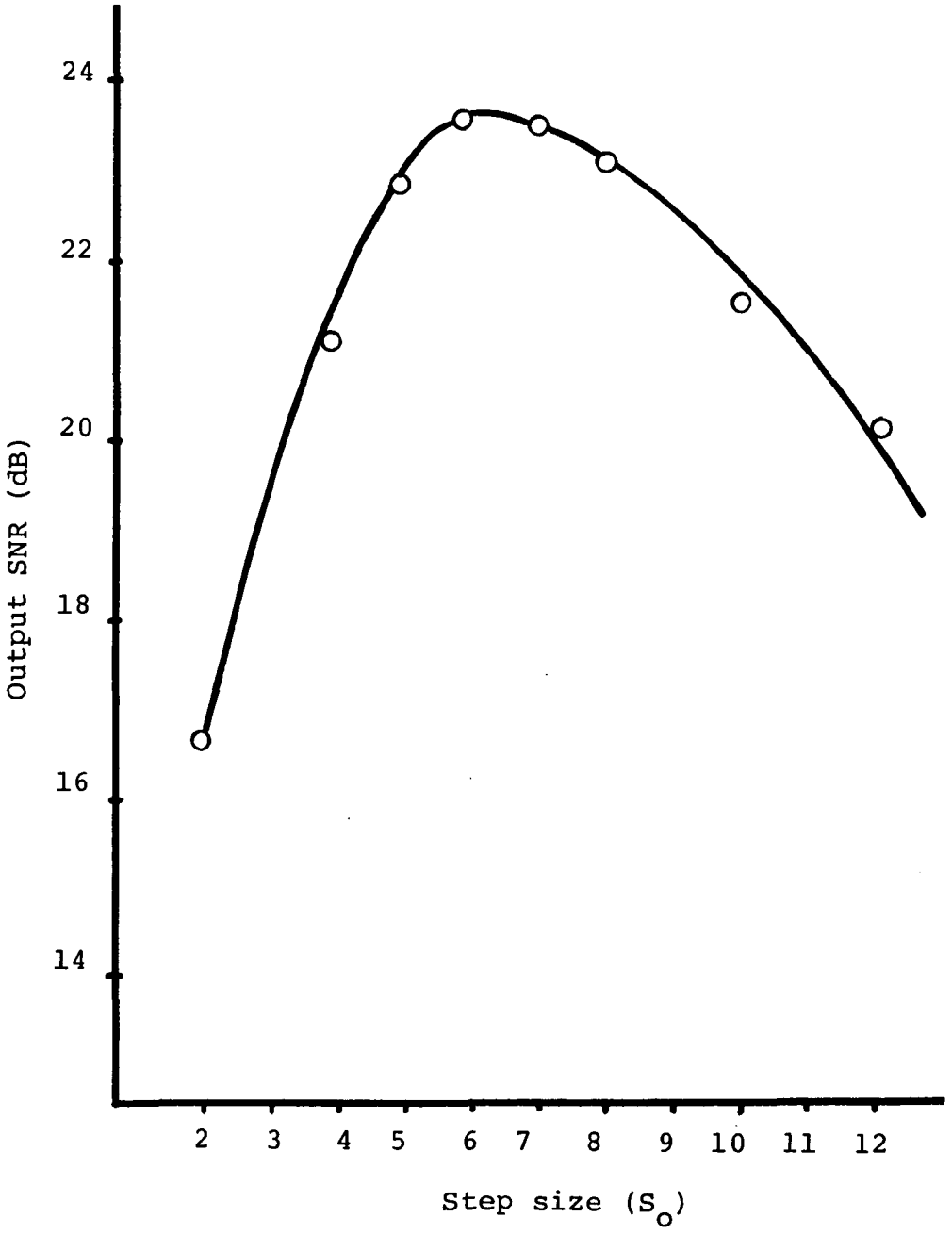


Fig. 2.4 Output SNR as a function of step size, "Boy", encoded by Linear DM.

(a)

(b)

Fig. 2.5 Linear DM encoded pictures, (a) $S = 5S_0$, (b) $S = 6S_0$.

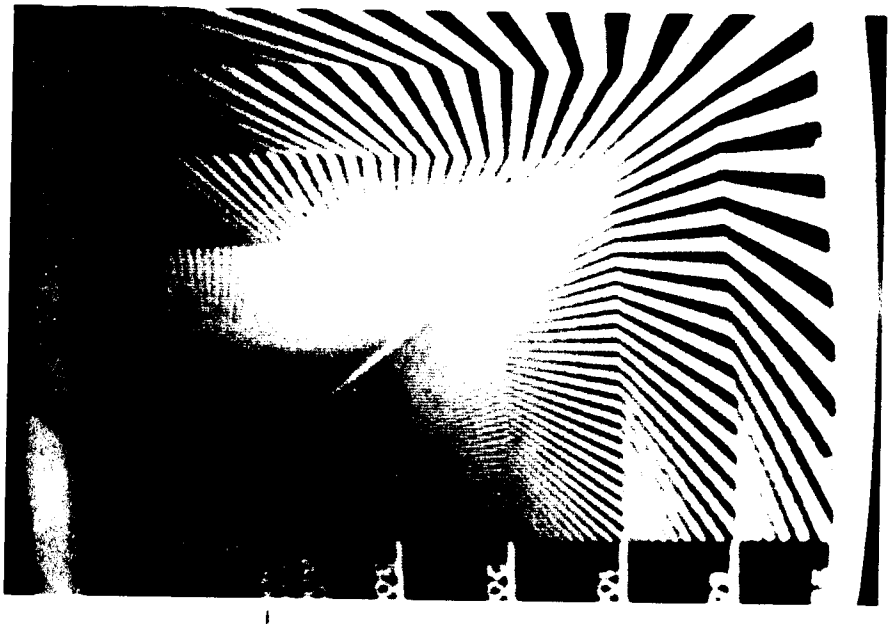


Fig. 2.6 Linear DM encoded picture, $S = 9S_0$.

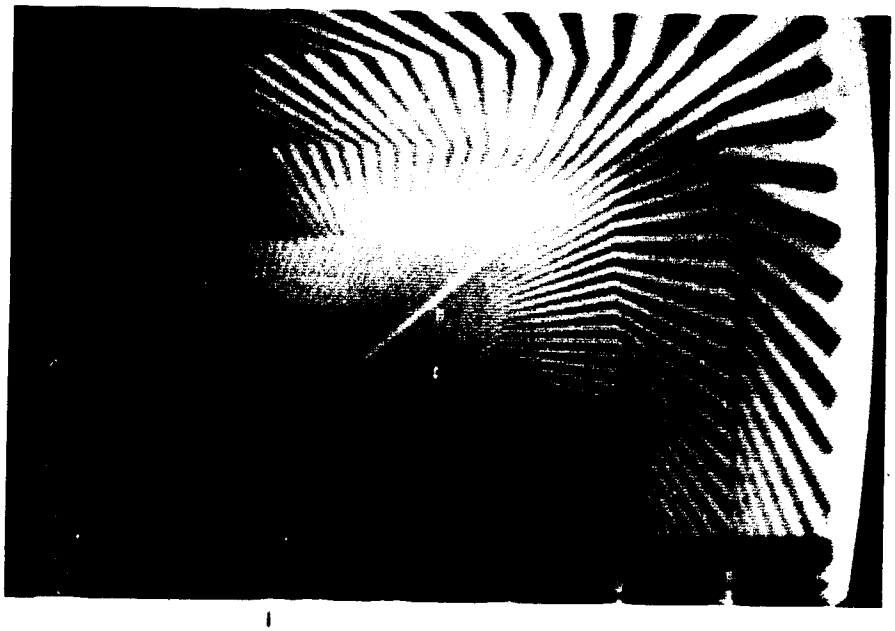


Fig. 2.7 Linear DM encoded picture, $S = 3S_0$.

Fig. 2.8 Abate mode ADM encoded picture, $S_{max} = 16S_0$.

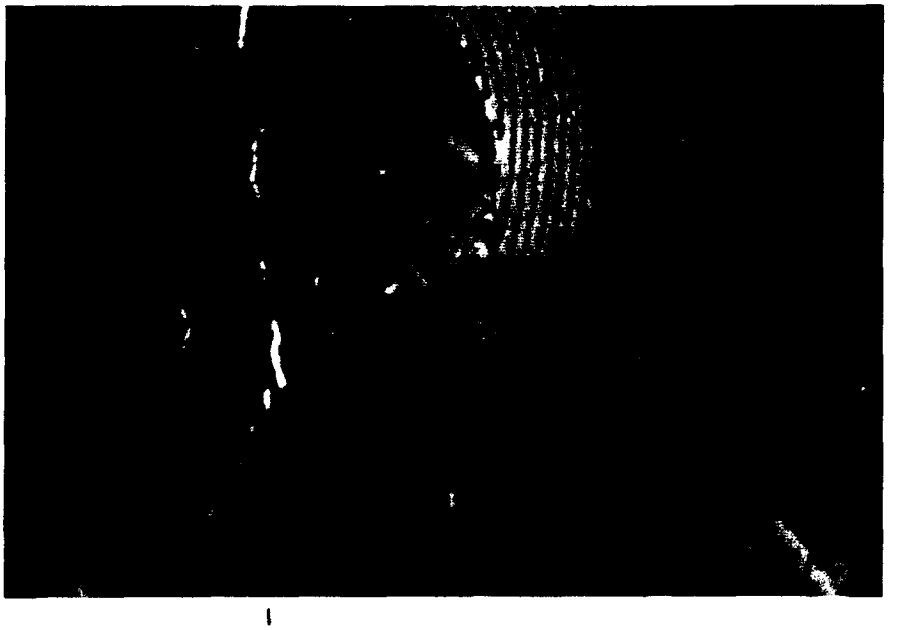


Fig. 2.9 The Difference picture of the Abate mode ADM.

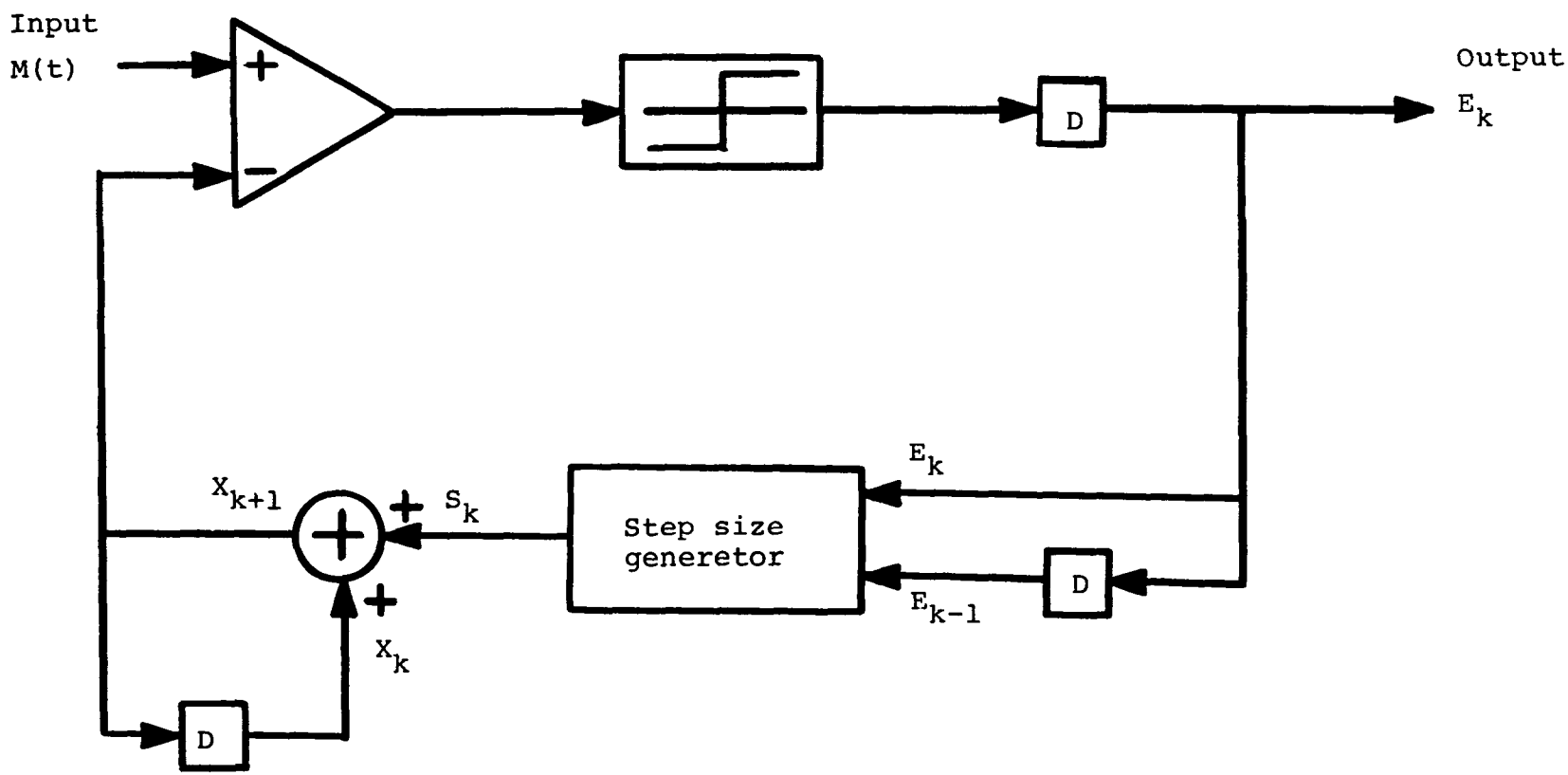


Fig. 2.10 Block diagram of Song mode ADM.

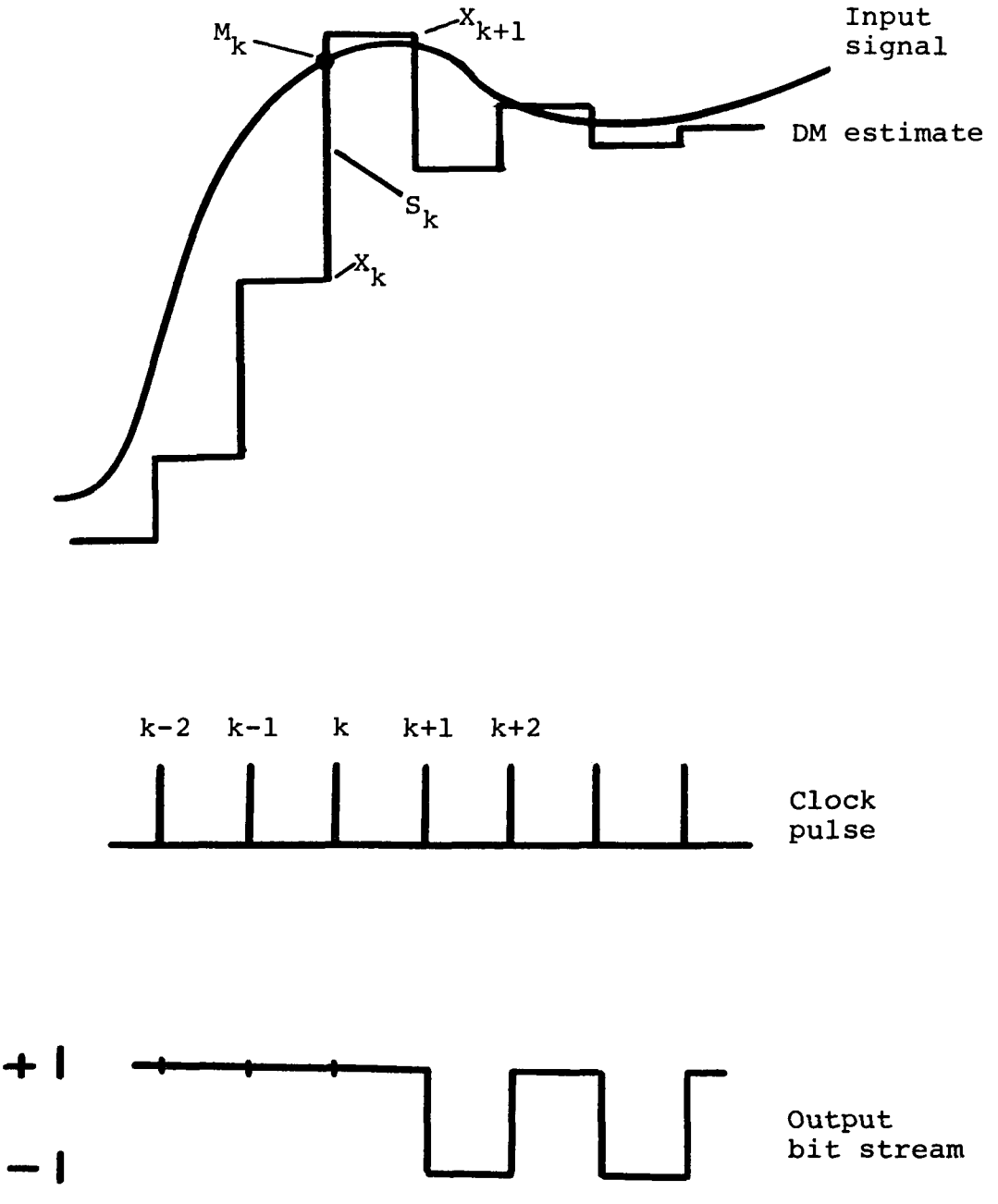


Fig. 2.11 Timing diagram of Song mode ADM.

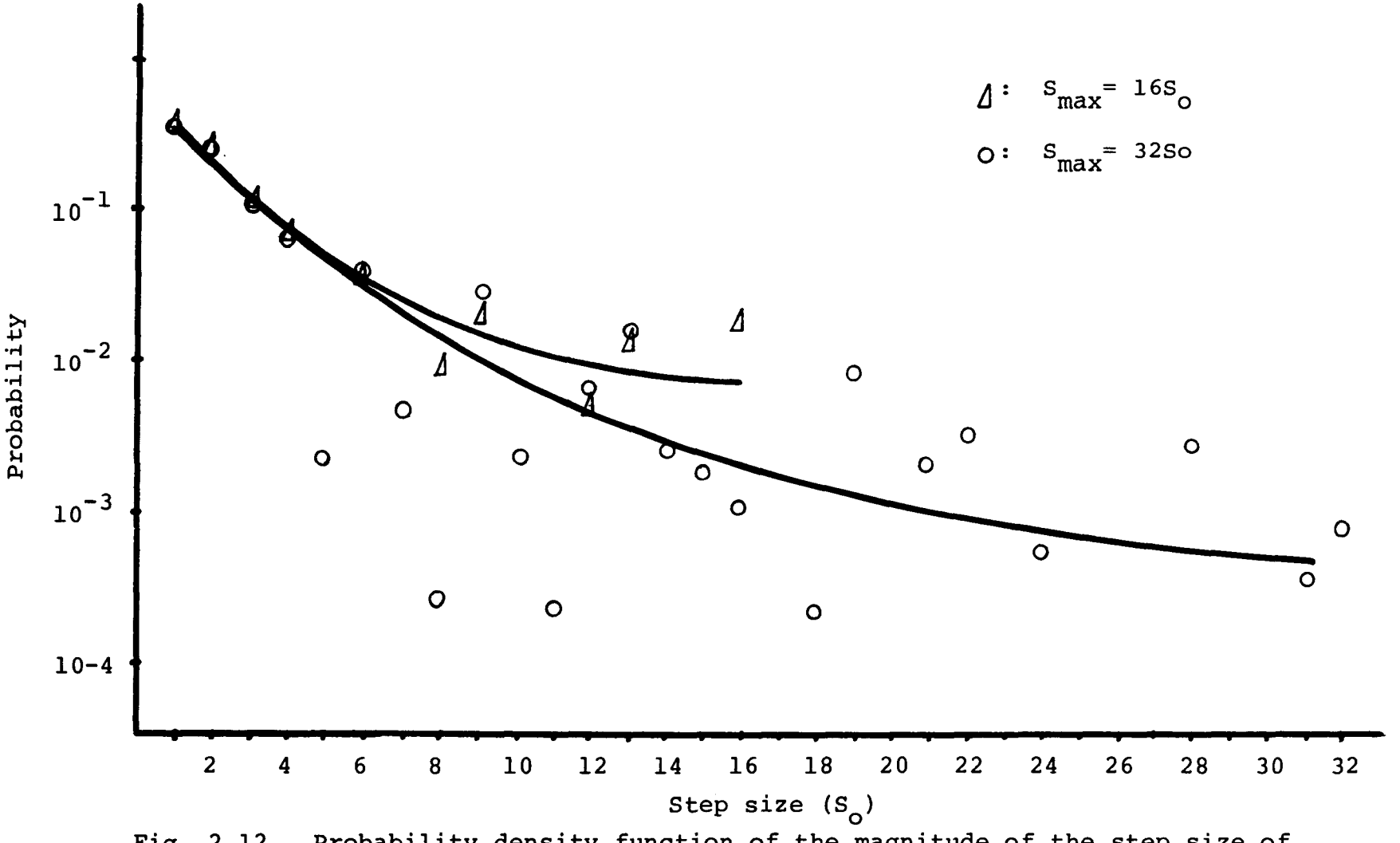


Fig. 2.12 Probability density function of the magnitude of the step size of the Song mode ADM.

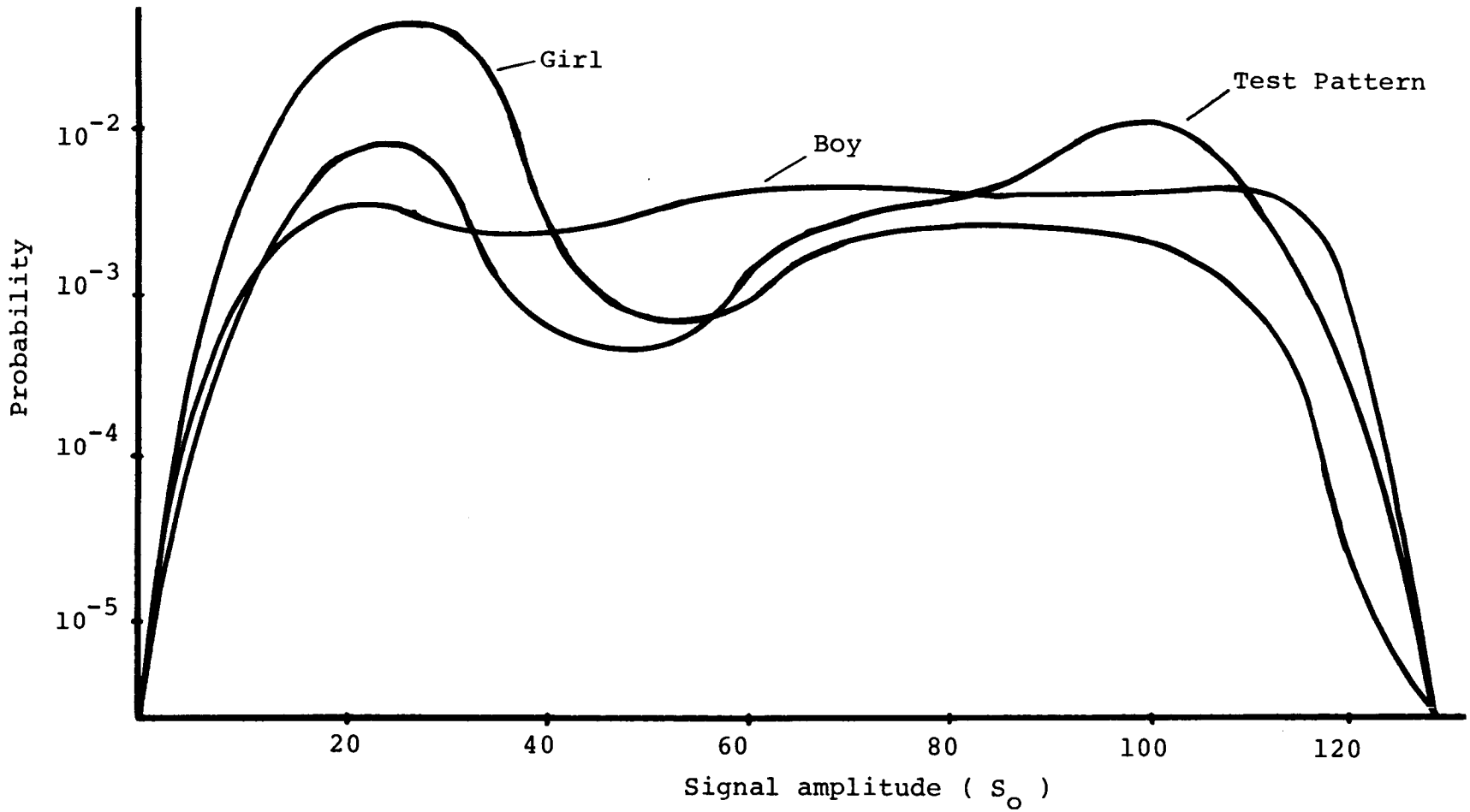


Fig.2.13 Probability density function of video signal amplitude.

(a)

(b)

Fig. 2.14 Song mode ADM encoded pictures.

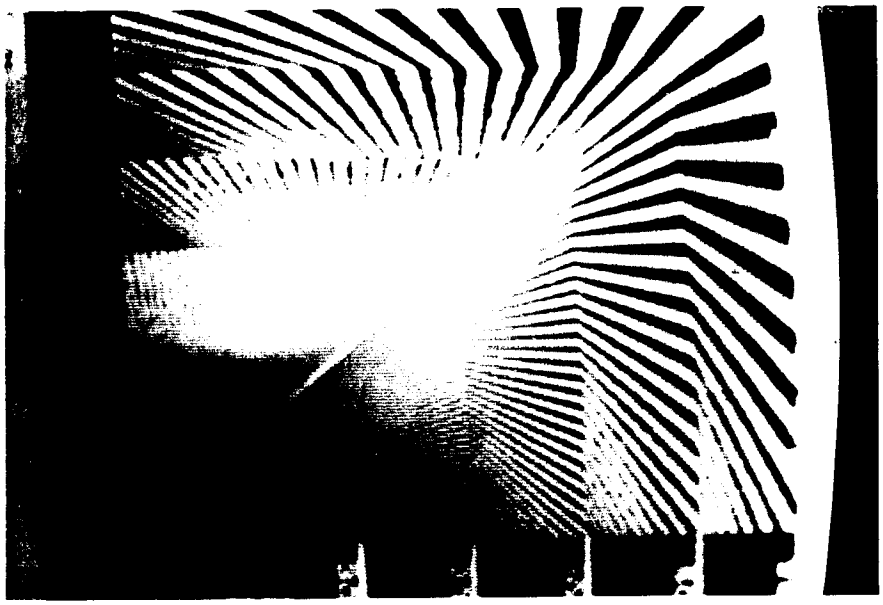


Fig. 2.15 Song mode ADM encoded "Test Pattern".

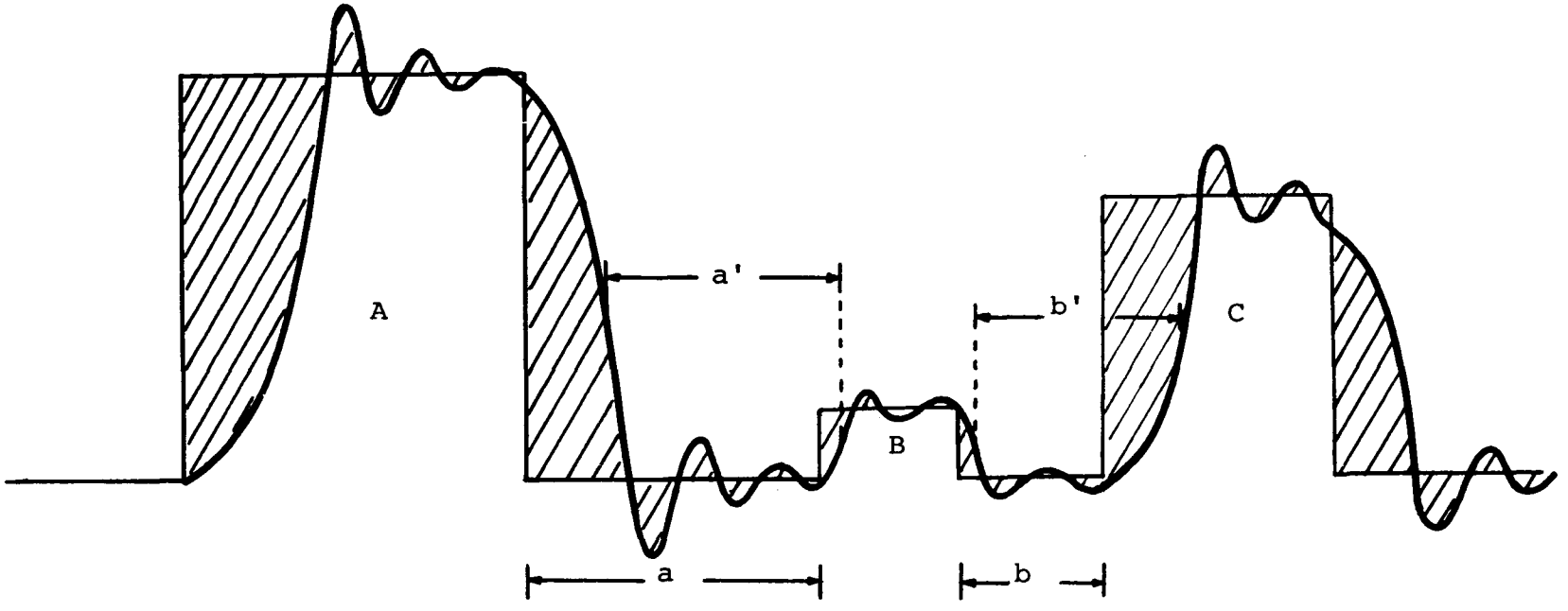


Fig. 2.16 Shaded area is the quantization noise. The encoded waveform is distorted not only in shape but also in relative positions.

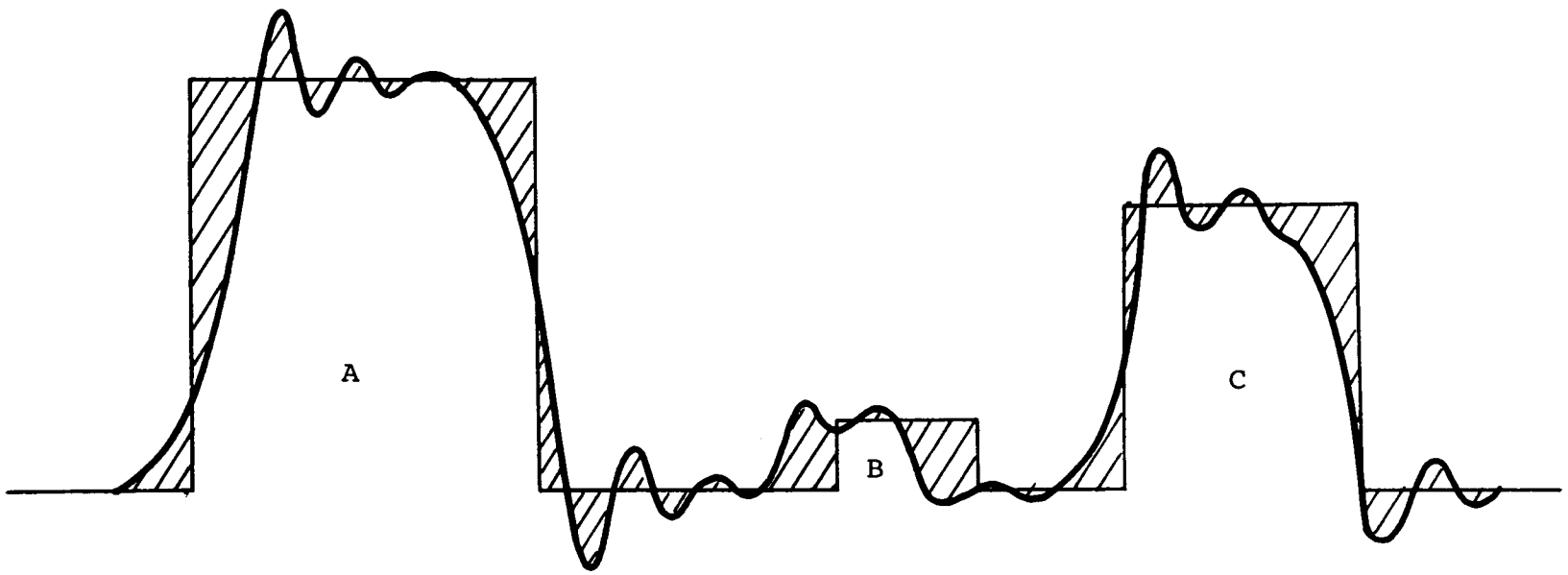


Fig. 2.17 Encoded waveform is shifted to the left. The quantization noise (shaded area) is reduced.

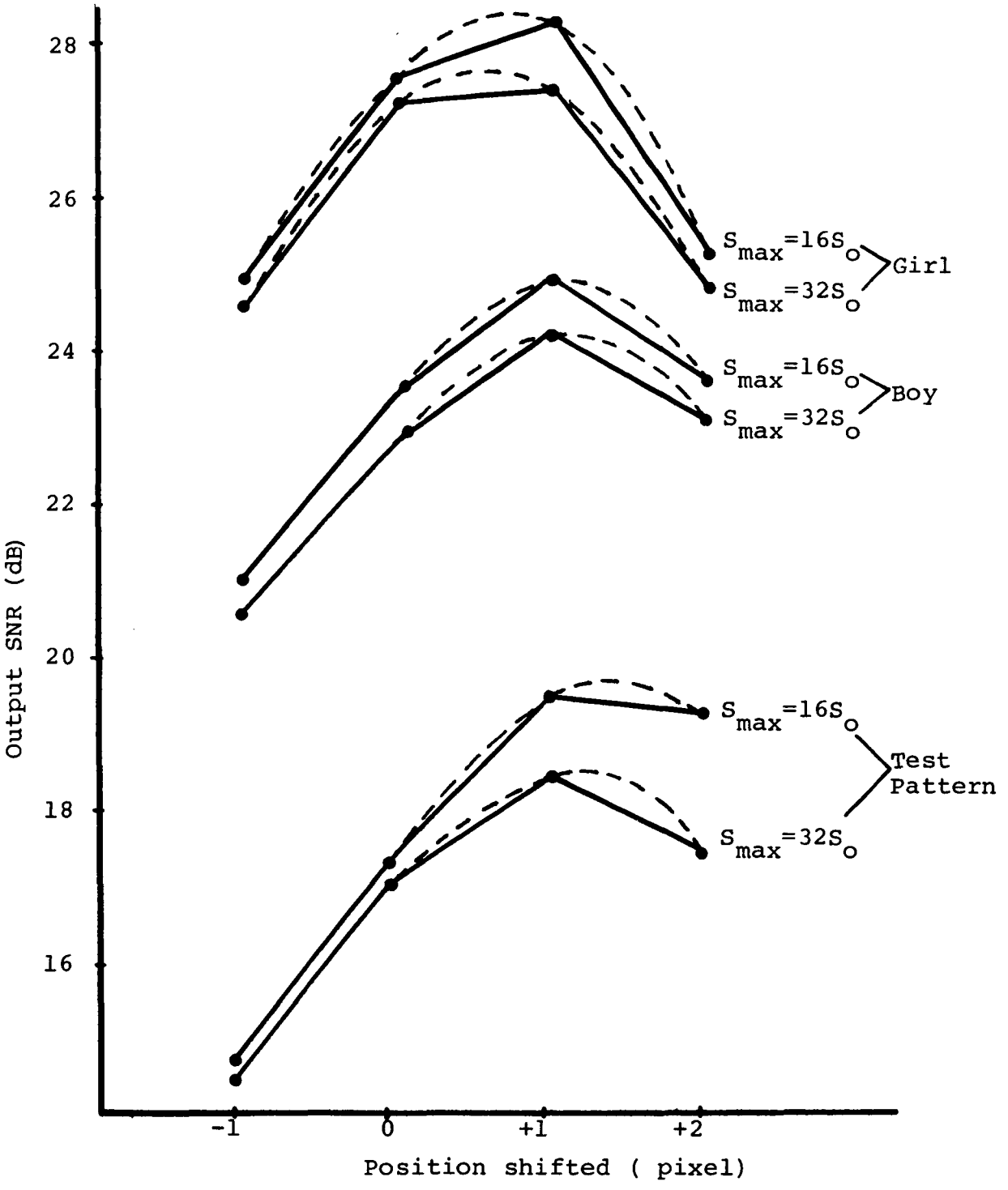
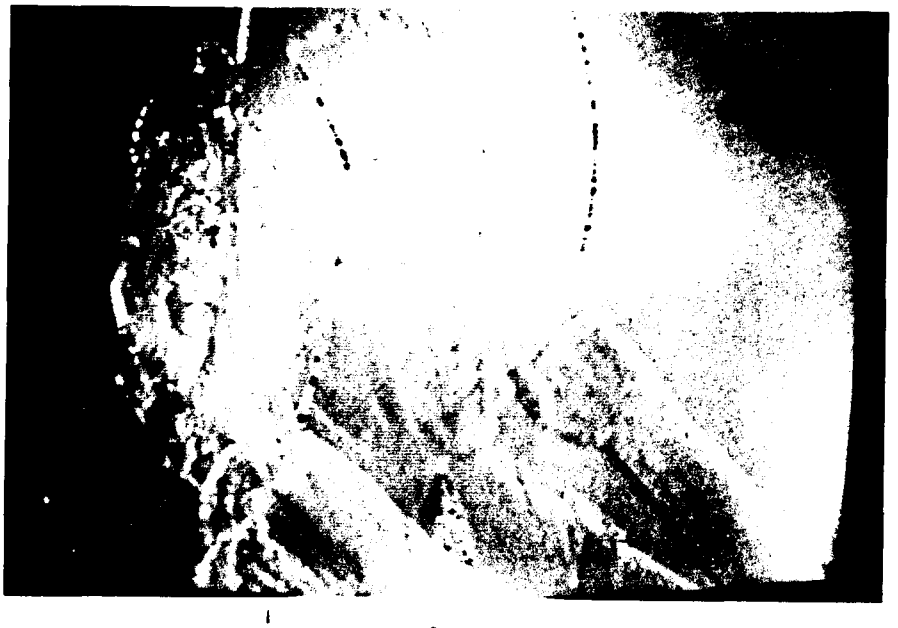


Fig. 2.18 SNR as a function of relative position shift between original picture and encoded picture.

(a)



(b)

Fig. 2.19 The difference pictures of the Song mode ADM,
(a) the encoded picture is not shifted,
(b) the encoded picture is shifted to the left by one pixel distance

(a)

(b)

Fig. 2.20 The difference pictures of the Song mode ADM,
(a) the encoded picture is not shifted,
(b) the encoded picture is shifted to the left by one pixel distance.

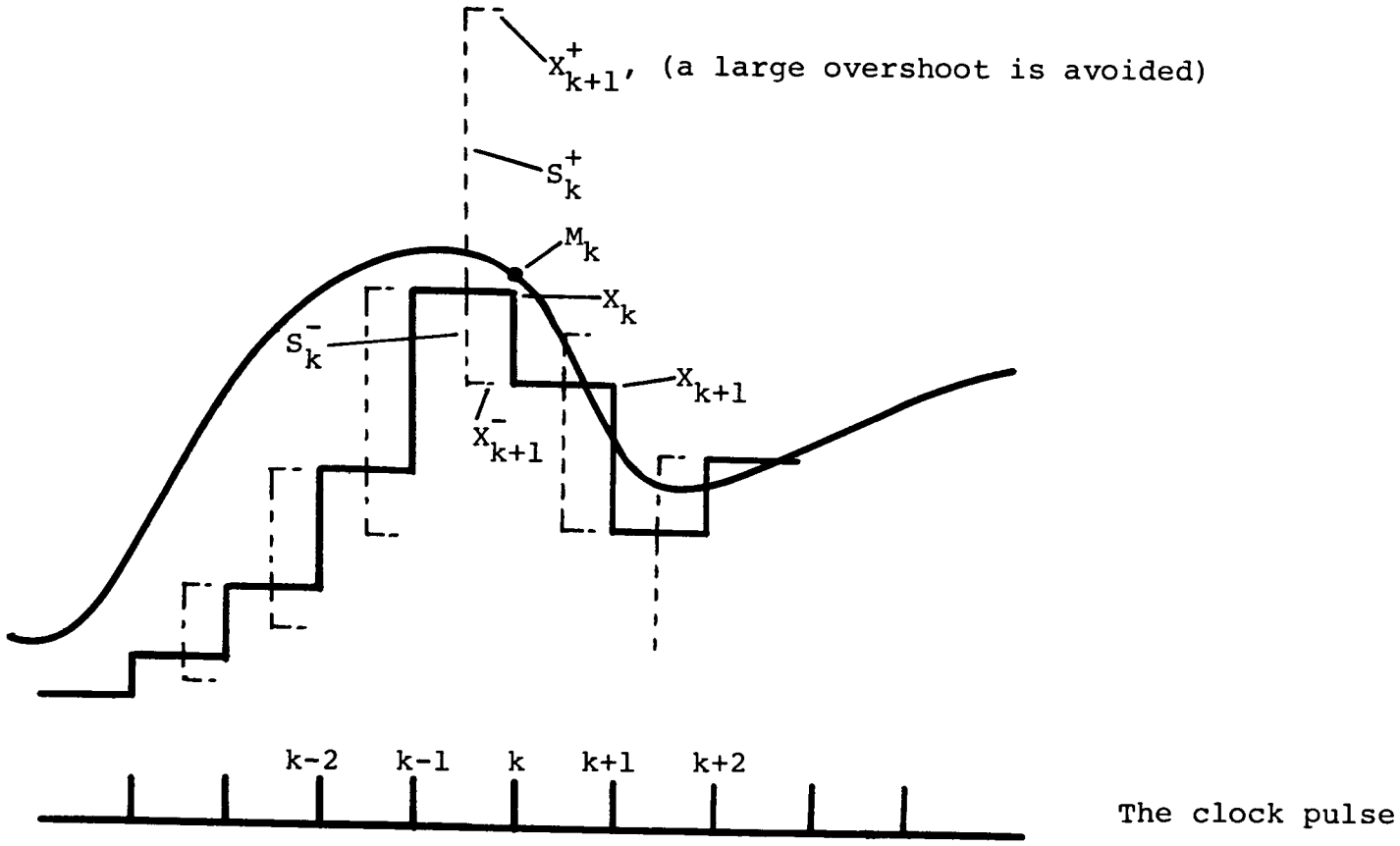


Fig. 2.21 Look-Ahead ADM

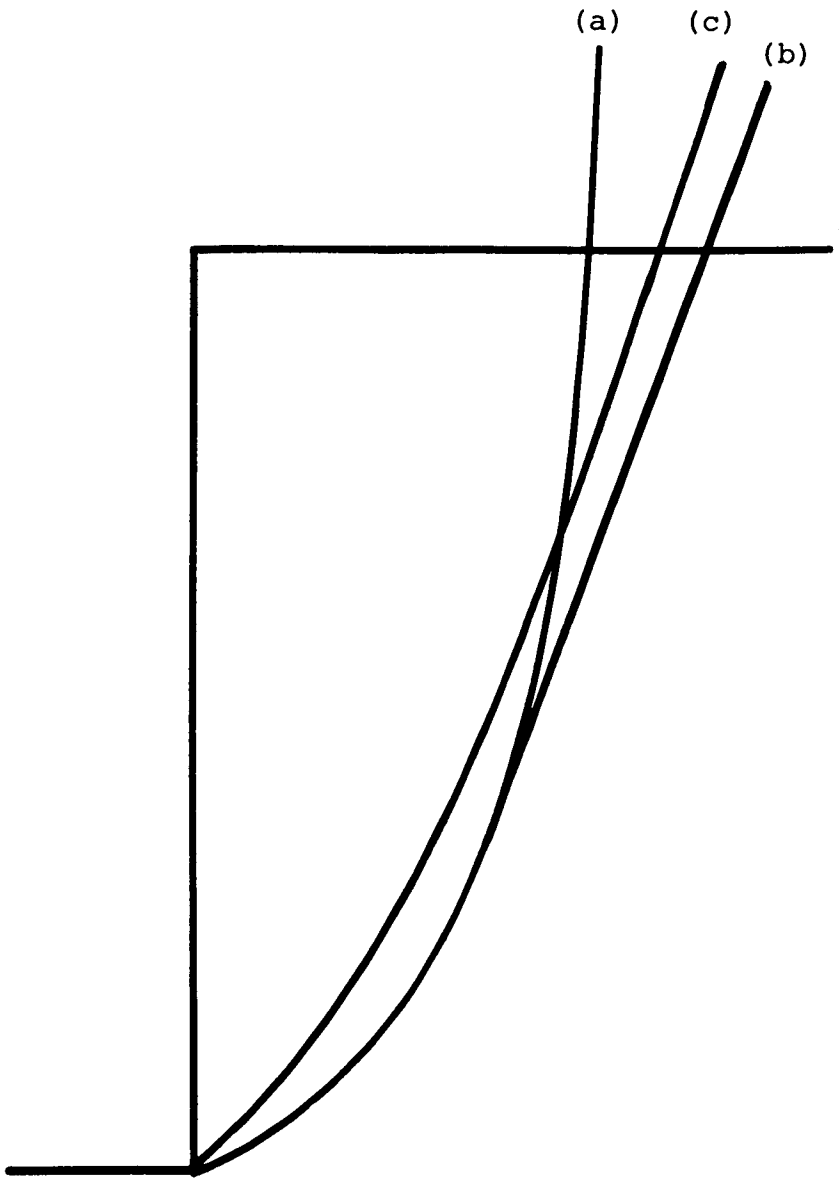
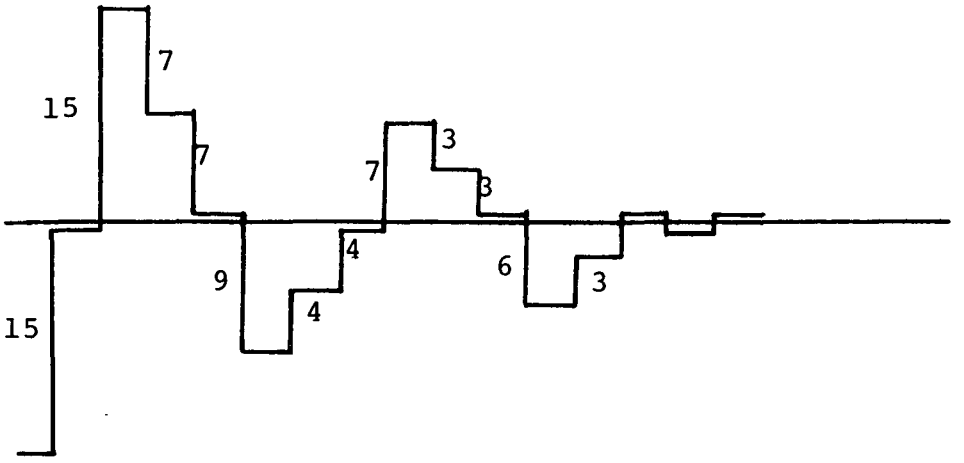
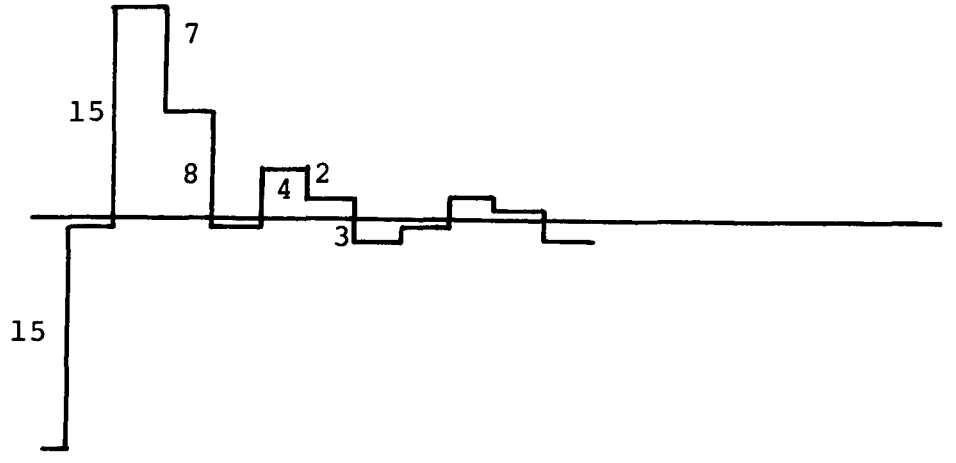


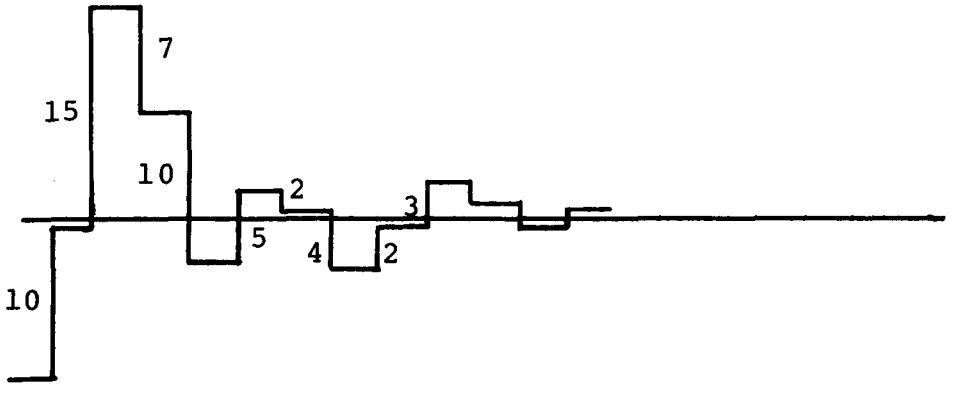
Fig. 2.22 The DM response to a step input,
(a) Song mode, (b) Song mode with a S_{max},
(c) A-mode.



(a) the A-mode ADM if a S_o is not included in Eq. (2.8(c)),



(b) the A-mode ADM, $S_{max} = 16S_o$, C=4,



(c) the Song mode ADM.

Fig. 2.23 The overshoot and oscillation patterns.

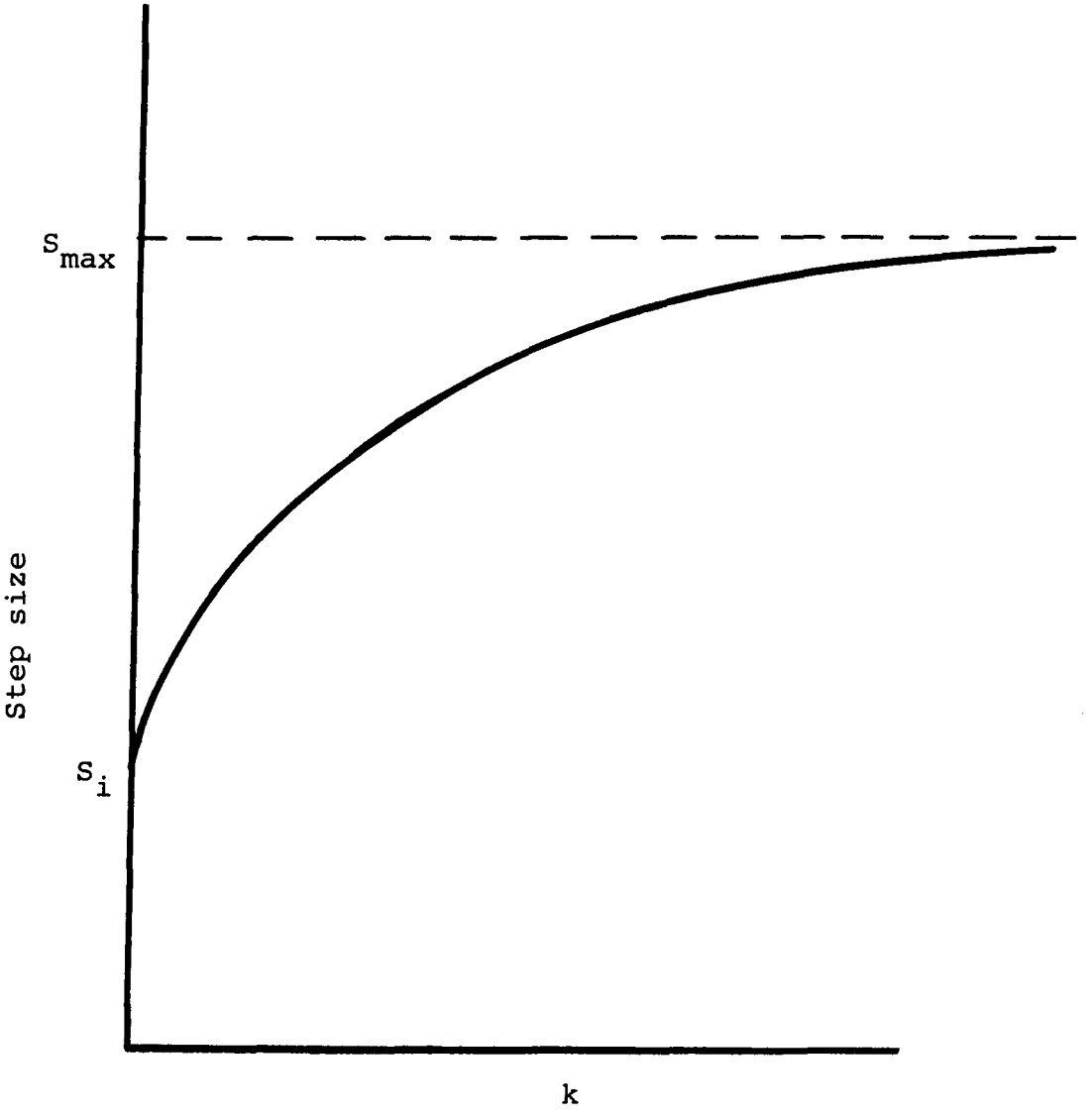


Fig. 2.24 The step size function of the A-mode ADM is a biased exponential function.

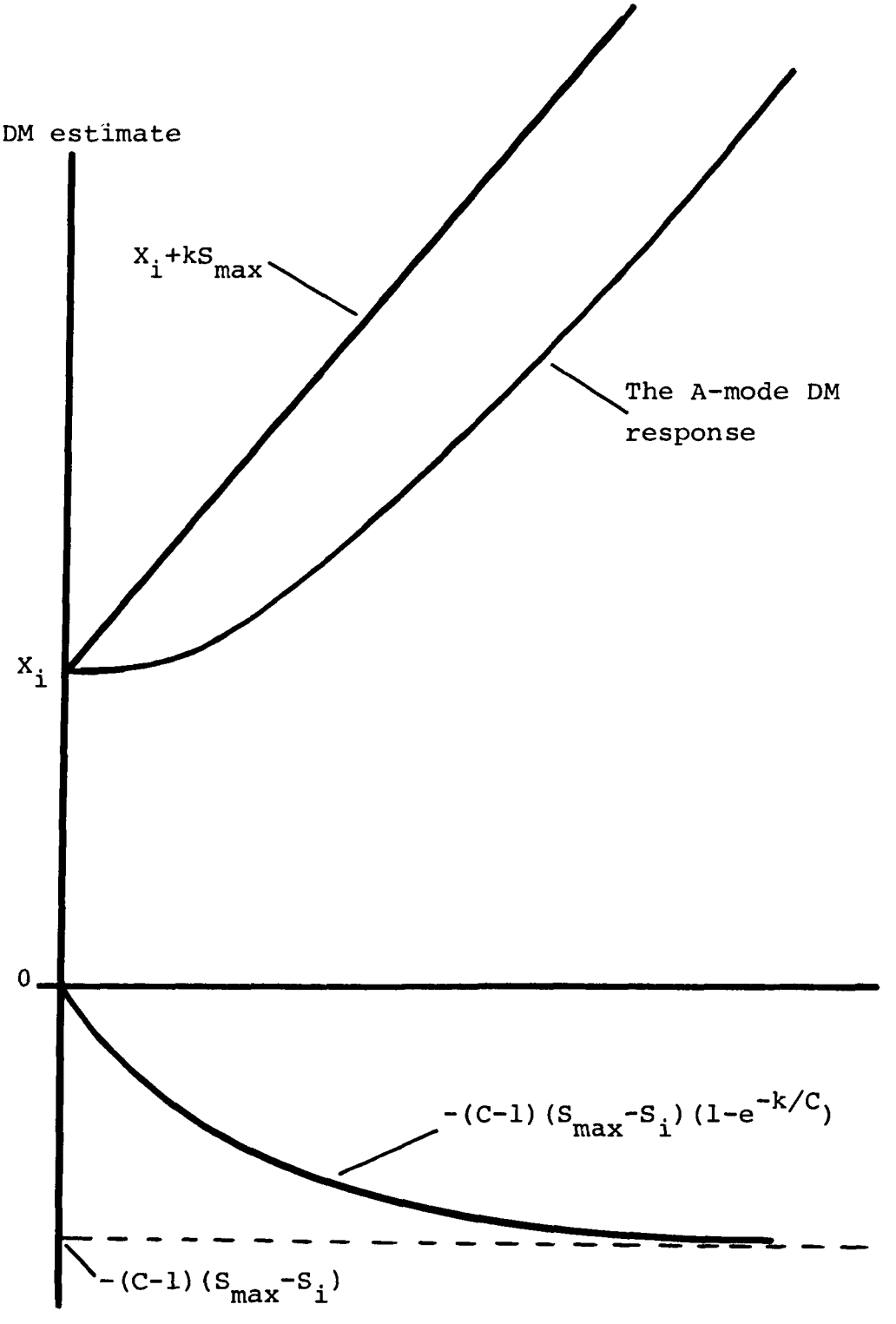


Fig. 2.25 The A-mode DM response to a step input is a sum of two equations.

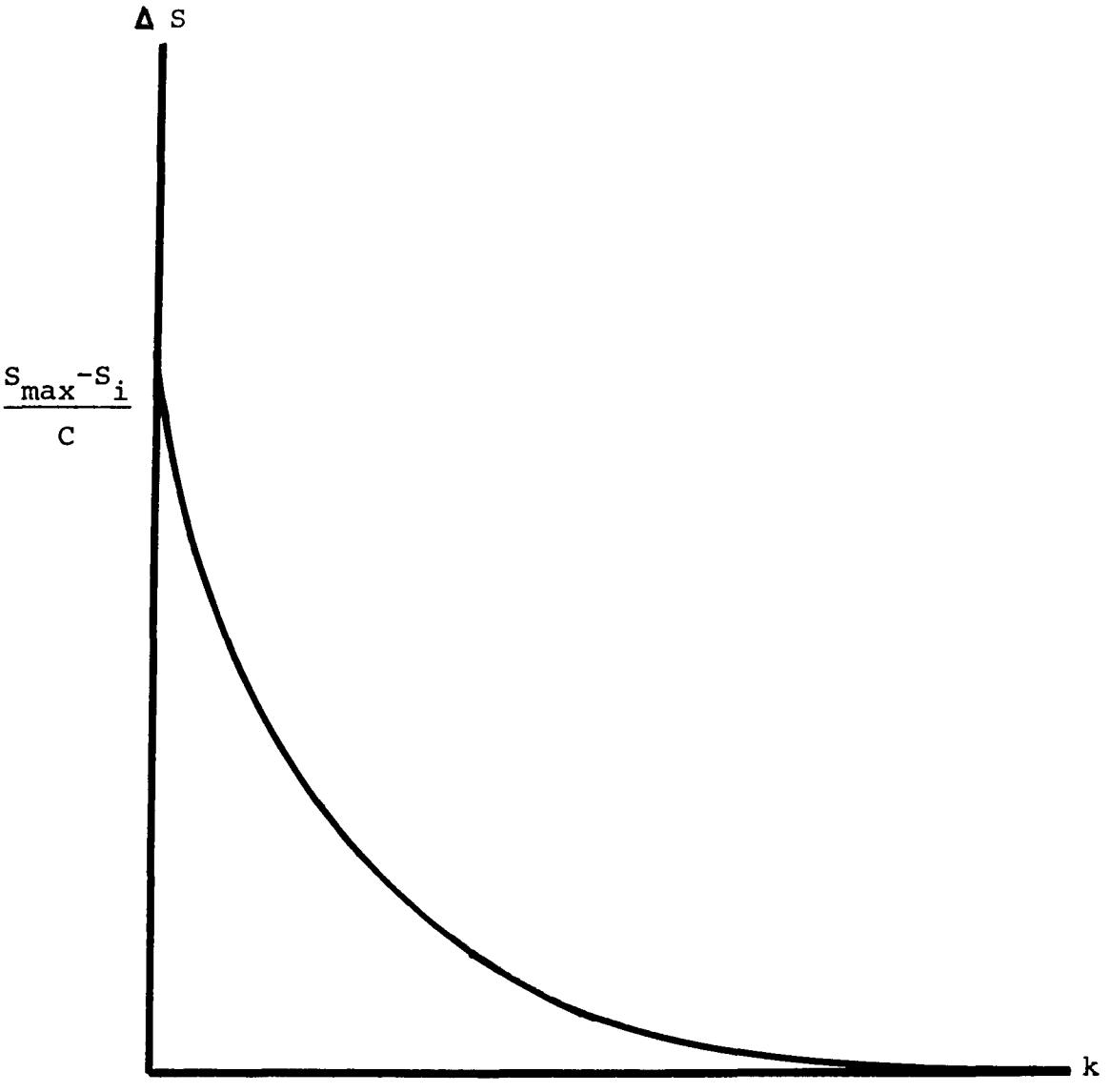


Fig. 2.26 The increment of step size in a A-mode ADM is exponentially decreasing.

(a)

(b)

Fig. 2.27 (a) A-mode ADM encoded picture,
(b) the difference picture of the A-mode ADM.

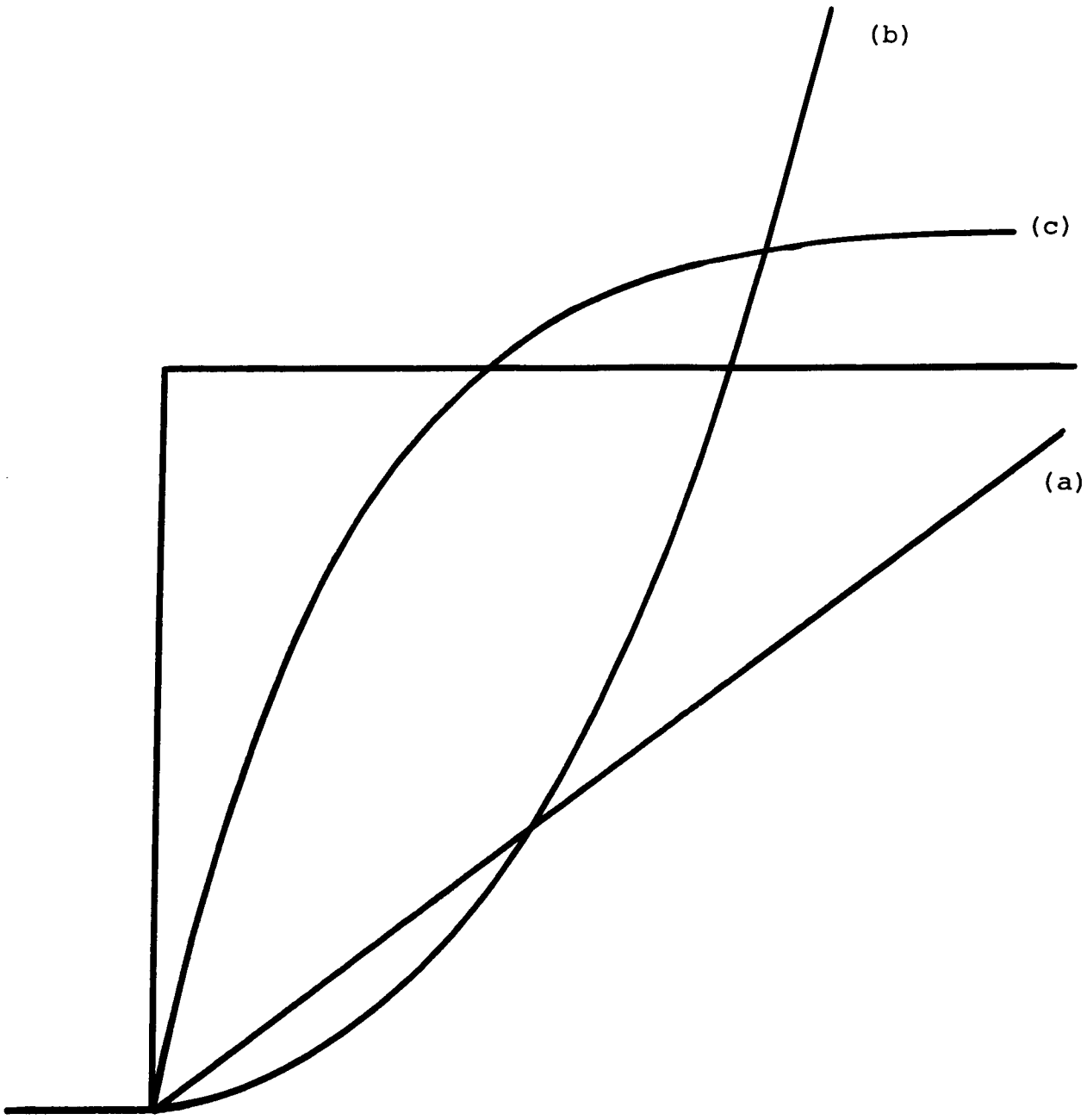


Fig. 2.28 The response curve, to a step input, of (a) the Abate mode ADM, (b) the Song mode ADM, (c) the ideal ADM.

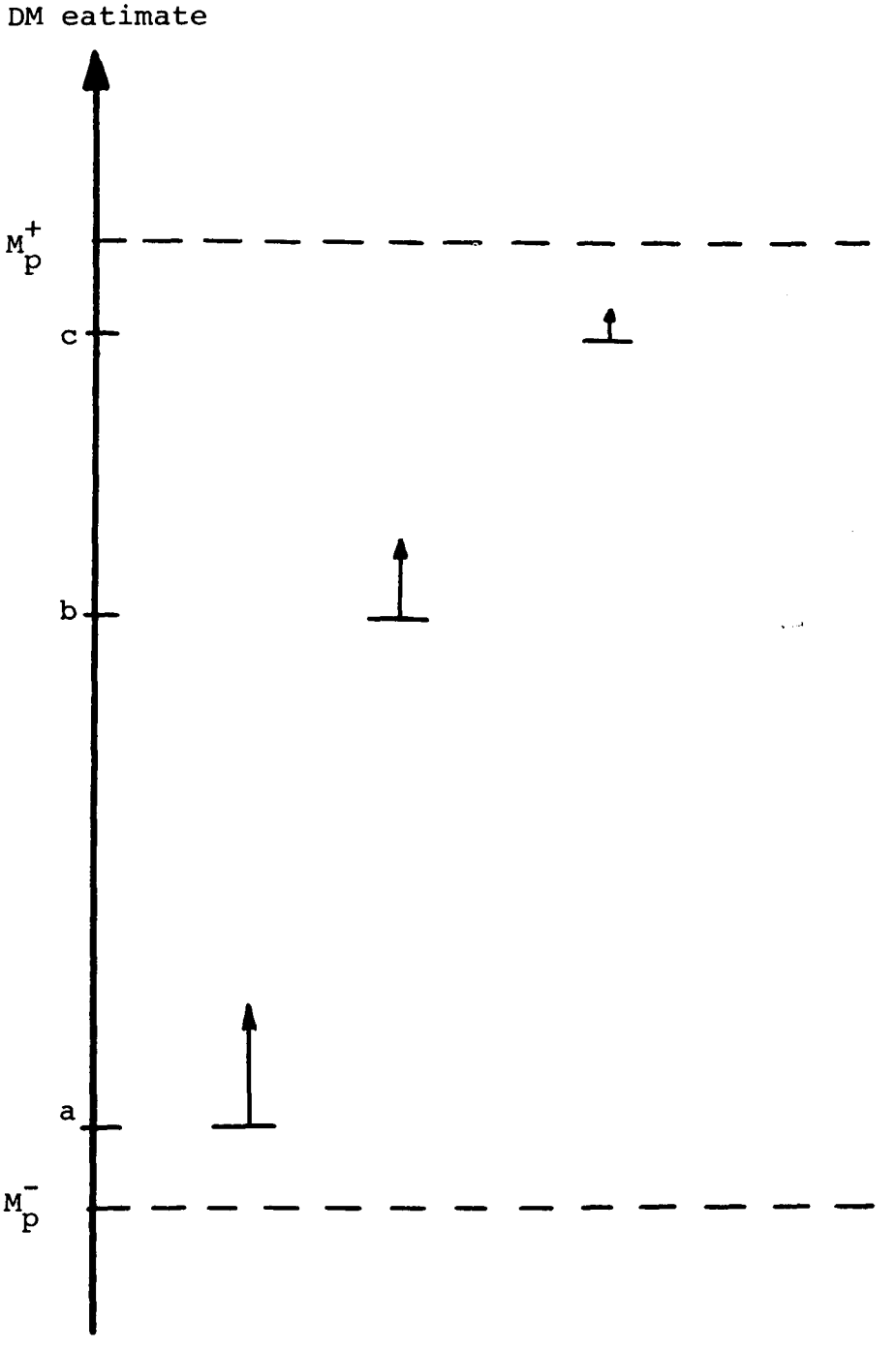


Fig. 2.29 The step size of the B-mode ADM is a function of the present DM estimate.

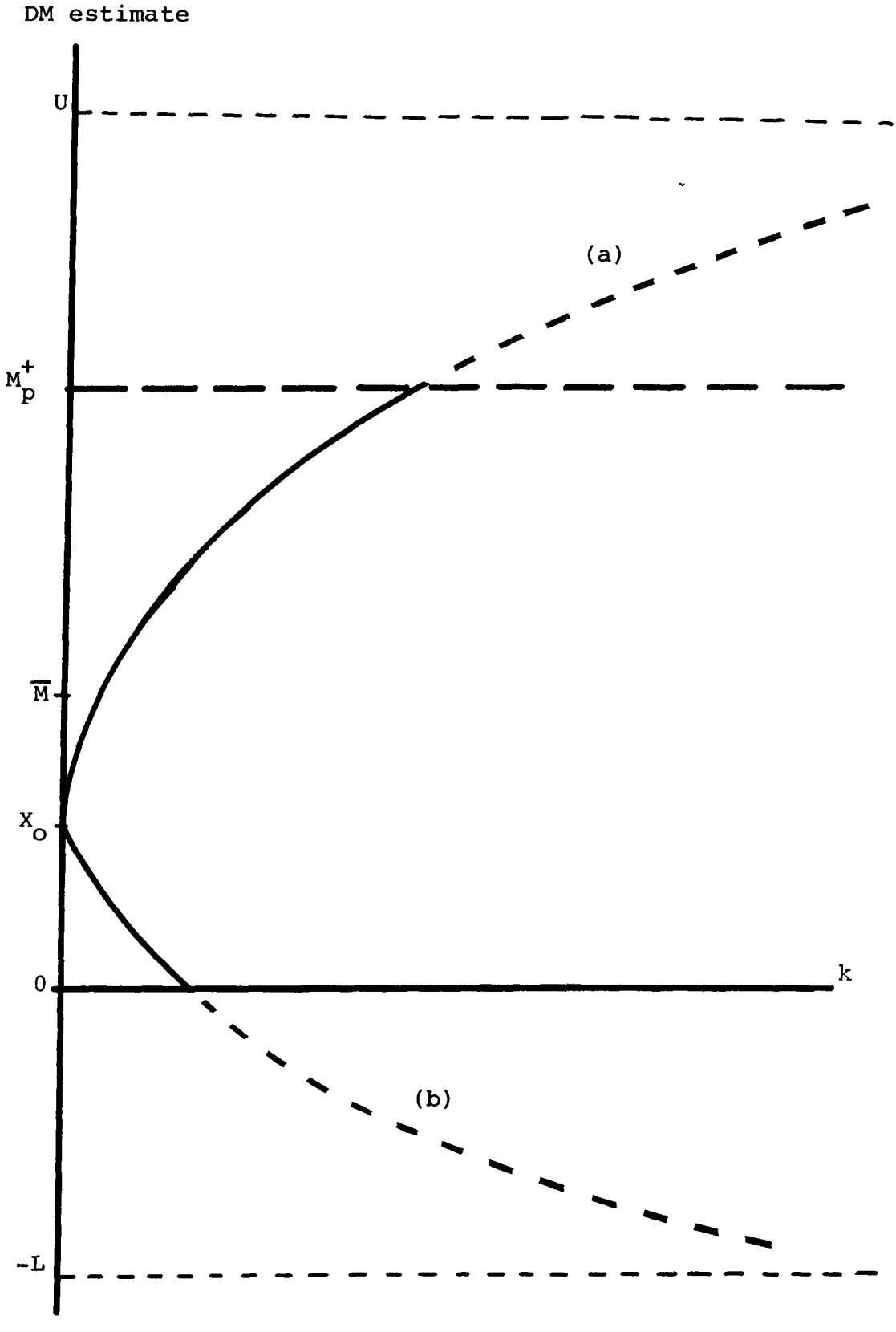


Fig. 2.30 The response curve to a step input of the B-mode ADM, (a) E_k's = +1, (b) E_k's = -1.

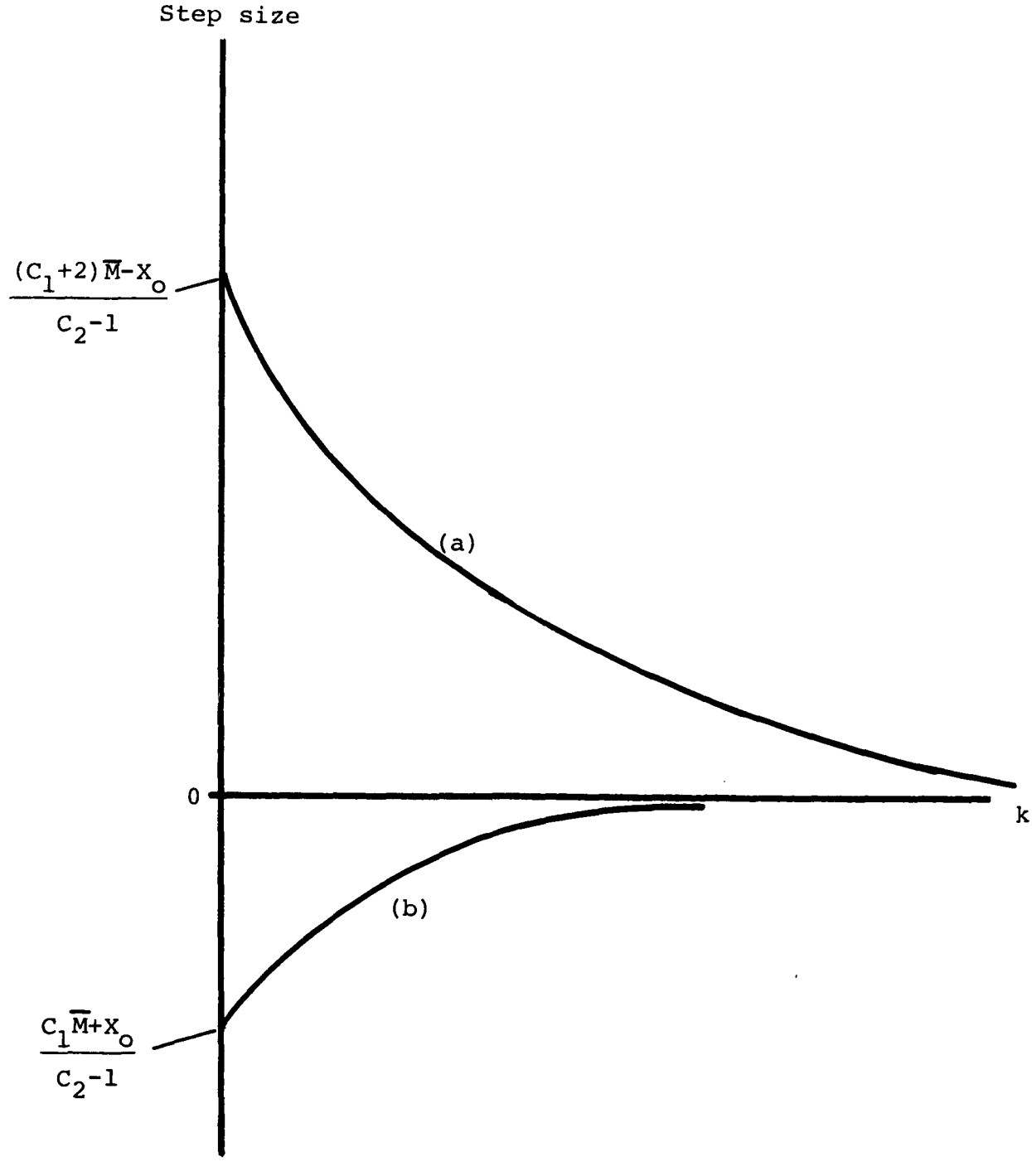


Fig. 2.31 The step size response curve of the B-mode ADM, starting at an initial DM estimate X_o. Curve (a) is for positive step size, curve (b) is for negative step size.

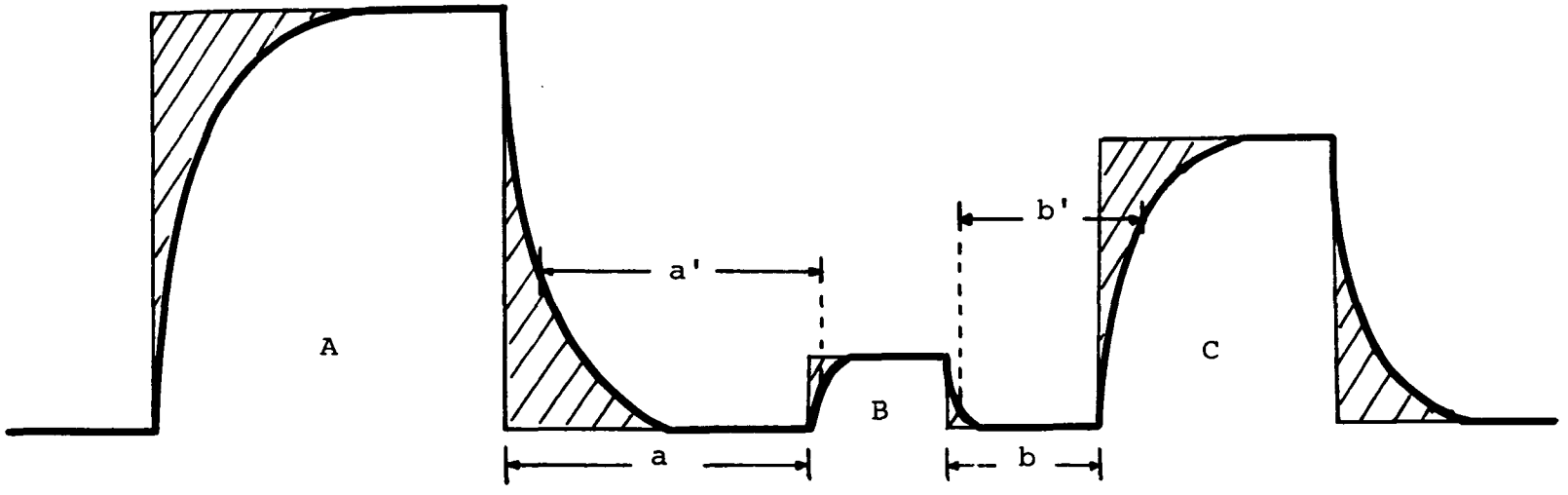


Fig. 2.32 The B-mode ADM generates less distortions in the encoded waveform.

(a)

(b)

Fig. 2.33 (a) B-mode ADM encoded picture,
(b) the difference picture of the B-mode ADM.

(a)

(b)

Fig. 2.34 1 bit/pixel encoded pictures,
(a) Song mode ADM, (b) B-mode ADM.

Chapter III

TWO DIMENSIONAL DELTA MODULATORS

3.1 Introduction

The video signal of a picture generated by the conventional scanning system is a collection of waveforms obtained by connecting sequentially the waveform of each scanning line into a long waveform. The delta modulators described in Chapter 2 treat the video signal as a continuous one dimensional waveform and neglect the waveform correlation between adjacent scanning lines. Thus, this type of DM encoding is called a one dimensional delta modulator (1DDM). The 1DDM encoded picture suffers two types of degradation known as "picture blur" and "edge busyness" as can be seen in the encoded pictures shown in Chapter 2 . The picture blur is due to "slope overloading" of the adjacent picture elements with a sharp change in brightness. The edge busyness is a result of the response of an 1DDM encoder to a vertical edge with different initial conditions. This occurs as a result of connecting each line together end to end thereby destroying the correlation that exists in the vertical axis of the picture.

Picture blur and edge busyness can cause extreme

picture degradation, particularly when the sampling rate is low. Much work has been done to deal with this problem. The edge enhancement technique emphasizes the presence of large brightness change to cope with the effect of picture blur. The overshoot suppression technique [67] and Look-Ahead algorithms [69-71] reduce the amount of overshoot and decrease the rise time thereby somewhat reduce the extent of edge busyness. However, none of these techniques completely eliminate picture blur and edge busyness.

In the research for this dissertation, the two dimensional characteristics of pictures are utilized and several two dimensional delta modulators are studied. The computer simulated results indicate a large amount of reduction in picture blur and edge busyness and a great improvement in picture quality.

Since pictures are two dimensional in nature, the correlation of picture elements is a function of distance between pixels and is not a function of the direction of orientation. One should not emphasize the correlation in one direction and neglect that of another direction as is the case of the 1DDM algorithm. Hence, a delta modulator, when used to encode pictures, should have the ability such that the coding path of the encoder is free to move in a two dimensional domain.

In a conventional scanning system, the picture element

P at row m and column k, as shown in Fig. 3.1, is surrounded by the up-to-date previous estimates α,β,δ and γ. In expanding the one dimensional delta modulator to a two dimensional delta modulator, the estimate of the pixel p should be formed by utilizing all four previous estimates α,δ,β, and γ and their associated error signals and step-sizes, such that the entire reconstructed picture has a minimum quantization error. Such a system is called a two dimensional delta modulator (2DDM). By utilizing all four previous estimates for each pixel the delta modulator will operate extremely well.

A major virtue of a DM is its low cost, therefore, in order not to lose this advantage, the 2DDM system discussed in this dissertation utilized only one vertical previous estimate. The α-δ pair was rejected since the pair lags the signal by one pixel, the 2DDM utilizing this pair cannot respond quickly to a sharp brightness change, thus, results in some degrees of picture blur and edge busyness. The α-γ pair leads the signal by one pixel. It, therefore, senses the brightness change ahead of time and is able to respond to it quickly. It is called the advanced mode (Adv.) 2DDM. However, this advantage of responding to an edge ahead of time becomes a disadvantage when the sampling rate is decreased so as to be equal to or less than the pixel rate, since at these low sampling

rates the predicted values have low correlation, false contours are produced by sensing the edge too early.

The 2DDM system utilizing pixel α and pixel β for the estimation of pixel p gives the best signal-to-noise ratios for all classes of video signals. It is called the nonadvanced mode (N. Adv) 2DDM.

3.2 Normal Mode 2DDM

The algorithm defining the 2DDM system using the α-β pair is given below. The relative positions of the pixels are defined in Fig. 3.2.

Let $D_{H,m,k}$ be the difference between the input sample and the estimate of α

or $$D_{H,m,k} = M_{m,k} - X_{m,k-1} \cdot \tag{3.1}$$

Let $D_{V,m,k}$ be the difference between the input sample and the estimate of β

or $$D_{V,m,k} = M_{m,k} - X_{m-1,k} \cdot \tag{3.2}$$

We define the reference DM estimate $X_{ref.,}$ the reference step size $S_{ref.}$ and the reference error sign $E_{ref.}$ as follows:

$$\left.\begin{aligned} X_{ref.m,k} &= X_{m,k-1} \\ S_{ref.m,k} &= Y_{m,k-1} \\ E_{ref.m,k} &= E_{m,k-1} \end{aligned}\right\} \quad \text{if } D_{H,m,k} \leq D_{V,m,k} \tag{3.3}$$

$$\left.\begin{aligned} X_{ref.m,k} &= X_{m-1,k} \\ S_{ref.m,k} &= Y_{m-1,k} \\ E_{ref.m,k} &= E_{m-1,k} \end{aligned}\right\} \quad \text{if } D_{H,m,k} > D_{V,m,k} \tag{3.4}$$

Then the complete DM alogrithm becomes

$$E_{m,k} = \text{sgn}(M_{m,k} - X_{ref.m,k}) \tag{3.5}$$

$$S_{m,k} = |Y_{ref.m,k}| (E_{m,k} + \frac{1}{2} E_{ref.m,k}) \tag{3.6}$$

$$X_{m,k} = X_{ref.m,k} + S_{m,k} \tag{3,7}$$

From the above equations, it is clear that the difference between a 1DDM and a 2DDM is in the previous references, $X_{ref.m,k}$, $S_{ref.m,k}$ and $E_{ref.m,k}$. In an 1DDM system, the previous reference pixel is always the pixel to the left of the present pixel, i.e., pixel α in Fig. 3.2. In this 2DDM system, the estimates of pixel α and pixel β are compared against the input sample as shown in Fig.3.3. The resultant difference D_H and D_V are then compared, as described in (3.3) and (3.4) to decide whether pixel α or pixel β will be used as the previous reference pixel. The algorithm is based on the assumption that by choosing the previous estimate which is closer to the input sample, a better estimate of the picture will be produced. This is called a Normal mode 2DDM. **Figure 3.4 is the picture encoded by a nonadvanced mode Normal 2DDM.The most noticeable effect** when comparing the 1DDM (Fig. 2.14(a)) and the 2DDM (Fig.

3.4) is seen to be the significant reduction of edge busyness. This effect is even more dramatic at the sampling rate of 1 bit/pixel when comparing Fig. 2.34 with Fig. 3.5. This is because that at any edge the movement of the coding path of a delta modulator is adjusted in the direction of the edge. Therefore, the edge will appear to be smooth and free of wiggle. This is clear if we compare the difference picture of the 2DDM (Fig. 3.6) with the difference picture of the 1DDM (Fig. 2.19(a)). In Fig. 2.19(a) the quantization noise is concentrated only along the vertical edges, but in Fig. 3.6 the noise is dispersed into the horizontal edges, the vertical edges and over the flat areas of the picture. As a result of the redistribution of the quantization noise, a 2DDM can produce higher quality pictures which give human observers a much more pleasant feeling. However, this non-advanced mode Normal 2DDM has a draw-back, it cannot encode edges along a 45 degrees angle. This is because of the nature of the direction of scanning, this 2DDM is not able to predict the input sample along a 45 degrees edge as shown in Fig. 3.7. In Fig. 3.7 the lines at different angles are encoded very well, except the lines at and near 45 degrees. The remedy for this is to use an advanced mode algorithm as explained in Fig. 3.8. In the advanced mode, the vertical reference pixel is advanced in distance by one pixel and is able to sense the change in signal amplitude

ahead of time to respond properly. The encoded "Test Pattern" with the advanced (ADV.) mode is shown in Fig. 3.9. The 45 degree lines are very smooth, however, the vertical edges appear to be jagged. The same effect can be seen from the Adv. mode encoded "Girl" shown in Fig. 3.10 and Fig. 3.11. Figure 3.12 is the Adv. Normal 2DDM encoded "Boy". The advanced mode and the nonadvanced mode each has its own advantages and disadvantages. Based on the laboratory model built at the City College of New York for real time 2DDM image encoding, the advanced mode is favoured by unbiased observers. However, the measurements of SNR from our simulated 2DDM consistantly showed a higher value for the nonadvanced mode. The values of SNR are listed in Table 3.1. An improvement in SNR over the Song mode ADM of up to 4.7 dB was obtained. On the average, the improvement is about 3 dB, which agrees with the results given by O'Neal [24]. Although a 3 dB improvement in SNR is not very high, the real improvements of a 2DDM, however, are the redistribution of noise power and the ability of encoding along edges. The Normal mode 2DDM needs twice as many hardware components to build, but the resultant improvement in picture quality is well worth the effort and the investment.

In a Normal mode 2DDM, two output bits are required, one bit is used to convey the encoding direction, the other is the DM error sign bit. If the picture is transmitted at

a channel rate of 2 bits/pixel, due to the requirement of 2 bits/sample in a 2DDM, the picture is, thus, sampled at the Nyquist rate which is one half as what would be in a 1DDM system. The correlation between the horizontal adjacent pixels is reduced due to the lower sampling rate. This idea of exchanging the pixel correlation with the information contained in the transmitted bits has been discussed earlier in Chapter 2. The results, as shown in Chapter 2, are not good when the Song mode and the B-mode algorithms are combined into one DM system. However, in a Normal mode 2DDM, instead of combining two ADM algorithms together, the DM system moves to the previous scanning lines to seek for a better pixel correlation. By doing this, the system performance is improved. In other words, the gain by utilizing the correlation of the previous lines is greater than the loss by the reduction in horizontal correlation cuased by halving the sampling rate.

3.3 Look-Ahead 2DDM

In the Normal mode 2DDM system, we assumed that choosing the previous estimate which is closer to the input sample will result in a better estimated picture. This assumption is not always correct. In Fig. 3.13, for a Normal mode 2DDM system, $X_{m-1,k}$ and $X_{m,k-1}$ are compared with $M_{m,k}$, therefore, $X_{m-1,k}$ would be chosen as $X_{ref.m,k}$ and the

error signal would be "down", consequently, $X^{-}_{m-1,k}$ would be the estimate for the input sample $M_{m,k}$. However, in Fig. 3.13 it is clear that $X^{+}_{m,k-1}$ is a better estimate than $X^{-}_{m-1,k}$ for $M_{m,k}$. The delta modulator would have made a better estimate if it was allowed to "look" one clock period ahead, thereby delaying its decision by one clock pulse. In other words, the delta modulator should wait until all four new estimates have been generated and then make the decision based upon the new estimates instead of the previous estimates. This look-ahead scheme, if generalized, by employing the tree search technique for the entire picture, will allow the best possible coding path to be picked. However, for the reason of cost effectiveness only the one bit look-ahead scheme was studied in detail. It is interesting to note that Cutler [69] pointed out that look-ahead (delayed decision) does not greatly improve the system SNR. It does, however, prevents the estimate from overshooting, thereby reduces the amount of edge busyness. Hence, the picture quality is improved.

The improvement of the DM system by using the look-ahead scheme is mainly in the prevention of overshoot. The noise power is reduced due to the elimination of overshoot noise. Consequently, the SNR is increased as shown in Table 3.2. However, the improvement in the encoded picture quality by using the look-ahead scheme, does not seem significant

enough to be noticeable. This is because the Normal mode 2DDM algorithm already eliminates most of the edge busyness, any further reduction in edge busyness is too insignificant to be subjectively detected. It can only be revealed by the measurement of signal-to-noise ratio. The nonadvanced mode Look-Ahead 2DDM encoded "Girl" is shown in Fig. 3.14.

Another noticeable difference between the Normal mode 2DDM and the Look-Ahead 2DDM is the statistics of the step size. Because of the lack of overshoot, which requires larger step size, the step size of the Look-Ahead 2DDM tends to have a smaller value. This is demonstrated in Table 3.3, Table 3.4 and Fig. 3.15. In fact, both the Normal mode 2DDM and the Look-Ahead 2DDM use small step sizes most of the time and larger step sizes are seldom used. This explains why the two dimensional delta modulators are not sensitive to the restriction of maximum step sizes. Table 3.5 shows the values of SNR of the Normal mode 2DDM with various maximum step sizes. The SNR function is almost flat regardless of the S_{max} used. For the "Girl", even a maximum step size of $4S_0$ gives a reasonable SNR as shown in Fig. 3.16. The dynamic range is, as expected, wide and flat, as shown in Table 3.6. It is also depicted in Fig. 3.17.

The Look-Ahead 2DDM requires almost twice as many electronic components to build as the Normal mode 2DDM. From the academic point of view, the Look-Ahead 2DDM

algorithm gives a superior performance. From the practical point of view, however, it remains a personal judgement as which of these two algorithms is of the best interest.

3.4 3-State 2DDM

In a Normal mode or a Look-Ahead 2DDM, one extra output bit is required to carry the message of the encoding direction, as a result, a total of 2 output bits per sample are needed. The bandwidth required for the DM system is, hence, doubled. Bandwidth is precious and we want to reduce as much bandwidth in a DM system as possible. Entropy encoding is a well known technique for this which will be discussed in Chapter 4. Another technique is to skillfully modify the code words of the 2DDM output bits by discarding some of the code words. By doing this, the subsequent source encoder can obtain a higher data compression ratio, thereby causing the DM system to occupy less bandwidth. In a 2DDM system, there are four output states; 11, 10, 01 and 00. If we can discard the least significant state and transmit only three states, the DM output bits per sample will be, $\log_2 3 = 1.6$ bits, instead of the original 2 bits. A twenty percent saving is readily obtained. In order to do this, the 2DDM scheme has to be modified. At every pixel, the DM estimates of the previous horizontal pixel and the vertical pixel are compared to determine which one is greater. The greater one

will be called X_{max} and the smaller one will be called X_{min} as illustrated in Fig. 3.18. If the input sample is greater than X_{max}, output bits "11" is transmitted. If the input sample is smaller than X_{min}, output bits "00" is transmitted. If the input sample falls in between X_{max} and X_{min}, output bits "10" will be sent out. The sequence "01" is not used. The reason for this is that most of the time the estimates of the previous horizontal and vertical pixels are very close; if the input sample is in between these two estimates, it does not make much difference which one of these two previous estimates is used as the reference estimate. However, at an edge, X_{max} and X_{min} could conceivably be far apart, then, the "past history" of the neighboring pixels are used to help make the decision. Each pixel is associated with a "history bit" as shown in Fig. 3.19. An "H" indicates that this pixel was encoded from the greater estimate of its two previous estimates, a "L" indicates that this pixel was encoded from the smaller estimate of its two previous estimates. Whenever the value of the input sample is within the two previous estimates, for example, take pixel p in Fig. 3.19, the history bits of its three neighboring pixels, α, β and δ are compared by a majority rule. In Fig. 3.19, two H's and one L are found, therefore, the result is an "H" . In other words, pixel p will be encoded from the greater previous DM estimate, X_{max}, of the pixels α and β.

This kind of encoding will inevitably create errors in the decision of the correct coding path. As a result, edge busyness will appear. However, if the decision error does not occur very often and if the length of edge busyness is not long, a good quality picture can still be produced. Since this scheme utilizes only three output states, it is call a 3-State 2DDM.

The 3-State 2DDM algorithm was first simulated by using a low speed flying spot scanner. The results have been presented at the International Symposium on Information Theory at Ronneby Sweden [72]. The results were not good. The encoded pictures showed notches along the edges as a result of incorrect decision on the encoding path. It was, then thought that the 3-State 2DDM algorithm is not a good algorithm. However, since the scan converters were installed for the DM simulation, the higher resolution equipments revealed that the 3-State 2DDM is indeed a good approach toward narrow bandwidth, high quality video transmission. The notches which were clearly seen in the 170 pixels/line flying spot scanner have disappeared due to the much reduced pixel space of real time TV system. At a nominal resolution of 500 pixels/line in a TV system, the notches along the edges become too small to be a serious defect of the algorithm. They just look like occasional edge busyness and are hardly noticeable, as can be seen from the 3-State 2DDM

encoded "Girl" in Fig. 3.20. The measured output SNR of the 3-State 2DDM is shown in Table 3.7. Notice that the SNR and the picture are obtained from a 3-State 2DDM sampled at the Nyquist rate, which is equivalent to a channel rate of 1.6 bits/pixel, whereas the SNR of other 2DDMs shown in Table 3.1 and Table 3.2 are taken at a channel rate of 2 bits/pixel. For a fair comparison, the sampling rate of the 3-State 2DDM should be 640 pixels/line, in order to obtain an equivalent channel rate of 2 bits/pixel. Unfortunately, our DM simulation system cannot generate this sampling rate. The SNR of the 3-State 2DDM at 1.6 bits/pixel is about 4 dB less than that of the Normal mode 2DDM at 2 bits/pixel. It is our believe that if the 3-State 2DDM is allowed to be sampled at 2 bits/pixel, its SNR will be as good as that of the Normal mode 2DDM. By examining the 3-State 2DDM encoded picture in Fig. 3.20 we further confirm our belief.

3.5 Weighted-Average 2DDM

Using one extra bit per sample for the transmission of encoding direction in the Normal mode and the Look-Ahead 2DDM may seem to be a waste of channel bandwidth. 3-State 2DDM was developed using only a fraction of a bit, on the average, to transmit the information of the encoding direction. In this section, we will discuss a 2DDM scheme, in which the direction of encoding is predicted from the

neighboring estimates, and thus, the directional bit is not transmitted at all. This is called a Weighted-Average 2DDM. A predicted value for an input sample is first generated from its neighboring estimates, then this predicted value is used instead of the actual input sample to compare with the previous horizontal and vertical pixels to determine the direction of encoding. Predicted input sample may differ from the actual input sample by a large quantity, resulting in an incorrect encoding direction, thus, messes up the edges. However, because of the savings from the elimination of the directional bit, the DM system can be sampled at twice the rate for the same channel bandwidth. Sampling at faster rate has two advantages. Firstly, the pixels become closer, the correlations among pixels are greater, thus, a better prediction can be obtained to avoid incorrect decisions on encoding directions. Secondly, when the pixel distance is decreased, so is the visibility of the jaggedness along the edges. In this section, we will discuss the results of two dimensional DM encoding using the weighted-average schemes.

The purpose of this study is two fold. One is to seek other alternatives for two dimensional encoding, the other is to determine whether or not the extra directional bit is essential, and whether the information of the encoding direction can be accurately predicted from the previous pixels.

A Weighted-Average 2DDM is basically the same as a Normal mode 2DDM, with the exception that in a Weighted-Average 2DDM, a predicted value is used to replace the actual input sample to determine the encoding directions. We use three previous pixels to predict the input sample. Refer to Fig. 3.1 for pixel positions. The prediction scheme is

$$\overline{P} = a\overline{\alpha} + b\overline{\beta} + c\overline{\delta} \tag{3.8}$$

where $\overline{\alpha}$, $\overline{\beta}$, and $\overline{\delta}$ denote the DM estimates of pixels α, β and δ respectively, and $\overline{P}$ denotes the predicted input sample for pixel p; a, b and c are the prediction coefficients to be determined. The prediction coefficients have to be properly chosen such that the predicted input $\overline{P}$ has to be as close to the actual input M as possible. The technique of choosing the prediction coefficient is discussed in O'Neal's paper [24]. We will briefly describe his technique here. The prediction error, e, between the actual input sample, M, and the predicted value, $\overline{P}$, is defined as

$$e = M - \overline{P} \tag{3.9}$$

where $\overline{P}$ is,as defined in Eq. (3.8),

$$\overline{p} = a\overline{\alpha} + b\overline{\beta} + c\overline{\delta}.$$

In order to minimize the prediction error, we have to minimize the expected value of the squared error. To find the values for a, b and c which satisfy the minimization condition, we take the partial derivation with respect to

a, b and c and let the derivatives equal to zeros,

$$\frac{d\, E[(M - \bar{P})^2]}{da} = 0 \tag{3.10}$$

$$\frac{d\, E[(M - \bar{P})^2]}{db} = 0 \tag{3.11}$$

$$\frac{d\, E[(M - \bar{P})^2]}{dc} = 0 \ . \tag{3.12}$$

Replacing $\bar{P}$ using Eq. (3.8) , we obtain

$$\frac{d\, E[(M - (a\bar{\alpha} + b\bar{\beta} + c\bar{\delta}))^2]}{da} = 0$$

or

$$- 2E[(M - (a\bar{\alpha} + b\bar{\beta} + c\bar{\delta}))\bar{\alpha}] = 0$$

or simply,

$$E[(M - \bar{P})\bar{\alpha}] = 0 \ . \tag{3.13}$$

If the video signal has a zero mean value, i.e.,

$$E[M] = 0,$$

then the covariance of pixels p and α can be written as,

$$R_{p\alpha} = E[\bar{P}\,\bar{\alpha}] \tag{3.14}$$

rearranging Eq. (3.13) we get

$$\begin{aligned} E[M\,\bar{\alpha}] &= E[\bar{P}\,\bar{\alpha}] \\ &= E[(a\bar{\alpha} + b\bar{\beta} + c\bar{\delta})\bar{\alpha}] \\ &= aE[(\bar{\alpha})^2] + bE[\bar{\beta}\,\bar{\alpha}] + cE[\bar{\delta}\,\bar{\alpha}]. \end{aligned} \tag{3.15}$$

Placing the covariance notation, Eq. (3.14), into Eq. (3.15) we have

$$E[M\,\bar{\alpha}] = aR_{\alpha\alpha} + bR_{\beta\alpha} + cR_{\delta\alpha} \ . \tag{3.16}$$

Since the $\bar{P}$ generated from this optimizing procedure will be very close to M, we can make the following approximation,

$$E[M\ \bar{\alpha}] \simeq E[\bar{P}\ \bar{\alpha}]. \tag{3.17}$$

Placing Eq. (3.17) into Eq. (3.16) and dropping the approximation sign, we get,

$$E[M\ \bar{\alpha}] = E[\bar{P}\ \bar{\alpha}] = R_{p\alpha} = aR_{\alpha\alpha} + bR_{\beta\alpha} + cR_{\delta\alpha}. \tag{3.18}$$

We can obtain a similar conclusion for pixel β and δ. Hence, a set of equations is derived for the optimization of the values a, b and c;

$$\begin{aligned} R_{p\alpha} &= aR_{\alpha\alpha} + bR_{\beta\alpha} + cR_{\delta\alpha} \\ R_{p\beta} &= aR_{\alpha\beta} + bR_{\beta\beta} + cR_{\delta\beta} \\ R_{p\delta} &= aR_{\alpha\delta} + bR_{\beta\delta} + cR_{\delta\delta}. \end{aligned} \tag{3.19}$$

Equation (3.19) indicates that if the covariances among pixels α, β, δ and p are known, we can find the optimum values for a, b and c in the prediction function, Eq. (3.8).

In this dissertation, we did not follow Eq. (3.19) to find the best set of values for a, b and c, but rather, a much simplified version of Eq. (3.19) was used for the following reasons. Firstly, the covariance function varies from picture to picture, there is hardly a typical picture which can represent a general class of video signals. Therefore the prediction coefficients that are calculated precisely from the covariance function of a particular slide have

little practical value except some degrees of academic interest. Secondly, even if a good set of prediction coefficients can be found, we still have to simplify and truncate those values for hardware implementation. Furthermore, O'Neal has, in his paper, given the covariance coefficients of certain scenes from his experiment. There is little value to repeat his experiment again. Therefore, we use the following approach to obtain the prediction coefficients for our interest.

Refer to Fig. 3.1, we assume that the covariances of adjacent pixels are the same, and that the covariances of diagonal pixels are also the same. In other words, we assume,

$$R_{p\alpha} = R_{p\beta} = R_{\alpha\delta} = R_{\beta\delta} = R_1\sigma^2$$

$$R_{\alpha\beta} = R_{p\delta} = R_2\sigma^2$$

and $$R_{\alpha\alpha} = R_{\beta\beta} = R_{\delta\delta} = \sigma^2$$

where σ^2 is the variance of the amplitude of pixels. We further assume that the horizontal and the vertical previous pixels, α and β, are equally important in the prediction scheme, therefore, they will have the same prediction coefficients, i.e.,

$$a = b. \tag{3.20}$$

Moreover, we normalize the covariances by letting the variance to be equal to 1, i.e.,

$$\sigma^2 = 1. \tag{3.21}$$

Equation (3.19) is, therefore, simplified as

$$\begin{aligned} R_1 &= (R_2 + 1)a + R_1 c \\ R_2 &= 2R_1 a + c \end{aligned} \tag{3.22}$$

where R_1 and R_2 are the covariance coefficients for the adjacent pixels and the diagonal pixels respectively. However, we still cannot see the value of finding these coefficients for our test slides from Eq. (3.22) for the same reasons stated earlier. We simply used Eq. (3.22) to determine the range of the prediction coefficients, a, b and c.

Solving Eq. (3.22) for a and c, we obtain,

$$a = \frac{R_1(1 - R_2)}{1 + R_2 - 2R_1^2} \tag{3.23}$$

$$c = \frac{R_2(1 + R_2) - 2R_1^2}{1 + R_2 - 2R_1^2} \tag{3.24}$$

Typical values of covariance coefficients given by O'Neal and others are

$$R_1 \simeq 0.9$$

$$R_2 \simeq 0.83.$$

Place these values of R_1 and R_2 into Eq. (3.23) and Eq. (3.24), we obtain,

$$a = 0.73$$

and $$c = -0.48.$$

In fact, the covariance coefficients vary within a large range of values. For instance, R_1 may vary between 0.82 and 0.96, R_2 may vary between 0.75 and 0.92, as given by O'Neal. Therefore, the values of a and c, by solving Eq. (3.23) and Eq. (3.24), also vary within a large range of values. In our experiment we chose,

$$a = b = 0.75$$

and $$c = -0.5$$

for the following reasons. First, these prediction coefficients are simple to implement. Second, the sum of a, b and c is equal to 1, which is essential in a prediction scheme. Third, these coefficients give rise to a sensitive edge detection 2DDM algorithm. Furthermore, this set of coefficients has been used by Lippman [105] and others in their DPCM simulations. But most of all, this set of coefficients can save a lot of computation time; because if

$$\overline{P} = 0.75\overline{\alpha} + 0.75\overline{\beta} - 0.5\overline{\delta} \qquad (3.25)$$

it will not be necessary to actually compute $\overline{P}$, and to compare $\overline{P}$ with $\overline{\alpha}$ and $\overline{\beta}$ in order to determine the encoding direction. A simple comparison of the differences between $\overline{\delta}$ and $\overline{\alpha}$, and $\overline{\delta}$ and $\overline{\beta}$ has the same effect. This is shown in the paper given by Lei, et al. [73], and is duplicated here.

The DM will choose pixel α as the reference pixel if α is closer to $\overline{P}$, i.e.

$$|\overline{P} - \overline{\alpha}| < |\overline{P} - \overline{\beta}|$$

or $$(\overline{P} - \overline{\alpha})^2 < (\overline{P} - \overline{\beta})^2 . \qquad (3.26)$$

Equation (3.26) can be expanded as

$$-2\overline{\alpha}\overline{P} + \overline{\alpha}^2 < -2\overline{\beta}\overline{P} + \overline{\beta}^2$$

Replacing $\overline{P}$ by the weighting equation (3.25), we get

$$-0.5\overline{\alpha}^2 - 1.5\overline{\alpha}\overline{\beta} + \overline{\alpha}\overline{\delta} < -.0.5\overline{\beta}^2 - 1.5\overline{\alpha}\overline{\beta} + \overline{\beta}\overline{\delta}$$

thus, $$0.5(\overline{\alpha}^2 - 2\overline{\alpha}\overline{\beta}) > 0.5(\overline{\beta}^2 - 2\overline{\beta}\overline{\delta})$$

or $$(\overline{\alpha} - \overline{\delta})^2 > (\overline{\beta} - \overline{\delta})^2 . \qquad (3.27)$$

Equation (3.27) indicates that pixel α will be chosen as the reference pixel if $\overline{\delta}$ is farther away from $\overline{\alpha}$ than $\overline{\beta}$. therefore, the whole weighted-average prediction becomes a simple amplitude comparison. The reference pixel will be the one whose estimate is farther away from the estimate of the diagonal pixel . In other words, pixel α will be the reference pixel if $|\overline{\delta} - \overline{\alpha}| > |\overline{\delta} - \overline{\beta}|$, and pixel β will be the reference pixel if $|\overline{\delta} - \overline{\alpha}| < |\overline{\delta} - \overline{\beta}|$. This is a good edge detecting prediction algorithm. It enables the DM to respond to steep slopes quickly.

In fact, the above conclusion of the encoding direction is valid as long as the coefficients of Eq. (3.8) satisfy the following conditions,

$$a = b$$

$$a + b + c = 1 \qquad (3.28)$$

and c is negative.

If c is positive, however, the result is just the opposite; the pixel whose estimate is closer to $\bar{\delta}$ will be the reference pixel. The main factor for this dramatic change is the different covariance coefficients of the pixels. If the pixels have low covariance coefficients, by solving Eq. (3.19) or the simplified version, Eq.(3.24), the prediction coefficient c will have a negative value. On the other hand, if the values of the covariance coefficients are high, c will be a positive value. With a negative c, the DM will be seeking for edges by encoding from the pixel which has a different amplitude from the others. This is reasonable because low covariance coefficients indicate that the video samples are not highly correlated, hence, the DM should concentrate on encoding edges. However, with a positive c, which is the result of high covariance coefficients, the DM will encode from the pixel which is not far away from others due to the fact that the video signal is sampled from a relatively smooth picture.

To obtain a positive value for c in our simplified version, Eq. (3.24), the values of the covariance coefficients, R_1 and R_2 have to be very high, for instance, $R_1 = 0.96$ and

R_2 = 0.95. These values are far greater than the covariance coefficients of our 2 bits/pixel simulation system. Therefore c is always a negative number as far as our experiment is concerned.

The resulting values of SNR of the Weighted-Average 2DDM are listed in Table 3.8; we can see that the SNR of the Weighted-Average 2DDM is less than that of the Normal mode 2DDM and the Look-Ahead 2DDM. The encoded picture is shown in Fig. 3.21. This Weighted-Average 2DDM is sampled at twice the Nyquist rate and transmits one output bit per sample, resulting in a channel rate of 2 bits/pixel. From the encoded picture and the measured SNR, we can conclude that the directional bit used in the Normal mode 2DDM and the Look-Ahead 2DDM is not a waste, but rather provides essential information for generating high quality pictures with high SNR. The predicted encoding direction is too frequent to be in error to generate good pictures. However, the Weighted-Average 2DDM still has an improvement over the one dimensional Song mode ADM.

The resulting low SNR of the Weighted-Average 2DDM led us to think that more encoding directions may be needed in order for the DM to select the best possible reference pixel. For this reason, one more encoding direction is added to the Weighted-Average 2DDM algorithm. In this scheme, the

the predicted input sample $\overline{P}$ is actually calculated, and compared with the previous estimates $\overline{\alpha}$, $\overline{\beta}$ and $\overline{\delta}$. Whichever is closest to $\overline{P}$ will be chosen as the previous estimate. In this case, the encoding direction will be free to move in any of the three directions; a horizontal direction, a vertical direction and a diagonal direction of -45 degrees. Unfortunately, this added degree of freedom has not significantly improved the SNR as shown in Table 3.8 under the heading of 3-directional Weighted-Average 2DDM. This is because that this extra encoding direction is, in fact, a redundancy. Since the horizontal scanning of a television is from left to right, this -45 degrees encoding direction from pixel δ cannot provide any more information about the future events than the other two directions can.

4-directional Weighted-Average 2DDM algorithm was also experimented; in which $\overline{P}$ is obtained by a weighted average of four pixels α, δ, β and γ . Refer to Fig. 3.19 for pixel positions, there are four encoding directions for pixel p. In this case one would think that any large change in input samples can be detected earlier by pixel γ, so that a much better prediction can be obtained. However, contrary to this believe, the resulting SNR is not improved as can be seen in Table 3.8. The reason is that too many degrees of freedom destroys the coordination between adjacent step sizes which is an essential requirement for a successful adaptive step

size algorithm. Although more degrees of freedom allows a better choice of encoding direction, it also prevents the same encoding direction to last for many pixels in order for the adaptive step size algorithm to properly adapt its step size to the input signal characteristics. A possible remedy is to use the look-ahead scheme to generate eight possible new estimates and to select the best one. However, this will make the DM system extremely complicated. Besides, in Chapter 4 we will see that the Weighted-Average 2DDM cannot tolerate channel errors and is not a good practical DM system. Therefore no further attempt is carried out.

3.6 Prediction 2DDM

As mentioned in Chapter 1, both DM and DPCM are feedback predictive encoders. One of the many differences between them is that the prediction coefficient of the predictor is "1" in a DM, whereas in a DPCM, the coefficients are calculated to optimize the performance. In this section, the prediction coefficients other than "1" is employed in a DM and it is called the Prediction 2DDM. The prediction equation is chosen to be the same as that of the Weighted-Average 2DDM,i.e.,

$$\overline{p} = 0.75\overline{\alpha} + 0.75\overline{\beta} - 0.5\overline{\delta}.$$

The DM algorithm equations then become

$$E_k = \mathrm{Sgn}[M_k - \bar{P}_k]$$

$$S_k = |S_{k-1}| E_k + S_0 E_{k-1} \qquad (3.29)$$

$$X_{k+1} = \bar{P}_k + S_k$$

where $\bar{P}_k$ is the predicted value of the k-th input sample.

This Prediction 2DDM is basically a one dimensional Song mode ADM except that the predicted value instead of the current estimate is used in determining the error sign bit and the next estimate. The resulting performance of the Prediction 2DDM is shown in Table 3.9. Although the SNRs in Table 3.9 are not very high, the encoded picture shows a very smooth transition between the edges and is, thus, free of edge busyness as can be seen in Fig. 3.22.

3.7 Conclusions

Two dimensional delta modulators utilize the correlation of the previous line for the encoding of video signals, resulting in an improvement in the DM performance. There are three techniques to utilize the information contained in the previous line. One is to transmit an extra bit or a partial of an extra bit to convey a more accurate encoding direction, such as the Normal mode, the Look-Ahead and the 3-State 2DDM algorithms. One is to use the predicted value to replace the actual input sample to control the encoding direction such as the Weighted-Average 2DDM algorithms.

The third one encodes from one direction only but it uses the predicted estimate to determine the DM error sign bit and to generate new estimate such as the Prediction 2DDM algorithm. All three techniques utilize the correlation between scanning lines, thus the performance of these two dimensional delta modulators are better than those of the one dimensional delta modulators. Consult Table 3.10 for a general reference list of all the DM performance. We can see a definite improvement of SNR in the two dimensional delta modulators. The frame differences, as another measure of merit,are listed in Table 3.11. The SNR performance as a function of input power is listed in Table 3.6 and depicted in Fig. 3.17.

In this dissertation, we assume that the video system is scanned line by line one after another. However, the common commercial scanning system is interlaced [81,82]. This interlaced scanning system causes the previous line to be two lines apart from the current one, thereby reduces the correlation between lines. The additional information of video signal obtained by utilizing the previous lines is, therefore, reduced and so is the improvement of the two dimensional delta modulators. The SNRs of the two dimensional delta modulators in an interlaced scanning system is listed in Table 3.12. However, the scanning system can always be modified for special applications to get the best out of

the two dimensional DM system.

The two dimensional DM algorithms discussed in this chapter can all perform well, but from the practical point of view, the Normal mode 2DDM is the most practical and realizable system. The others either are too complicated in their hardware implimentation for present state of art, or require a follow-up entropy encoder, or do not have good channel error responses. In the subsequent chapters, we will use the Normal mode 2DDM as a representative of the two dimensional DM system. Unless otherwise specified, the term 2DDM will be used for the Normal mode 2DDM only, just as the term 1DDM is used for the Song mode one dimensional ADM.

	Girl	Boy	Test Pattern
N.Adv.	32.7	28.7	23.3
Adv.	30.5	27.8	22.9

Table 3.1 SNR (dB) of Normal Mode 2DDM.

	Girl	Boy	Test Pattern
N.Adv.	33.4	29.4	23.0
Adv.	30.4	27.8	22.9

Table 3.2 SNR (dB) of Look-Ahead 2DDM.

Boy Step size	N.Adv. Normal Mode		N.Adv. Look-Ahead	
	No. of occurrence	Probability	No. of occurrence	Probability
1	134526	0.5172	161250	0.6200
2	70553	0.2713	54599	0.2100
3	25516	0.0981	20234	0.0780
4	13092	0.0503	11181	0.0430
6	7960	0.0306	6769	0.0260
8	843	0.0032	771	0.0030
9	3813	0.0147	3146	0.0120
12	541	0.0021	200	0.00077
13	1852	0.0071	1573	0.0060
16	1405	0.0054	887	0.0034

Table 3.3 Step size statistics of 2DDM , encoding the "Boy", with $S_{max} = 16S_o$. Total number of pixels is 260,100.

Boy, step size	N.Adv. Normal Mode No. of occurrence	Probability	N.Adv. Look-Ahead No. of occurrence	Probability
1	34528	0.5172	161700	0.622
2	69986	0.2690	54411	0.210
3	25557	0.0980	20098	0.0773
4	13029	0.0500	10898	0.0419
5	131	0.0005	70	0.00027
6	7833	0.0300	6562	0.0252
7	273	0.0009	193	0.00074
8	12	0.00005	4	0
9	4462	0.0170	3758	0.0144
10	221	0.00085	86	0.00033
11	18	0.00007	3	0
12	32	0.00012	3	0
13	2136	0.0082	1711	0.0066
14	217	0.00083	182	0.0007
15	112	0.00043	18	0.0007
16	50	0.00019	4	0
18	8	0.00003	1	0
19	938	0.0036	698	0.00268
21	149	0.00057	23	0.00009
22	23	0.00009	4	0
24	38	0.00015	2	0
27	0	0	1	0
28	301	0.00116	159	0.00061
31	21	0.00008	5	0
32	62	0.00024	2	0

Table 3.4 Step size statistics of 2DDM, encoding the "Boy", with $S_{max} = 32S_o$. Total number of pixels is 260,100.

	Maximum step size, S_{max}		
	$8S_o$	$16S_o$	$32S_o$
Girl	32.7	32.1	31.7
Boy	28.7	28.6	28.5
Test Pattern	23.3	23.1	23.1

Table 3.5 SNR (dB) of N.Adv. Normal Mode 2DDM with different S_{max}.

	Input signal power (dB)			
	+6	0	-6	-12
N.Adv. Look-Ahead	23.5	29.4	29.5	28.4
3-State	22.2	23.6	23.4	22.8
Weighted-Average	23.3	27.4	26.0	24.8

Table 3.6 SNR (db) of two dimensional DM with different input signal power.

	N.Adv. Normal Mode	Adv. Normal Mode	N.Adv. Look-Ahead	Adv. Look-Ahead
Girl	28.8	25	28.3	26.8
Boy	24.5	20.1	23.6	22.4
Test Pattern	21.2	16.1	20.2	18.8

Table 3.7 SNR (db) of 3-State 2DDM.

	2-directional Weighted-Avg.	3-directional Weighted-Avg.	4-directional Weighted-Avg.
Girl	30.3		
Boy	27.4	27.5	27.6
Test Pattern	22.5	23.2	21.5

Table 3.8 SNR (dB) of Weighted-Average 2DDM.

	Girl	Boy	Test Pattern
SNR (dB)	28.6	26.7	22.2

Table 3.9 SNR (dB) of Prediction 2DDM.

		Girl	Boy	Test Pattern
One dimensional DM	Linear mode	24.6	23.7	20.1
	Abate mode	27.1	23.2	19.2
	Song mode	28.0	24.8	19.4
	A-mode	29.6	25.5	21.8
	B-mode	30.0	26.5	22.6
Two dimensional DM	Normal mode	32.7	28.7	23.3
	Look-Ahead	33.4	29.4	23.0
	3-State (1.6b/p)	28.8	24.5	21.2
	Weighted Avg.	30.3	27.4	22.5
	Prediction 2DDM	28.6	26.7	22.2

Table 3.10 SNR (dB) of the delta modulators.

	N.Adv. Normal	N.Adv. Look-Ahead	3-State N.Adv. Normal	3-State N.Adv. Look-Ahead	Weighted-Average
Girl	28.1	29.5	25.8	26.1	26.9

Table 3.11 Signal-to-frame difference ratio (dB) of two dimensional DM.

N.Adv. Normal	Adv. Normal	N.Adv. Look-Ahead	Adv. Look-Ahead	3-State N.Adv. Normal	Weighted-Average
26.8	26.1	27.6	26.6	22.8	25.7

Table 3.12 SNR (dB) of two dimensional DM as a result of interlaced scanning. (Boy)

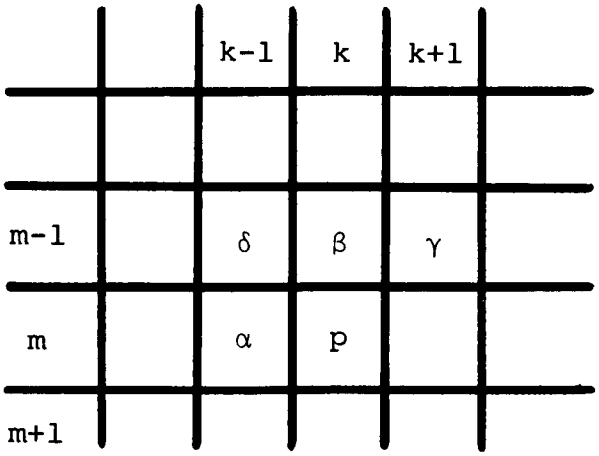


Fig. 3.1 Relative positions of pixels.

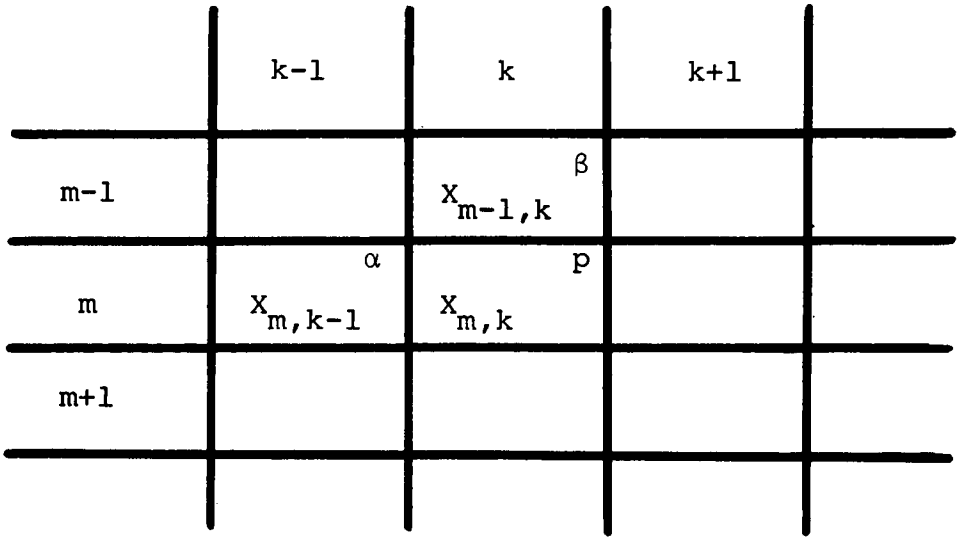


Fig. 3.2 Relative positions of present input sample and previous estimates.

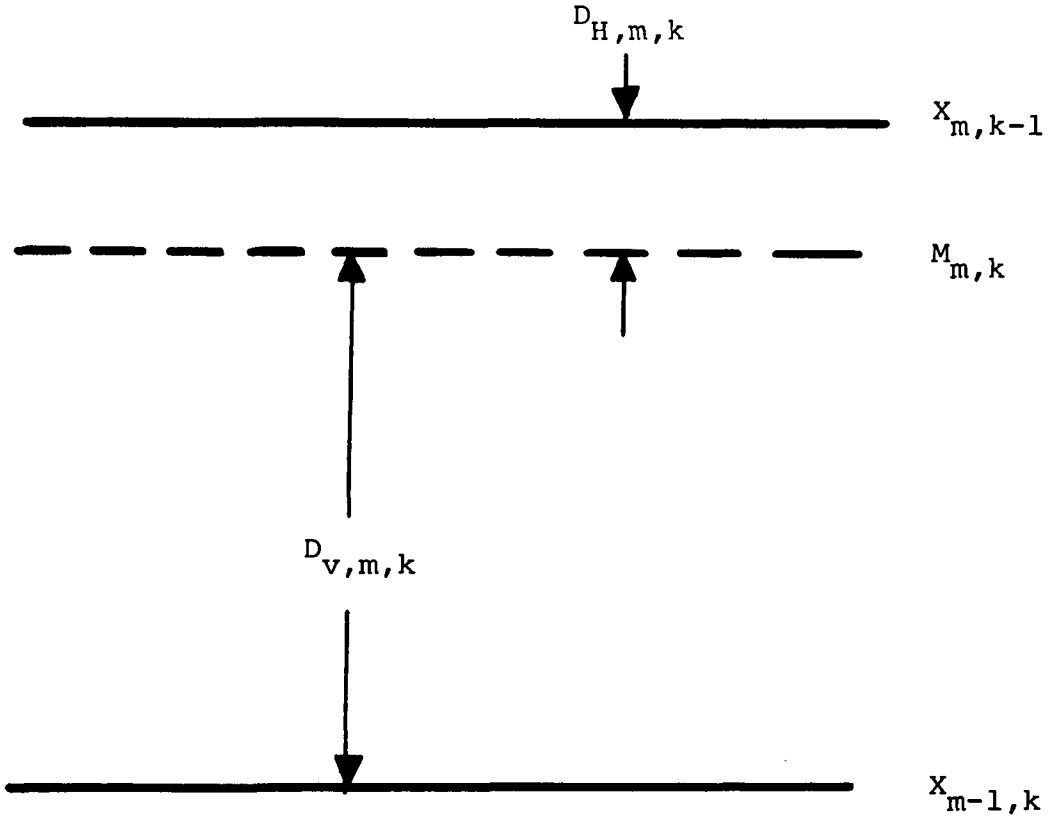


Fig. 3.3 The comparison scheme used in a Normal mode 2DDM.

Fig. 3.4 N.Adv. Normal 2DDM encoded "Girl".

Fig. 3.5 N.Adv. Normal 2DDM encoded "Girl" at 1 b/p.

Fig. 3.6 Difference picture of N.Adv. Normal 2DDM.

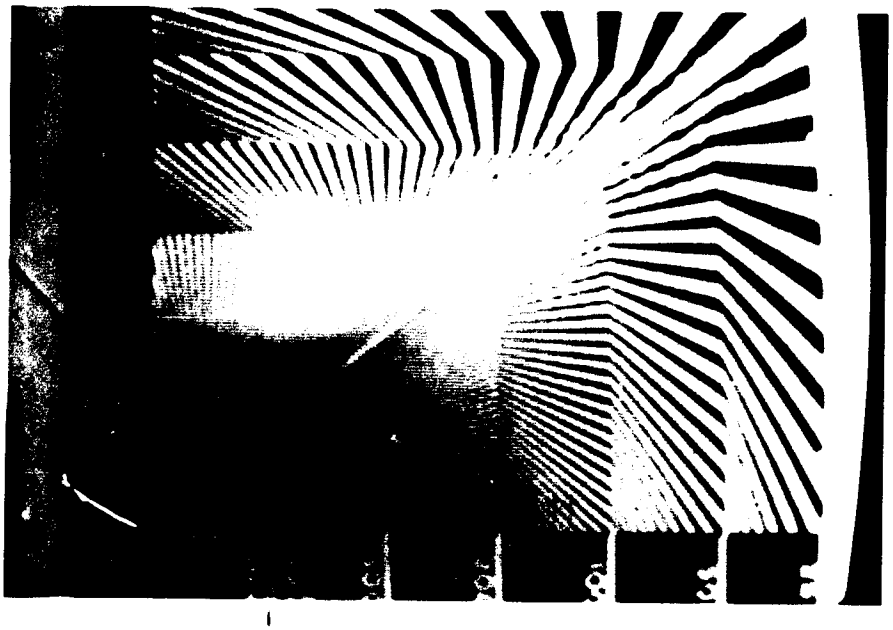
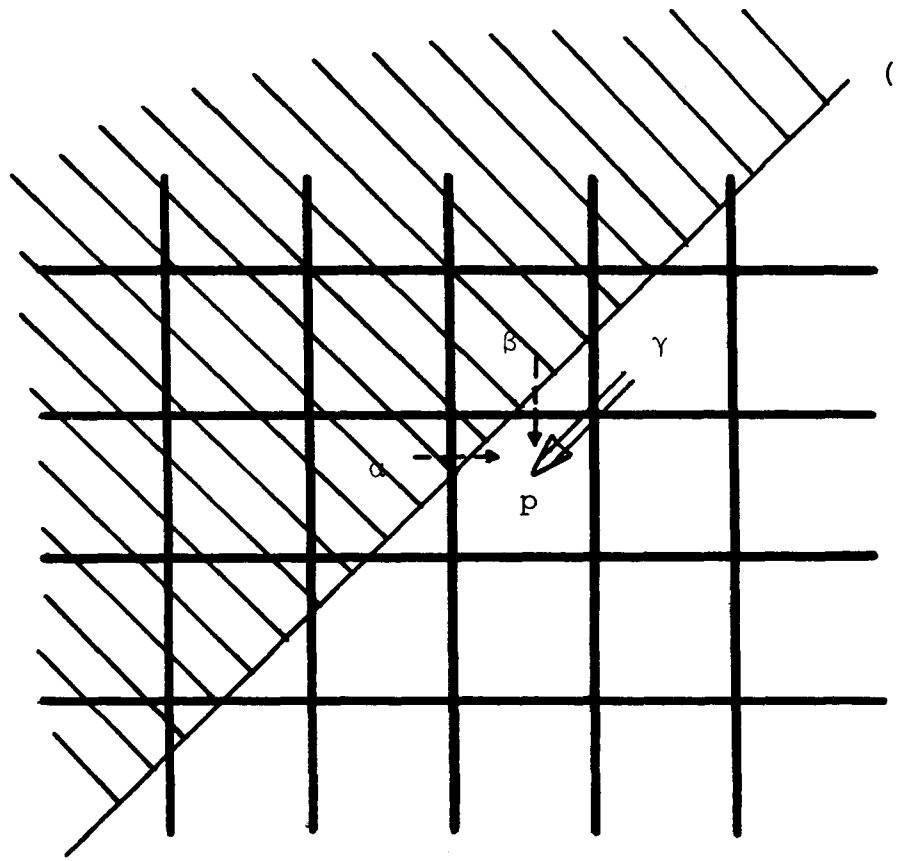


Fig. 3.7 N.Adv. Normal 2DDM encoded picture.



In a N.Adv. 2DDM, from either α or β to encode p it is a sudden jump of voltage, however, in an Adv. 2DDM, pixel p can be encoded accurately from pixel γ.

Fig. 3.8 N.Adv. 2DDM cannot accurately encode edges at an angle of 45 degrees.

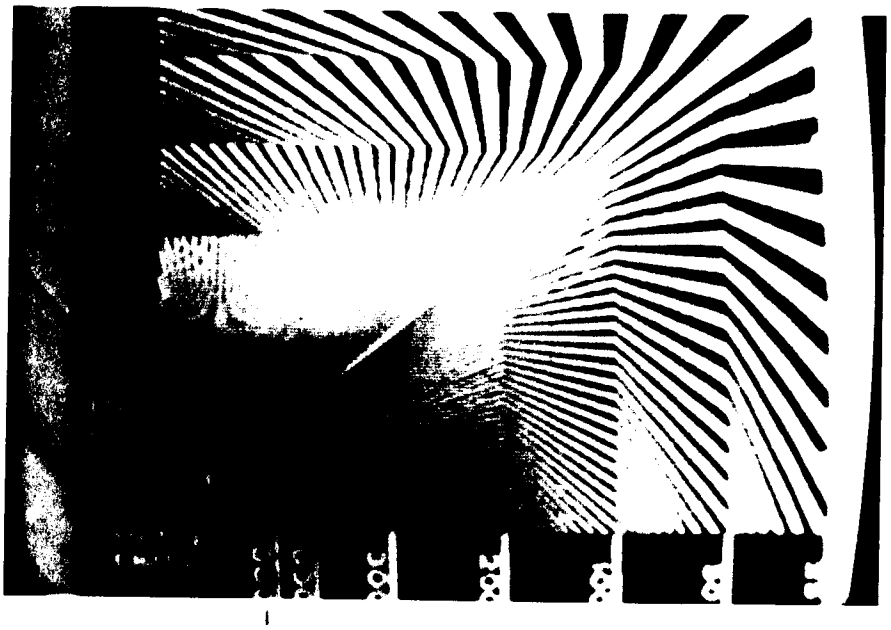


Fig. 3.9 Adv. Normal 2DDM encoded "Test Pattern".

Fig. 3.10 Adv. Normal 2DDM encoded "Girl".

Fig. 3.11 Adv. Normal 2DDM encoded "Girl" at 1 b/p.

Fig. 3.12 Adv. Normal 2DDM encoded "Boy".

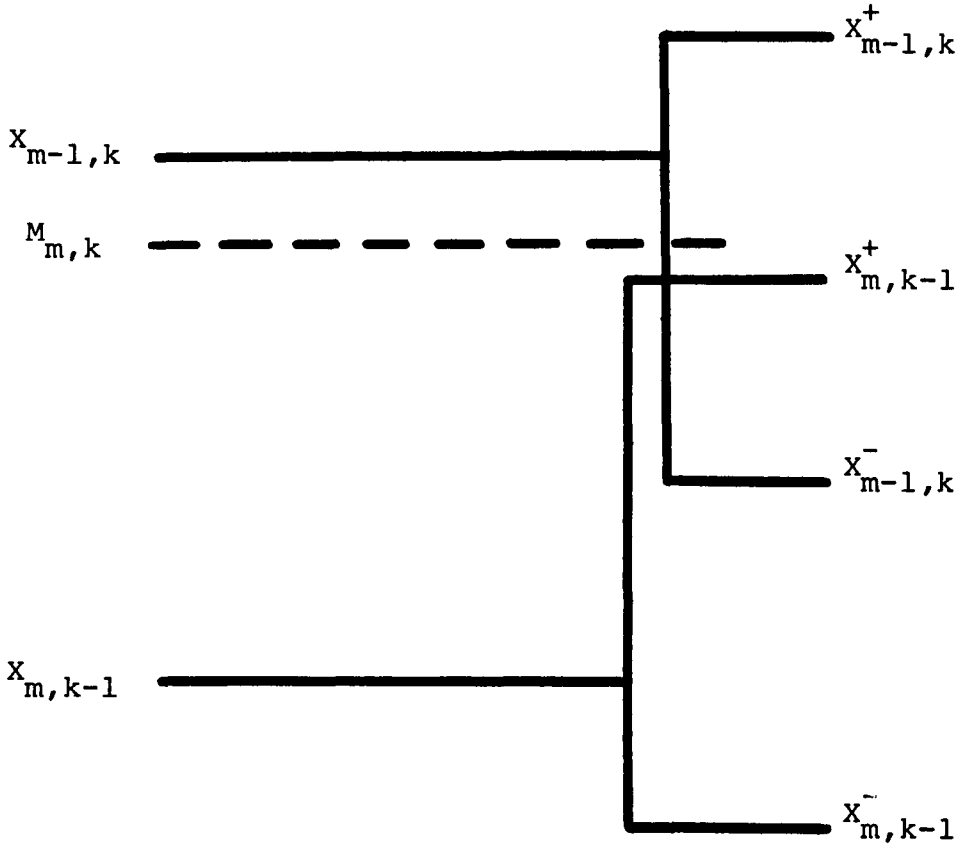


Fig. 3.13 The input sample compares with the four new estimates instead of the two previous estimates in a Look-Ahead 2DDM system.

Fig. 3.14 N.Adv. Look-Ahead 2DDM encoded "Girl".

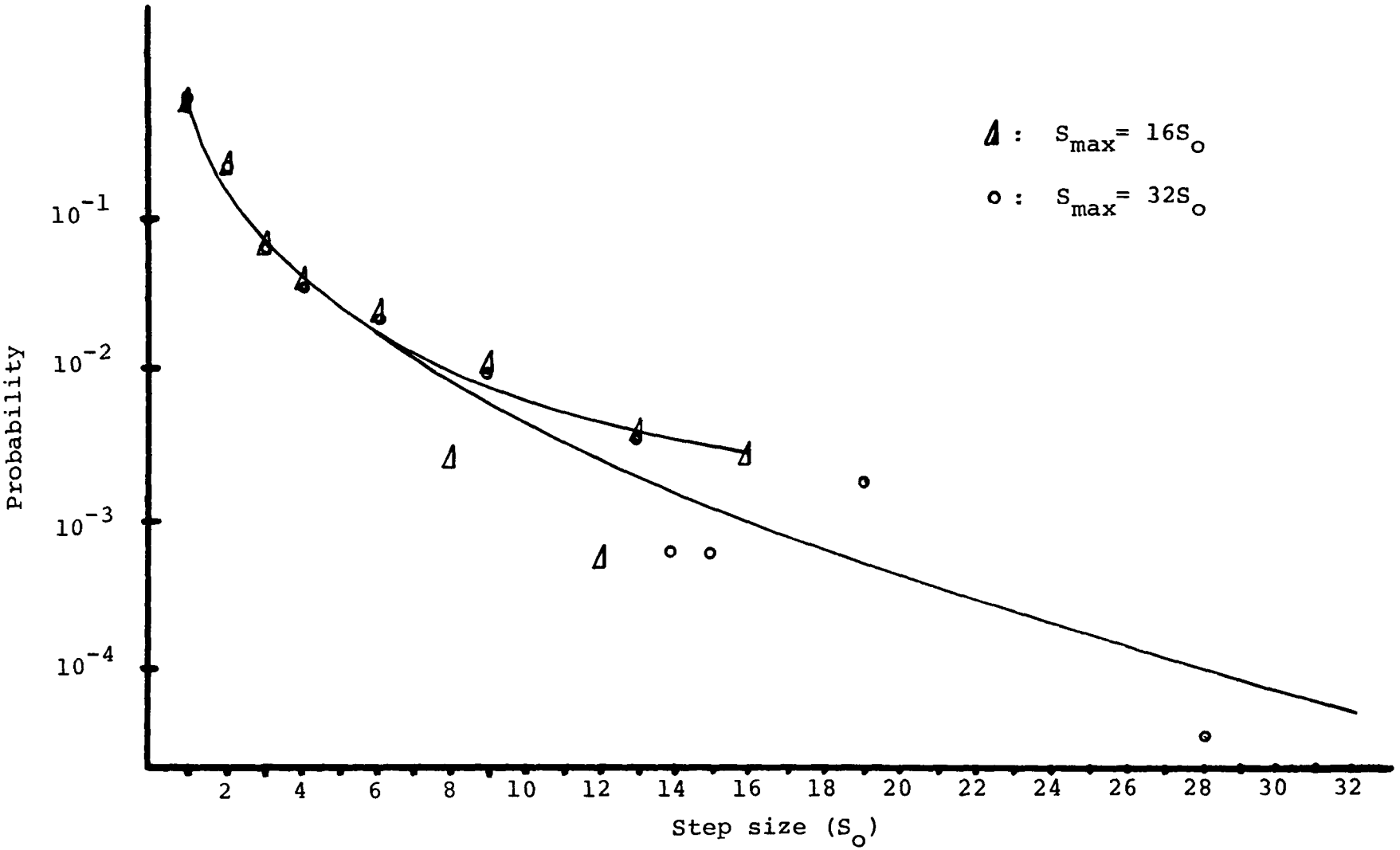


Fig. 3.15 Probability density function of the magnitude of the step size of Look-Ahead 2DDM encoding the "Boy".

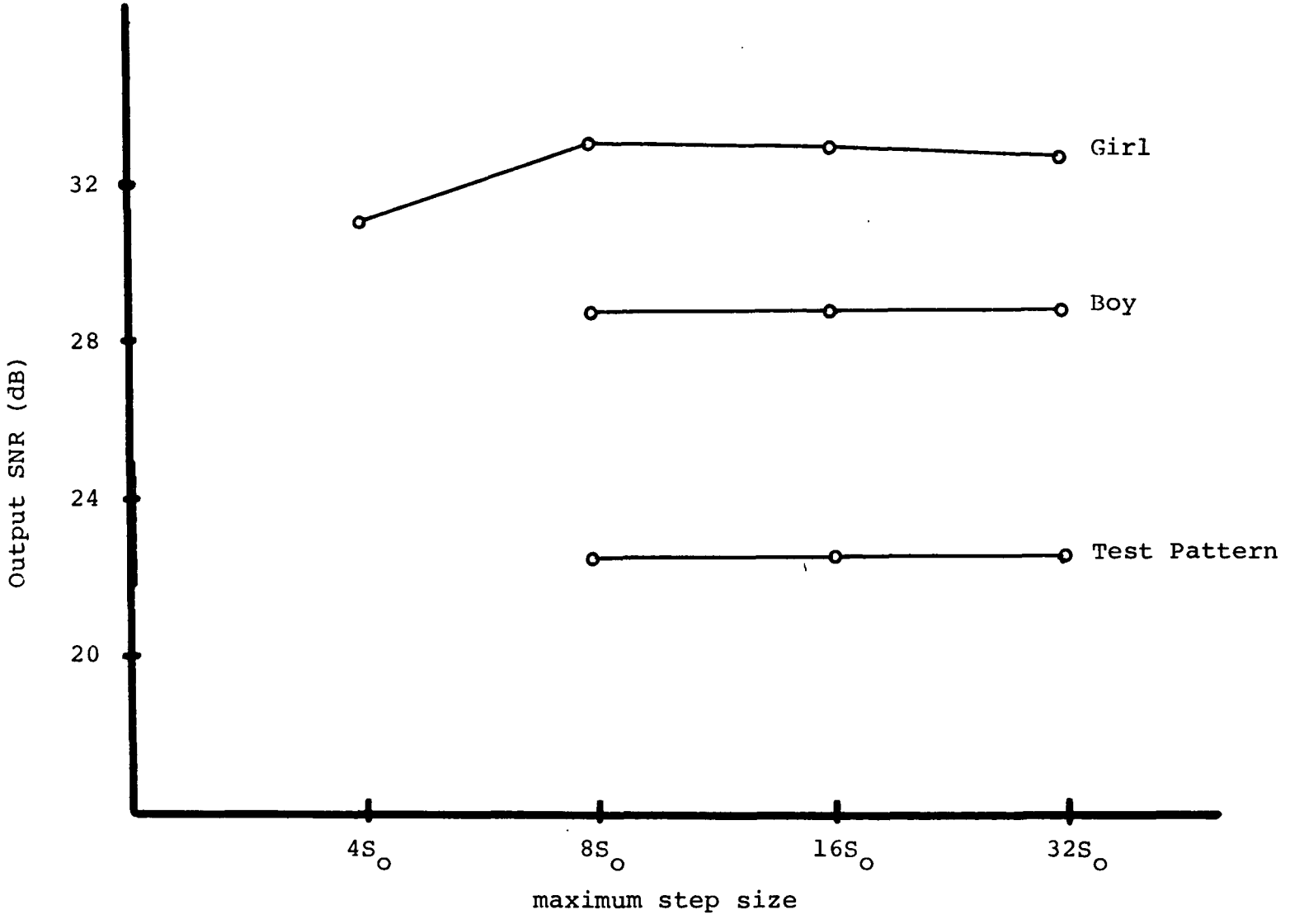


Fig. 3.16 The output SNR as a function of S_{max} in a N.Adv. Normal 2DDM.

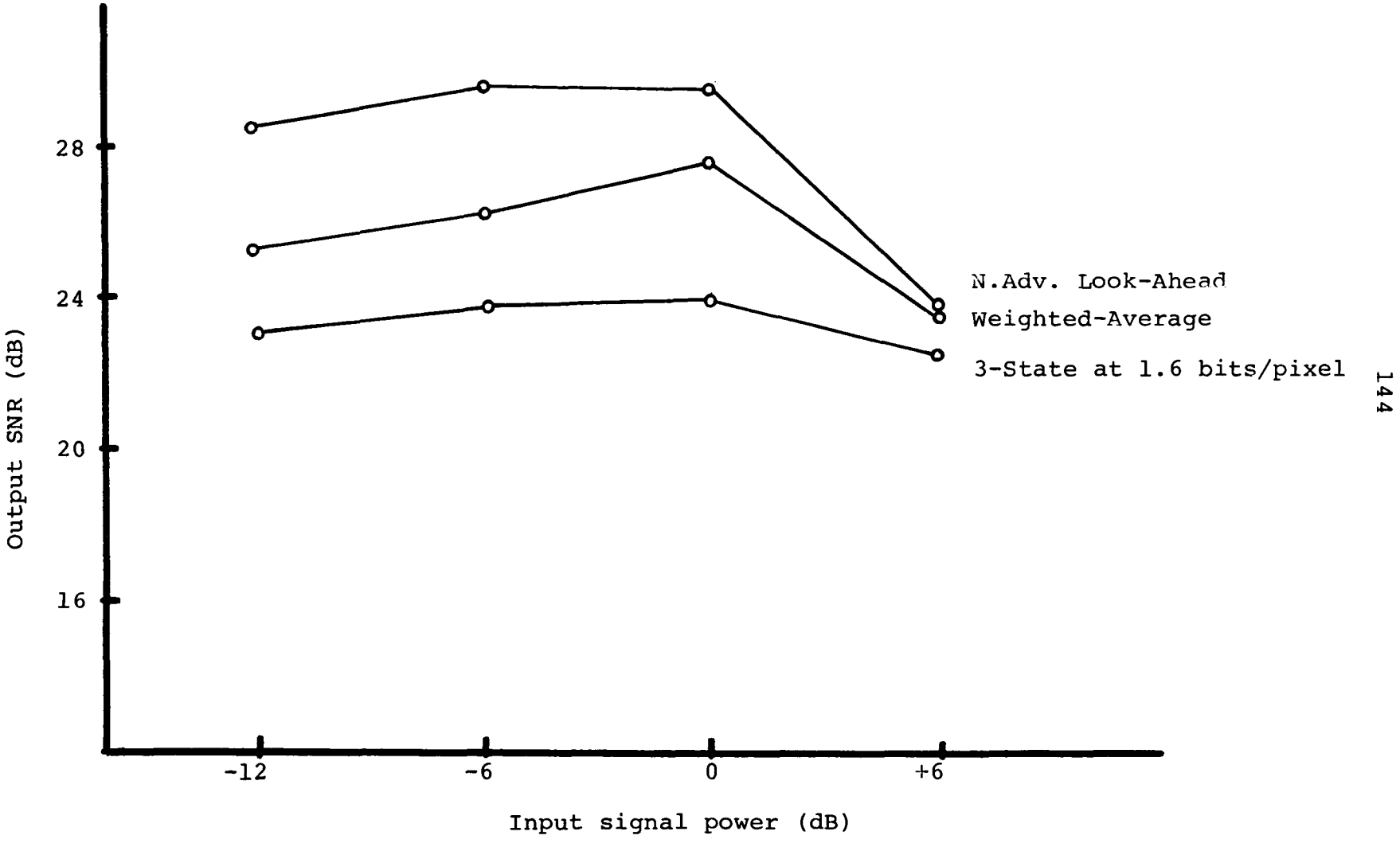


Fig. 3.17 Dynamic range of two dimensional DM.

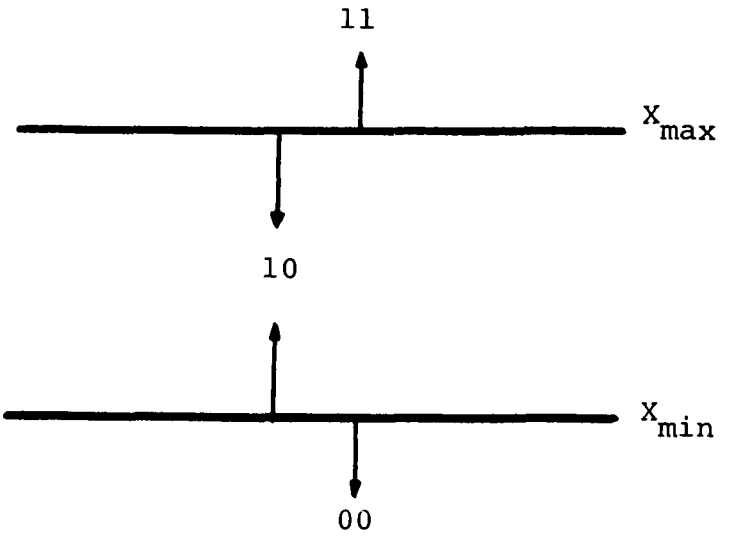


Fig. 3.18 There are only three output states in a 3-State 2DDM.

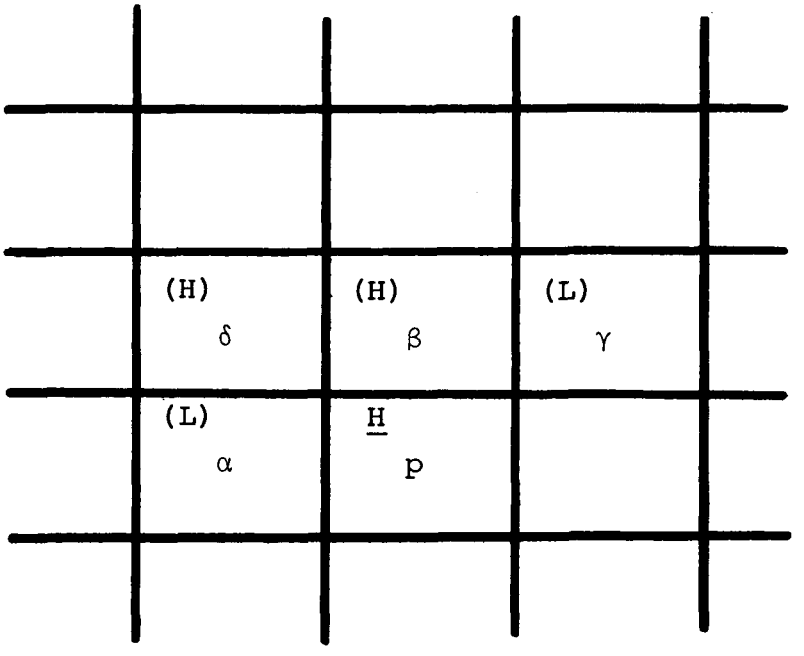


Fig. 3.19 A "history" bit is associated with each pixel in the 3-State 2DDM.

Fig. 3.20 3-State 2DDM encoded "Girl".

Fig. 3.21 Weighted-Average 2DDM encoded "Girl".

Fig. 3.22 Prediction 2DDM encoded picture.

Chapter IV

ENTROPY ENCODING OF DM OUTPUT BITS

4.1 Introduction

A delta modulator converts analog input waveforms into a string of binary digits. These binary digits can be transmitted directly or can be processed before transmission. This kind of processing prior to transmission is called an entropy encoding. Entropy encoding is a technique by which a string of digits are rearranged into a shorter sequence and it still carries the exact information of the original sequence. Entropy encoding can be applied to any finite discrete random variable. For instance, if we divide the DM output bit stream into blocks of n digits, then the digits in every block is a random variable, X. There are 2^n possible outcomes of the random variable X. If the probabilities of the occurrence of each of the 2^n combinations of digits are not the same, entropy encoding technique can be employed to redefine those digits into a more efficient decipherable variable length code thereby shortens the total transmitted digits and reduces the required bandwidth. If one particular sequence occurs in groups very often, such as the steady state pattern in a delta modulator, entropy encoding

techniques can again be applied to transmit the length of the groups of sequence instead of the original sequence. In this case, the total number of transmitted digits is also shortened. There are many different entropy encoding techniques. However, for the entropy encoding of DM output bits, we are interested in only two types of codings; the Huffman code [89] and the run length code [91-96]. According to information theory entropy/bit is a measure of effeciency of a digital transmission system. Entropy/bit is obtained by dividing the average code word length into the entropy of the source. For a binary digital source that consists of k code words, each with a probability of p_i, the entropy, H, of the source is defined as

$$H = -\sum_{i=1}^{k} p_i \log_2 p_i \quad . \tag{4.1}$$

The average code word length is defined as

$$\bar{n} = \sum_{i=1}^{k} n_i p_i \tag{4.2}$$

where n_i is the length of the i-th code word. The entropy /bit, h, is hence obtained as

$$h = \frac{H}{\bar{n}} \quad . \tag{4.3}$$

If the code words of a digital source is arranged in such a way that the channel capacity is fully utilized, h will

be equal to 1. Otherwise, h is always less than 1 and greater than 0. A high value of h indicates that the code words chosen for the information source match the statistical property of the source, hence the code used is an efficient one. If the value of h is low, it indicates that either other sets of code words ought to be used for a more efficient transmission or entropy encoding techniques should be employed to save channel bandwidth. However, different values of h may be obtained under different conditions for the same digital source. For instance, given a stream of DM output bits, depending on the block length used to calculate the entropy, h can have different values. h is a monotonically decreasing function of the block length. The absolute minimum value of h is obtained when the block length contains all the bits from the beginning of the transmission to the end of the transmission.

Huffman code is a coding scheme by which the bit stream is broken up into blocks of N bits length, each block contains one of the 2^N possible sequences. The least probable message is then assigned a code word containing the longest sequence of bits, whereas the more probable messages are assigned code words of shorter length. Huffman code is the optimum coding scheme for entropy encoding if the block length is very large. However, in practice, the block length N is often restricted because of the

large number of sequences, 2^N, involved in the statistical computation. Run length code is not an optimum coding scheme, but is a less complex one. For the same amount of hardware complexity, a good run length code may exceed the performance of the Huffman code. Because of that, run length code has been extensively researched to optimize the coding scheme in relation to the statistics of the source.

In this chapter, the statistics of the DM output bits are studied. The entropy of these bits are calculated along with the average code word length. Run length coding schemes which match the statistical properties of the DM output bits are also discussed.

4.2 Entropy Encoding of 1DDM Output Bit Stream

The entropy/bit of the one dimensional Song mode ADM output bits generated by using the low speed flying spot scanner sampled at 2 bits/pixel is 0.82, as shown in Table 4.1. This value is a result of the sequence statistics calculated as described by Eq. (4.1) and Eq. (4.3). The block length used is 10 bits. Longer block length will result in lower entropy value. However, it is beyond the capacity of our PDP-8 computer to process a block length longer than 11 bits. Therefore a block length of 10 bits was used in our studies of entropy encoding. This value,

0.82, indicates that only 82 percent of the channel capacity are utilized, and 18 percent of redundancy is contained in the DM bit stream. We can tabulate the probabilities of the 2^{10} = 1024 different sequences and use the Huffman's method to remove the 18 percent redundancy, so that a saving in bandwidth of up to 18 percent can be achieved. However, this is a extremely complex procedure because of the large number of sequences involved. Other indirect ways of applying Huffman code to a smaller set of sequences should be studied, they are discussed next.

Adaptive DM generates a sequence of consecutive 1's or 0's whenever a slope overload occurs. Different length of runs have different probabilities, thus, entropy can be measured from these runs to determine the amount of redundancy. Table 4.2 showns the probabilities of different length of runs of consecutive 1's, p(L). The probabilities of runs of consecutive 0's should be identical to p(L). Therefore the conditional probability of runs of 1's, p(L|1), is twice that of the p(L) as shown in Table 4.2. The entropy of these runs cannot be calculated directly using Eq. (4.1), because of the two different kinds of runs (runs of 1's and runs of 0's) involved.Meyr et al. have shown in their paper [96] that in a two-tone facsimile run length coding, if the length of a run depends only on the grey tone of the run, and if the grey tone is statistically dependent only on

the grey tone of the previous run, the entropy per bit is given by

$$h = \frac{H(L|0) + H(L|1)}{E(L|0) + E(L|1)} \tag{4.4}$$

where $H(L|0)$ and $H(L|1)$ are the entropies associated with run length distributions of 0's and 1's respectively, $E(L|0)$ and $E(L|1)$ are the average run length of 0's and 1's respectively.

Equation (4.4) is directly applicable to our case. $H(L|1)$ is defined as

$$H(L|1) = - \sum_{i=1}^{k} p_i(L|1) \log_2 p_i(L|1). \tag{4.5}$$

For example, placing the probabilities of the "Girl" given in Table 4.2 into Eq. (4.5), we get

$$H(L|0) = H(L|1) = 1.6607.$$

The average run length of 1's is defined as

$$E(L|1) = \sum_{i=1}^{k} n_i p_i(1|1). \tag{4.6}$$

Placing the probabilities of the "Girl" into Eq. (4.6), we have

$$E(L|0) = E(L|1) = 1.7685.$$

From Eq. (4.4) the entropy/bit is thus,

$$h = \frac{2(1.6607)}{2(1.7685)} = 0.94.$$

This indicates that the code chosen (consecutive 1's or consecutive 0's) to represent slope overload condition matches the probability of the interval during which the DM is slope overloaded. In other words, when the DM is slope overloaded, the Song mode ADM output bits utilize the channel capacity very efficiently to convey this information. Although entropy encoding can still be applied to remove the remaining 6 percent redundancy, it is hardly worth the effort.

The Song mode ADM generates a sequence of consecutive "10" pattern in the steady state. The length of the steady state pattern with the associated probabilities are listed in Table 4.3. We know that the DM output bits stream has at least 18 percent redundancy (Table 4.1). We also know that the DM encoding is very efficient when it is slope overloaded, therefore, the low entropy in the DM output bits must be resulted from the inefficient DM encoding of the steady state sequences. We can calculate the entropy of the steady state sequences from Table 4.3. Notice that in this table, 0 run length indicates that a "1" is followed by another "1" or "0" is followed by another "0", thus no run of "10" sequence can be counted. From this table the entropy of the "10" runs of the "Girl" is computed to be $H = 1.667$, and the average run length to be $\overline{n} = 2.059$, thus

$$h = \frac{H}{\bar{n}} = 0.81.$$

We can conclude that when DM is considered as a "10" sequence (steady state) generator, the channel capacity utilization is 81 percent. The Huffman code can easily be implemented to these steady state sequences, because there are only 16 different sequences involved. The average run length of the sequences of the "Girl" in Table 4.3 after being coded by the Huffman's method can be shown to be 1.812. Compare it to the original average run length, $\bar{n} = 2.059$, a saving of

$$\frac{2.059 - 1.812}{2.059} = 12\ \%$$

in bandwidth can be achieved. The theoretical limit of

$$\frac{2.059 - 1.677}{2.059} = 19\ \%$$

cannot be achieved. The best approximation is 12%, because the best known source code at present is the Huffman code.

It is interesting to point out that the entropy/bit calculated directly from the 10-bit block is 0.82, whereas the value calculated from the steady state sequence is 0.81. There are two reasons for this discrepancy. First, different equipments were used in the simulations; the 10-bit block statistics were obtained by using a flying spot scanner as the signal source, whereas the run length statistics

were obtained from the TV camera and scan converters. Therefore it is meaningless to compare these two values in great detail. Second, the entropy resulted from the direct measurement of the 10-bit block cannot actually represent the entropy of the signal source. It is only an approximation, because the block length has to be at least one video frame in length to obtain the actual entropy of the source. However, in general, the entropy measured from direct calculation of sequence blocks should be lower than the entropy obtained by any run length statistics.

4.3 Entropy Encoding of 2DDM Output Bit Stream

Two dimensional DM generates two output bits at each sample. The first bit is the directional bit and the second bit is the DM sign bit. These two bits are statistically independent. Although Huffman code can apply directly to blocks of 2DDM output bits, it will be much easier to implement if these bits are seperated into two groups, one group consists of 2DDM directional bits and the other consists of DM error sign bits. Run length statistics of the directional bits are tabulated in Table 4.4. Because the distance between scanning lines in a non-interlaced TV system is shorter than the distance

between the horizontally adjacent pixels, the vertical directional run length is, on the average, longer than the horizontal directional run length. The entropy and the average run length of the horizontal directional run can be calculated by using Eq. (4.5) and Eq. (4.6) respectively,

$$H_o(L|0) = 1.641$$

and $$E_o(L|0) = 1.692.$$

Similarly, for vertical directional run, $H_1(L|1)$ and $E_1(L|1)$ can also be calculated,

$$H_1(L|1) = 3.232$$

$$E_1(L|1) = 4.021.$$

Placing these values into Eq. (4.4), the entropy/bit of the 2DDM directional bits is therefore,

$$h = 0.853.$$

This is the entropy/bit of the directional bits measured from the run length statistics when they are separated from the DM sign bits and grouped together. The entropy/bit of the complete 2DDM output bit stream calculated by using the slow scan flying spot scanner system is 0.93 as shown in Table 4.1. From this number we can predict that the entropy/bit of the 2DDM error sign bits is very high, hence the overall percentage reduction in bandwidth by using entropy encoding is very small. The statistics of the DM sign bits cannot be changed. However, we can modify the run length

statistics of the directional bits by producing longer runs to lower the entropy and, consequently, to reduce the channel bandwidth.

In order to obtain longer directional runs in a 2DDM system, the 2DDM algorithm is modified. At every sample, the previous horizontal estimate and the previous vertical estimate are compared, if the difference between them is less than a threshold value, the 2DDM will encode from the horizontal direction regardless of the value of the input sample. This is called a biased 2DDM. If the threshold value,T, is reasonably small, the degradation of the quality of the picture caused by this biasing scheme is negligible. The resulting output SNR of the "Girl" and the "BOY", with $T=4S_o$, are 30.1 dB and 28.0 dB respectively as compared to 32.7 dB and 28.7 dB of the unbiased 2DDM. The resulting run length statistics of the horizontal directional bits are shown in Table 4.5. Long horizontal runs are now generated. It is too long to tabulate all the probabilities, hence the average probabilities are shown for groups of longer runs. Because the DM encoding is biased to the horizontal direction, the vertical directional runs are drastically cut short, as shown in Table 4.6. More than 75 percent of the vertical directional bits are of a single run. This happened because only one vertical encoding is actually needed at the edge to make a discrete jump of DM estimate, the following

pixels can then be encoded horizontally. The entropy/bit of the horizontal directional bits calculated from Table 4.5 is,

$$h_o = \frac{H_o(L|0)}{E_o(L|0)} = \frac{4.48}{10.53} = 0.425.$$

This low value of h indicates that a large amount of redundancy exists in the biased 2DDM system. Entropy encoding can result in a reduction of appreciable amount of tranmission bandwidth.

Direct Huffman encoding of the horizontal directional runs listed in Table 4.5 is one method of entropy encoding. The other simpler technique is the run length encoding. Because the probability distribution of large runs are more prevalent than an exponential distribution model, the B_1 code described by Meyer [96] is the ideal run length code for these horizontal directional bits. The B_1 code is shown in Table 4.7, in which "D" indicates encoding directions. If a "0" represents horizontal encoding, then "D" will all be replaced by "0" in the codewords. Since the vertical runs are short, run length encoding will not gain anything, so it is left un-coded. The resulting average length of codeword for horizontal runs is 4.5 bits. The average run length of the vertical runs is 1.37. Therefore, the ratio of the average codeword length of the directional bits after run length encoding to that before encoding is

$$\frac{4.5 + 1.37}{10.53 + 1.37} = 0.49.$$

This ratio indicates that, on the average, one out of every two directional bits can be eliminated by using entropy encoding techniques. If we take the DM sign bits into account, the ratio of the resulting number of transimitted bits to the original transmitted bits can easily be calculated to be

0.75.

This is the entropy/bit of the biased 2DDM with run length encoding on the directional bits. It is interesting to compare this result to that of the 3-State 2DDM, in which the redundancies are introduced into the DM output bits directly from the 3-State algorithm. The entropy/bit of the 3-State 2DDM is measured to be 0.76 as shown in Table 4.1. This value is very close to the theoretical predicted value, $(\log_2 3)/2 = 0.8$. The results are very interesting, the run length encoded biased 2DDM and the 3-State 2DDM both showed the same value of entropy. Although they are two different DM systems, they both obtained a lower entropy by degrading the DM performance slightly. We can raise the threshold value of the biased 2DDM to 8 S_o to obtain a lower entropy, but the degradation of the picture quality is then too great to be ignored. We can conclude that within a acceptable limit of degradation, a simple entropy encoder can reduce the channel bandwith of a 2DDM by

$$\frac{1 - 0.75}{1} = 25\%$$

In conclusion, the entropy encoding can reduce the bandwidth of the one dimensional Song mode ADM by 12% (19% theoretical) without any degradation, whereas the bandwidth reduction of the 2DDM is 25% with a slight amount of degradation, in which the SNR is still higher than that of the 1DDM. This is another evidence that the two dimensional DM encoding is superior to the one dimensional DM encoding. With the same channel bit rate, a biased 2DDM has higher SNR and less edge busyness than the 1DDM, and still be able to reduce 25% of the bandwidth if properly run length encoded. However, in practice, entropy encoder cannot be properly operated if the information bits are not protected by error correction channel code. The combination of entropy encoder and channel encoder will make the DM system so complicated which violates the principal merit of a DM system of being a simple and effective digital encoding system. The entropy encoding study in this chapter is a very interesting and valuable work. However, any further pursue in this subject will lead to a level beyond practical interest.

One dimensional Song mode ADM	N.ADV. Normal mode 2DDM	3-state 2DDM
0.82	0.93	0.76

Table 4.1 Entropy/bit of the DM output bits. Results are obtained by using the low speed flying spot scanner with a resolution of 171 by 171 pixels. The entropies are taken from blocks of 10 bits with the DM sampled at 2 bits/pixel. Signal source is a slide with geometric shapes.

Run length	Girl		Boy	
	P(L)	P(L|1)	P(L)	P(L|1)
1	0.2357	0.4714	0.2297	0.4595
2	0.1874	0.3948	0.1817	0.3634
3	0.0421	0.0842	0.0478	0.0955
4	0.0115	0.0230	0.0160	0.0320
5	0.0064	0.0128	0.0091	0.0182
6	0.0030	0.0059	0.0062	0.0123
7	0.0016	0.0032	0.0036	0.0073
8	0.0011	0.0022	0.0020	0.0041
9	0.0007	0.0015	0.0016	0.0032
10	0.0005	0.0010	0.0014	0.0028
11	0	0	0.0008	0.0016
12	0	0	0.0001	0.0002

Table 4.2 Probabilities and conditional probabilities of the run length of consecutive 1's (or 0's) of the output bit stream of the one dimensional Song mode ADM sampled at 2 bits/pixel.

Run length	Girl	Boy
0	0.4328	0.4598
1	0.4332	0.4261
2	0.0783	0.0729
3	0.0291	0.0252
4	0.0132	0.0091
5	0.0062	0.0039
6	0.0033	0.0018
7	0.0016	0.00076
8	0.0009	0.0003
9	0.0005	0.00017
10	0.0004	0.00008
11	0.0002	0.00002
12	0.00006	0.00001
13	0.00006	0.00001
14	0.00006	0
15	0.00003	0

Table 4.3 Probabilities of the run length of consecutive runs of "10" pattern of the output bit stream of the one dimensional Song mode ADM.

Run length	Probabilities (GIRL) Consecutive 0's (horizontal encoding) P(L|0)	Consecutive 1's (vertical encoding) P(L|1)
1	0.6240	0.2716
2	0.2068	0.1658
3	0.0938	0.1426
4	0.0428	0.1026
5	0.01634	0.0756
6	0.00766	0.0342
7	0.00392	0.0426
8	0.00194	0.0298
9	0.00128	0.0252
10	0.00062	0.0190
11	0.00026	0.0144
12	0.00026	0.0116
13	0.00006	0.0092
14	0.00016	0.0074
15	0	0.0060
16	0.00014	0.0042
17	0.00006	0.0038
18	0.00006	0.0026
19	0	0.0022
20	0	0.002
21	0	0.00146
22	0	0.00118
23	0	0.001
24	0	0.00058
25	0	0.0006
26	0	0.00056
27	0	0.00042
28	0	0.0003
29	0	0.00026
30	0	0.00026
31	0	0.00022
32	0	0.00020

Table 4.4 Run length statistics of the directional bits in a Normal mode 2DDM. A directional bit "0" indicates horizontal encoding and a "1" indicates vertical encoding.

Run length	P(L|H) (GIRL)
1	0.2759
2	0.1461
3	0.0989
4	0.0625
5	0.0503
6	0.0394
7	0.0310
8	0.0251
9	0.0189
10	0.0162
11	0.0146
12	0.0115
13	0.0106
14	0.0089
15	0.0074
16 - 21	0.00567
22 - 28	0.00355
29 - 35	0.00284
36 - 42	0.00199
43 - 54	0.00170
55 - 83	0.00113
84 - 114	0.00057
115 - 189	0.00028
190 - 225	0.00006

Table 4.5 The run length statistics of the horizontal directional bits in a biased 2DDM with $T = 4\ S_o$.

Run length	P(L|V) (GIRL)
1	0.755
2	0.154
3	0.061
4	0.022
5	0.0056
6	0.0013
7	0.0003
8	0.0002
9	0

Table 4.6 The run length statistics of the vertical directional bits in a biased 2DDM with $T = 4\ S_o$.

Run length	Codeword
1	D 0
2	D 1
3	D 0 D 0
4	D 0 D 1
5	D 1 D 0
6	D 1 D 1
7	D 0 D 0 D 0
8	D 0 D 0 D 1
.	.
.	.
.	.
15	D 0 D 0 D 0 D 0
16	D 0 D 0 D 0 D 1
.	.
.	.
31	D 0 D 0 D 0 D 0 D 0
.	.
.	.

Table 4.7 The B_1 run length code.

Chapter V

THE EFFECTS OF CHANNEL ERRORS

5.1 Introduction

In a digital channel, the binary digits might be decoded incorrectly by the decoder due to the corruption by various kinds of noise in the transmission system. These incorrectly decoded digits are called channel errors or transmission errors. The channel errors will cause a deviation between the received and the transmitted waveforms. In a video system, the channel errors will result in a erroneous brightness level [99-104].In a PCM transmission system, these erroneous brightness levels are isolated to a single pixel. Therefore, the total channel error noise power is low, and the channel errors are less likely to be detected by human observers. However, in a delta modulation system, because of the interdependencies between the picture elements, the erroneous brightness levels caused by channel errors will propagate into many pixels. The total channel error noise power is, therefore, high, and the channel errors are easily detectable. The reason is the well known theory of the trade-off between the channel bandwidth compression and the channel error immunity. The savings in

channel bandwidth of a delta modulator from utilizing the pixel correlations to remove redundancy is paid for in the form of higher channel error noise power in the decoded waveform [24]. Signal redundancy protects the signal from being corrupted by channel errors. Delta modulators remove much of the signal redundancy in order to save channel bandwidth, hence, the DM system is vulnerable in a noisy transmission channel.

The response of the delta modulator decoder to a transmission error is a permanent DC level shift. As a result of the DC shift, erroneous brightness patterns appear in the decoded picture. In this chapter, the effects of the channel errors on the one dimensional delta modulators and the two dimensional delta modulators are discussed in separate sections.

5.2 The Effects Of Channel Errors On One Dimensional Delta Modulators

In an one dimensional DM, the DC level shift in the decoded waveform caused by channel errors will propagate from line to line throughout the entire picture. However, if the encoder and the decoder are both reset at the end of a horizontal sweep, the error will propagate only until the end of the scanning line. Leaky integrator can be employed to both the encoder and the decoder to shorten the error

stripes appeared in the decoded pictures [99].A leak factor of 1/32 to 1/64 is sufficient to significantly reduce the length of the error streaks while maintaining the quality of the encoded pictures. The insertion of a leaky integrator into a one dimensional DM system will degrade the SNR. However, with a leak factor of 1/32 or less, the degradation is very small as shown in Table 5.1. Figure 5.1 is a DM decoded picture at a channel error rate of 10^{-4} without using leaky integrators. The dark and bright stripes are clearly visible in the picture. The DM decoded pictures with leaky integrators can be seen from the literatures by Scheinberg [66,99]. All of the one dimensional delta modulators described in Chapter 2 can use the leaky integrator to reduce the effects of channel errors. The results are equally good for all one dimensional DM algorithms. The real time demonstration of the Song mode 1DDM with channel errors at the Communications laboratory of the City College of New York showed that with a channel error rate of 10^{-4} and a leak factor of 1/64, the decoded channel errors are barely noticeable; they are clearly visible only if the channel error rate is 10^{-3} and higher. Because of the nature of the video scanning system, one dimensional delta modulators restrict the propagation of channel errors to a single scanning line. This, together with the leaky integrator and horizontal sweep reset cause the 1DDM to have a much better channel error immunity than the 2DDM as we shall see in the next section.

5.3 The Effects Of Channel Errors On Two Dimensional Delta Modulators

In a two dimensional delta modulator, because of the nature of the encoding path, the channel error pattern can spread out in both the horizontal and the vertical directions. In a 2DDM decoder, each channel error, depending on the pixel position of the initiating error in a particular coding path, causes a different kind of shape in the error pattern. Refer to a particular coding path as shown in Fig. 5.2, if an error occurs at pixel x, the error will not propagate because pixel x is isolated and will not be used for the estimation of other pixels. Thus, the error at pixel x is like a PCM error, and a brightness deviation will be observed only at that point. Because of the small pixel size and that there is a good possibility that the deviation is small, this kind of error is not clearly visible. It can be readily shown that the probability of a pixel in a coding path being isolated is 1/4. Refer to Fig. 5.3(a), if pixel x is in error, in order not to spread the error, pixels β and γ have to be encoded from pixels δ and ε respectively. Each pixel has an equal probability of being encoded horizontally or vertically. Therefore, the probability that pixels β and γ are not encoded from pixel x is

$$1/2 \times 1/2 = 1/4.$$

This implies that one out of every four errors in the

decoded picture will not propagate and therefore, will not be noticed.

If an error occurs at pixel y in Fig. 5.2, the error will propagate along the coding path, and a horizontal streak of 5 pixels in length will appear on the decoded picture. The probability that an error will cause a horizontal or a vertical streak covering n pixels is $(1/2)^{2n}$. This is shown in Fig. 5.3(b) where the channel error started at pixel y and propagates horizontally to cover n pixels until pixel ν. In order to produce this kind of error pattern, pixels β to ν have to be encoded horizontally. The pixels in the next scanning line, pixels 1 to k have also to be encoded horizontally to avoid the vertical spreading of the error, and pixel λ has to be encoded vertically to terminate the error propagation. There is a total of 2n pixels involved after pixel y. Therefore the probability is $(1/2)^{2n}$. In the example of pixel y in Fig. 5.2, n is equal to 5. The probability that the error will move only horizontally for 5 pixels is $(1/2)^{+10} = 10^{-3}$ which is very small. Hence, generally three out of four errors received will spread into a two dimensional area of various sizes and shapes. This kind of error spreading can be seen in Fig.5.2; if pixel z is in error, the entire area covered by the coding path following pixel z will be affected. The convergence and divergence property of the coding path after the erroneous pixel determines whether

the error pattern is constrained or not. The worst case happens when there are multiple errors in a divergent coding path. The result is often the total destruction of a part of the picture covered by the coding path following several erroneous pixels. Figure 5.4(a) is the 2DDM decoded picture with a probability of error of 10^{-4}. The visible error patterns propagate diagonally resulted in a very noisy picture. The total channel error noise power of this picture is more than that of the 1DDM encoded picture in Fig. 5.1. This, again, can be explained by the same statement of signal redundancy and error immunity. The 2DDM decoder results in a more effective channel error noise power due to the fact that the 2DDM encoder has removed more signal redundancy at the transmitting end. The 2DDM is a more efficient delta modulator in the sense that less bandwidth is required by the 2DDM to transmit video signals with the same SNR. On the other hand, the 2DDM is more vulnerable to channel errors due to the more efficient removal of signal redundancies. Figure 5.4(a) is sampled at 1 bit/pixel. There are 25×10^4 pixels in the picture. At an error rate of 10^{-4}, 25 errors should appear in Fig. 5.4(a), Some of these errors are isolated, some of these errors do not propagate long enough to meet the next error, hence, die down shortly after the pixel that originated it. But some of the errors happened to be located in a divergent

coding path, thereby causing multiple errors to be linked together resulting in a total destruction of a portion of the decoded picture.

However, channel errors occur randomly. All the registers of the delta modulator encoder and decoder are reset to zero prior to the beginning of each new frame. The errors in each frame appear at different places, and pictures seen on the TV screen should not be as severely distorted by the error patterns as the pictures taken from a single frame.

The deviation between the transmitted picture and the decoded picture due to channel errors in a 1DDM system can be reduced by employing a leaky integrator into a delta modulator. The same technique, however does not work in a 2DDM system. Figure 5.4(b) is the same picture as Fig. 5.4(a) with a leak factor of 1/32. The error pattern can be seen decaying as is expected. However, in Fig. 5.4(b) fewer errors are isolated and more errors are propagated. It is believed that the decaying effect of the leaky integrator disturbs the steady state of the delta modulator, hence the step size at each pixel is larger, and switching between coding directions is more often, channel errors, therefore, cause larger deviations and have a higher tendency to propagate. Another reason for the leaky integrator's ineffectiveness in reducing the channel errors in a 2DDM system is the two

dimensional coding characteristics of the 2DDM. The two dimensional coding path causes the channel error to spread into the neighboring lines more quickly causing it to meet the next channel error and to drift further away from the correct DM estimate before the leaky integrator can effectively reduce the noise amplitude. The leaky integrator will be effective only if the coding path stays in a horizontal or vertical direction for a sufficient length of time to accumulate enough decaying effect of the leaky integrator for correcting errors. The biased 2DDM has this kind of property. The biased 2DDM is biased to encode from one direction, it changes encoding direction only if necessary, such as along steep slopes. This biased 2DDM scheme not only provides long directional run length for entropy encoding as described in Chapter 4, it also restricts the error propagation in a one dimensional fashion. Figure 5.5(a) is the horizontally biased 2DDM decoded picture at a channel error rate of 10^{-4} and with a leak factor of 1/32. This picture happened to have a very bad quality. It is not due to the effect of biasing, but due to the inaccurate adjustment of the scan convertor, the monitor and the camera. However, the purpose of this picture is to see how the error propagates. We should not be too concerned about the picture quality at this moment. The errors propagate horizontally as is expected and leave

a large part of the picture unaffected. Figure 5.5(b) is the vertically biased 2DDM decoded picture at a channel error rate of 10^{-4} and a leak factor of 1/32. The errors can be seen to propagate vertically only. The biased 2DDM gives an enormous improvement over the unbiased 2DDM in handling channel errors. A diagonal flash at the upper right hand corner of Fig. 5.5(b) is, again, a photographic effect and is not due to the propagation of channel errors.

The two dimensional delta modulators described in Chapter 3 can be divided into two categories. One category sends extra information concerning the encoding direction to the receiver, such as the Normal mode 2DDM, the Look-Ahead 2DDM, the 3-State 2DDM, and the biased 2DDM. The other category uses the previous DM estimates to control and to predict the DM encodings, such as the Weighted-Average 2DDM and the Prediction 2DDM. The first category can utilize the biased scheme and a leaky integrator to reduce the effect of channel errors as described earlier. However, the second category has an entirely different channel error response. In a Prediction 2DDM, no error will be isolated, every channel error will affect other pixels and will propagate diagonally. However, because of the averaging effect of this scheme, a slight modification of the weighting coefficients, making the sum of the coefficients to be less than 1, will have the leaky effect which greatly

reduces the length of error propagation. This is a well established and standard technique in DPCM picture encodings [2,3,4,6]. By employing this technique, error propagations will decay very fast, as a result, an acceptable video image will be produced. In a Weighted-Average 2DDM, however, the results are astonishing. Not a single channel error will be tolerated by this Weighted-Average 2DDM as shown in Fig. 5.6(a). The reason can be stated as follows.

The Weighted-Average scheme is a very sensitive edge detector. The equation of this scheme as described in Chapter 3 Eq.(3.25),

$$\overline{P} = 0.75\overline{\alpha} + 0.75\overline{\beta} - 0.5\overline{\delta}$$

is actually a comparision of DM estimates among pixels α, β, and δ as shown in Fig. 5.7(a). If $\overline{\alpha}$ is closer to $\overline{\delta}$ and $\overline{\beta}$ is farther away from $\overline{\alpha}$ and $\overline{\delta}$, the chosen encoding direction will be vertical. It is because in this scheme a discontinuity is assumed to be at the vertical line k, and the DM decides to encode vertically. This is a very good scheme in a noiseless transmission medium. The encoded pictures and SNR shown in Chapter 3 have demonstrated that this Weighted-Average 2DDM is a good alternative in two dimensional DM encoding. However, in the presence of channel noise, any decoded error will confuse the DM decoder. As a result, false edges are detected and channel errors are

propagated into the 4-th quadrant at an angle of 90 degrees. For instance, in a flat area, each pixel has a DM estimate of 5 units as shown in Fig. 5.7(b), an error occured at pixel P causing the decoder to incorrectly decode the DM estimate to be 10 units. As a result, at pixel P', the DM decoder will assume that there is a horizontal edge, and subsequently decode the signal horizontally from the erroneous pixel P. At pixel P", the decoder will assume that there is a vertical edge and decode the signal vertically from pixel P. The decoded DM estimates of pixels P' and P" are highly unlikely to have a value of 5 units as that of the rest of the pixels. Therefore, the horizontal edge and the vertical edge will continue to be erroneously detected at the subsequent pixels P_2' and P_2''. These false detection of edges will continue till the end of the scanning line and the end of a video frame, at that point, the registers are reset. This kind of edge detecting characteristic of the Weighted-Average 2DDM produced the unusual channel error patterns shown in Fig. 2.6(a). The leaky integrator will not ease this problem but only makes it worse as shown in Fig. 2.6(b). Therefore, from the transmission point of view, the Weighted-Average 2DDM is not a good algorithm and has little value in practical applications.

5.4 Conclusions

Decoded channel errors in a DM system will propagate; depending on the algorithm employed, different error patterns will result. Channel errors in a one dimensional DM generate error streaks along the scanning lines. Leaky integrator can be employed in the one dimensional DM system to correct the errors. The real time demonstration of the one dimensional Song mode ADM system showed that the system is highly acceptable if the channel error rate is 10^{-4} or lower. In a two dimensional DM system, because of the nature of the coding path, channel errors spread into a two dimensional domain. Channel errors in a Prediction DM system will decay because of the averaging effect. If the prediction coefficients are modified to have a sum of less than 1, the channel errors will decay even faster. Therefore, from the transmission point of view, it is a good two dimensional DM algorithm. In a Normal mode 2DDM, a Look-Ahead 2DDM, or a 3-State 2DDM not every channel error will spread; only 3 out of 4 channel errors will cause an error spreading. To reduce the effect of channel error, this kind of DM system should be biased to one direction to restrict the error propagation to be within a few lines before the leaky integrator can correct the error. Satisfactory results can be obtained if the channel error rate is 10^{-5} or lower.

In a Weighted-Average 2DDM system, however, the picture will be destroyed even there is only one single decoded channel error because of the sensitive edge detecting characteristic of this scheme. Decoded channel error causes the DM decoder to sense a false edge and to switch the coding direction to the erroneous pixel. The higher the amplitude of the error the more likely the DM decoder will decode from that erroneous pixel. Therefore, unless accompanied by some other good error correction techniques the Weighted-Average 2DDM cannot tolerate any channel error and cannot be used in actual transmission.

ADM mode	Song		A-mode		B-mode	
Leak factor	0	1/32	0	1/32	0	1/32
SNR (dB)	24.8	23.6	25.5	23.4	26.5	23.0

Table 5.1 Comparison of the one dimension ADM systems with and without leaky integrators.

Fig. 5.1 Song mode 1DDM decoded picture at a channel error rate of 10^{-4}, no leaky integrator is used but line-reset is employed.

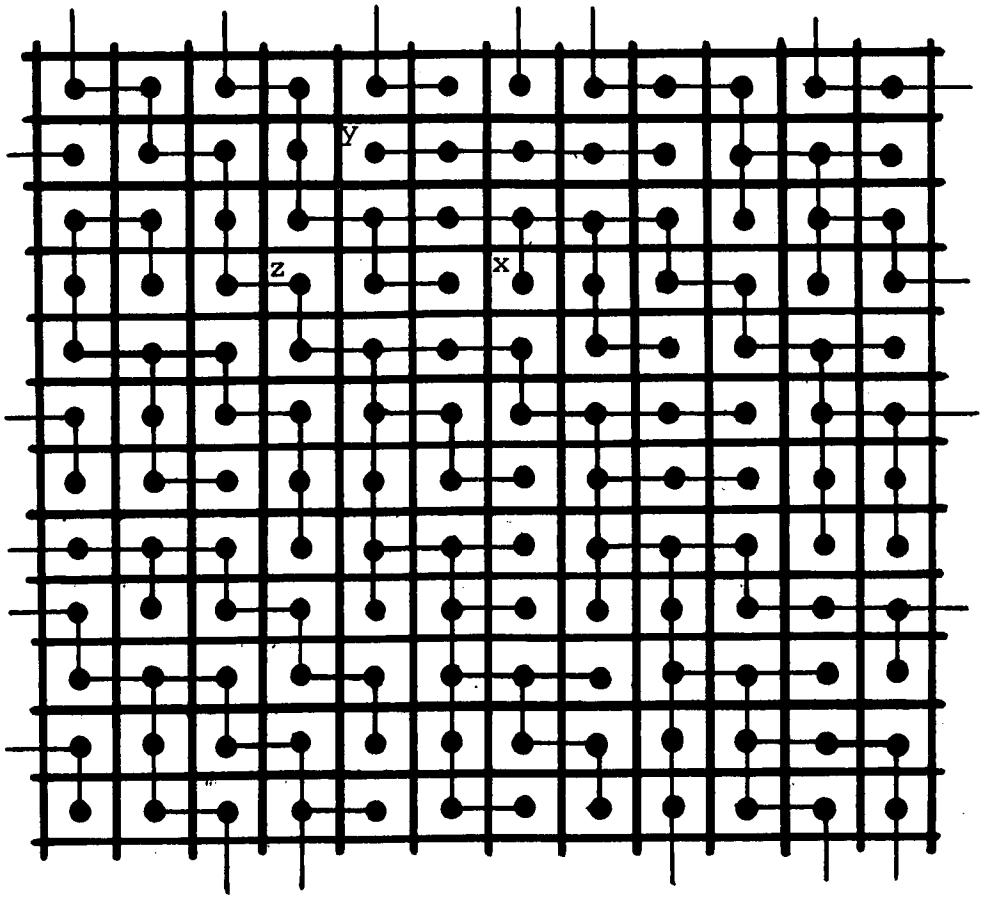


Fig. 5.2 portion of a coding path.

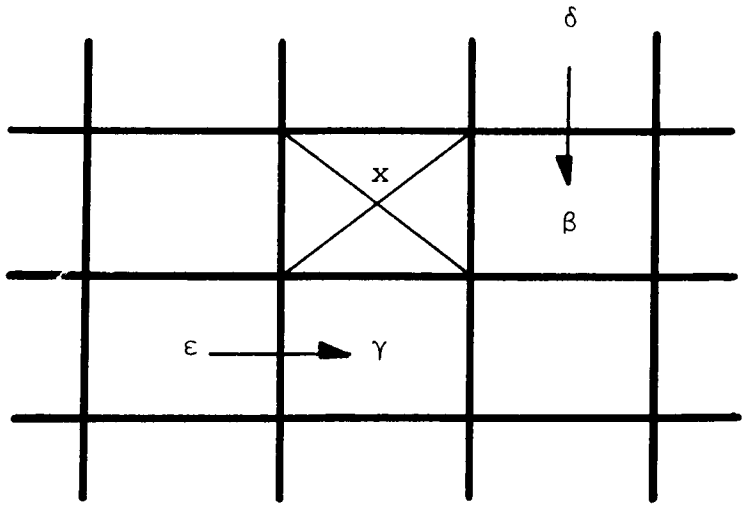


Fig. 5.3(a) If pixel x is in error, the probability of isolating pixel x is 1/4.

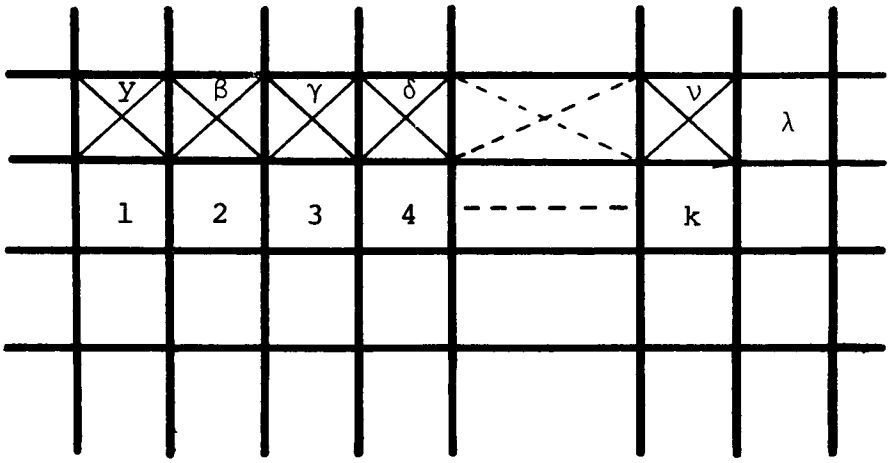


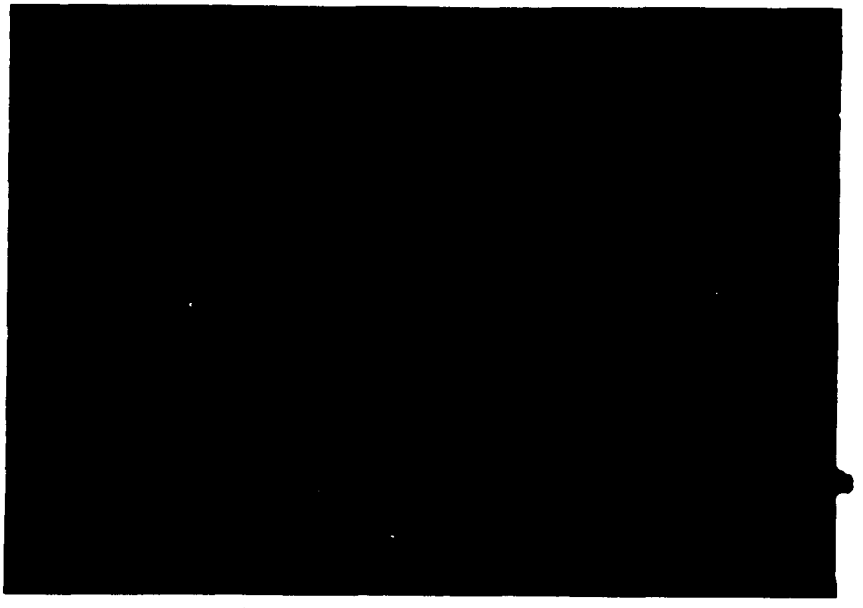
Fig. 5.3(b) The probability of error propagating horizontally covering n pixels is $(\frac{1}{2})^{2n}$.

(a)

(b)

Fig. 5.4 Normal 2DDM decoded pictures at a channel error rate of 10^{-4}, sampling at 1 bit per pixel, (a) no leaky integrator is used, (b) with a leak factor of 1/32.

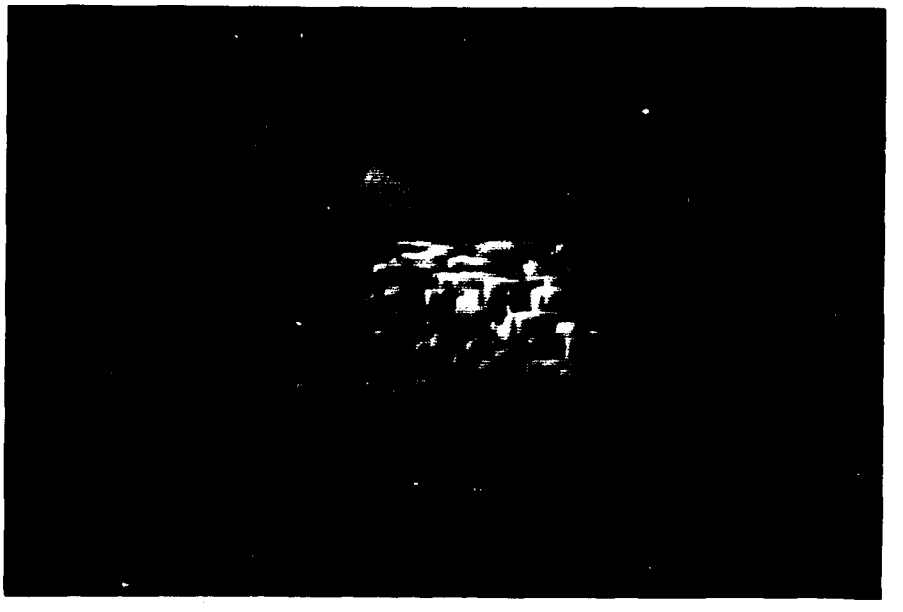
(a)



(b)

Fig. 5.5 Biased Normal 2DDM decoded pictures at a channel error rate of 10^{-4}, sampling at 1 bit/pixel, (a) horizontally biased, $T=4S_o$, (b) vertically biased, $T=4S_o$.

(a)



(b)

Fig. 5.6 Weighted-Average 2DDM decoded pictures at a channel error rate of 10^{-4}, sampling at 1 bit/pixel, (a) no leaky integrator is employed, (b) with a leak factor of 1/32.

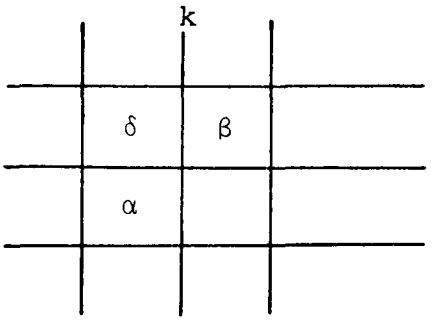


Fig. 5.7(a) Weighted-Average 2DDM detects an edge by comparing the DM estimates of pixels α, β and δ.

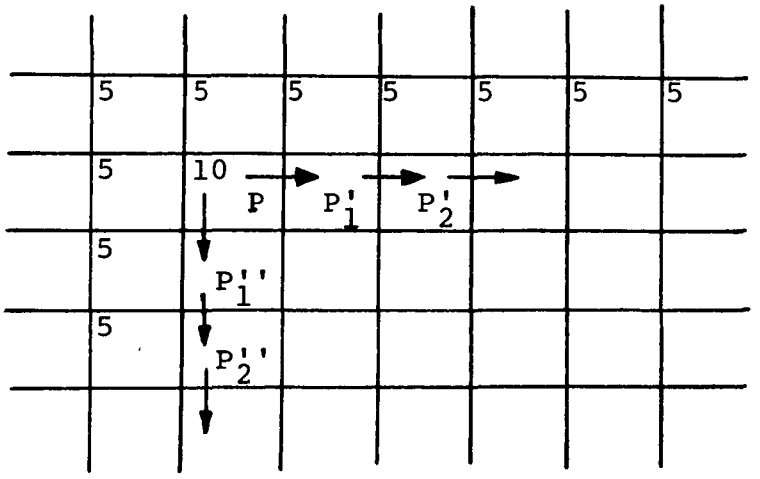


Fig. 5.7(b) Every channel error in Weighted-Average 2DDM will cause a false dectection of a horizontal edge and a vertical edge.

Chapter VI

CONCLUSIONS

Delta modulator, since its invention in 1946, has gone through several stages of modification. The double integration DM, the high information DM, and the continuous DM, etc. have improved the dynamic range of a delta modulator considerably. The real improvement, however, is the adaptation of DM with digital circuits. Digital delta modulators have many advantages, such as higher accuracy, free of constant adjustment, better control of parameters, etc.. Adaptive step size algorithms, such as those described by Abate, Jayant, Song and others, together with the digital techniques have made the delta modulator a more practical and economical communications system. In this dissertation, several new types of delta modulator algorithms are discussed, because of them the delta modulation techniques in real time video transmission will be further advanced.

In Chapter 2, the delta modulation using the Song mode algorithm has been discussed. The inherent edge busyness of an ADM can be reduced by using the look-ahead scheme. The edge busyness can further be reduced if the DM response is modified to minimize the slope overload noise. For this purpose, A-mode ADM algorithm and B-mode ADM

algorithm are proposed and simulated. The results are shorter length in edge busyness and a general improvement in output SNR. However, in order to totally eliminate edge busyness, the DM should be able to encode along the edges. Based on this understanding, two dimensional delta modulators are studied. The redistribution of the quantization noise in a 2DDM results in a good quality picture which is free of edge busyness. Several two dimensional DM algorithms are studied. They all show an improvement over the one dimensional DM algorithms. Channel error response of the 1DDM and the 2DDM are also studied. The results are discussed in Chapter 5. Based on the present available data, we can state that among all these DM algorithms the Normal mode 2DDM is the best. It is the DM algorithm that should be used in the future video transmissions. However, the research work is not completed; further research must be continued to investigate the following areas.

- The simulated ADM using the A-mode and the B-mode algorithms has shown good results. But these step size adaptation algorithms are too complicated for a simple and economical DM system. If the A-mode and the B-mode ADM are to be implemented, a good circuitry ought to be designed to ease the complication in step size generator of the DM.

- The real time Normal mode and Look-Ahead 2DDM encoders have been built at the Communications laboratory of the City College of New York. The frame-to-frame discrepancy, when the camera looks into a still picture, varies at a slow rate. The human eyes can detect this kind of slow variations. The phenomenon of flickering and speckling were described by observers. Further studies should be carried out to obtain a better understanding of this phenomenon and to find ways to reduce it.
- Weighted-Average 2DDM algorithm can generate good quality images. It is also easy to implement. However the practical applications is restricted because of its poor channel error response. Further studies on good error correction techniques are needed.
- Prediction 2DDM uses the prediction techniques, which has been used in DPCM successfully, to generate DM estimates. The results shown in Chapter 3 suggest that this is a good DM algorithm. Further studies should be continued in this direction.
- Entropy encoding can remove the redundancy in the DM output bits in order to save bandwidth. However, a DM system becomes more vulnerable to channel errors when more redundancies are removed. Channel coding techniques should be considered to protect the information bits in the case of actual transmission.

- Color television system is in great demand at present. Delta modulation of colored video signals is the immediate follow-up research.
- The simulation equipments used in this dissertation consists of analog devices. These analog devices have degraded the performance of the overall system. We suggest in the future research,slide projector and TV camera should be replaced by magnetic tapes to store the digitized sampled pictures so that the signal source will be identical for repeated experiments. The analog scan converters should also be replaced by digital storage devices to reduce the distortion and degradations.

Appendix I

The DM Response Of The Abate-mode ADM

Starting with an initial DM estimate X_o, and an initial step size S_i, the DM response of the Abate-mode ADM to a step input is derived as follows:

$$X_1 = X_o + S_i$$

$$X_2 = X_1 + S_{i+1} = (X_o + S_i) + S_{i+1}$$

$$= (X_o + S_i) + (S_i + S_o)$$

$$= X_o + 2S_i + S_o$$

where S_o is the minimum step size. Continuing the iterative step, we get

$$X_3 = X_2 + S_{i+2}$$

$$= (X_o + 2S_i + S_o) + (S_i + 2S_o)$$

$$= X_o + 3S_i + 3S_o$$

$$X_4 = X_3 + S_{i+3}$$

$$= (X_o + 3S_i + 3S_o) + (S_i + 3S_o)$$

$$= X_o + 4S_i + 6S_o$$

...

$$X_k = X_o + kS_i + S_o \sum_{j=0}^{k-1} j$$

$$X_k = X_o + kS_i + S_o \frac{k(k-1)}{2}$$

$$= X_o + \frac{S_o}{2} k^2 + (S_i - \frac{S_o}{2})k \ . \qquad \text{(A.1.1)}$$

Eq. (A.1.1) is the general form of a parabolic curve. To obtain the vertex and the focus, we add $\frac{S_o}{2}\left(\frac{2S_i - S_o}{2S_o}\right)^2$ to both sides of (A.1.1), and get

$$X_k + \frac{S_o}{2}\left(\frac{2S_i - S_o}{2S_o}\right)^2 = X_o + \frac{S_o}{2}\left(k + \frac{2S_i - S_o}{2S_o}\right)^2 \ . \qquad \text{(A.1.2)}$$

Equation (A.1.2) can also be written as

$$4 \cdot \frac{1}{2S_o}\left[X_k - X_o + \frac{S_o}{2}\left(\frac{2S_i - S_o}{2S_o}\right)^2\right] = \left(k + \frac{2S_i - S_o}{2S_o}\right)^2 \ . \qquad \text{(A.1.3)}$$

From Eq. (A.1.3) it is clear that this parabola has a vertex at

$$\left(-\frac{2S_i - S_o}{2S_o}\ ,\ X_o - \frac{S_o}{2}\left(\frac{2S_i - S_o}{2S_o}\right)^2\right)$$

and a focus at

$$\left(-\frac{2S_i - S_o}{2S_o}\ ,\ X_o - \frac{S_o}{2}\left(\frac{2S_i - S_o}{2S_o}\right)^2 + \frac{1}{2S_o}\right).$$

For a numerical example, let $S_o = 1$ and $S_i = 4$; the vertex of the parabola is then at $(-3.5,\ X_o - 6.13)$, and the focus at $(-3.5,\ X_o - 5.63)$. This is shown in Fig. A.1. The solid line of the parabola is the DM response of the Abate mode ADM.

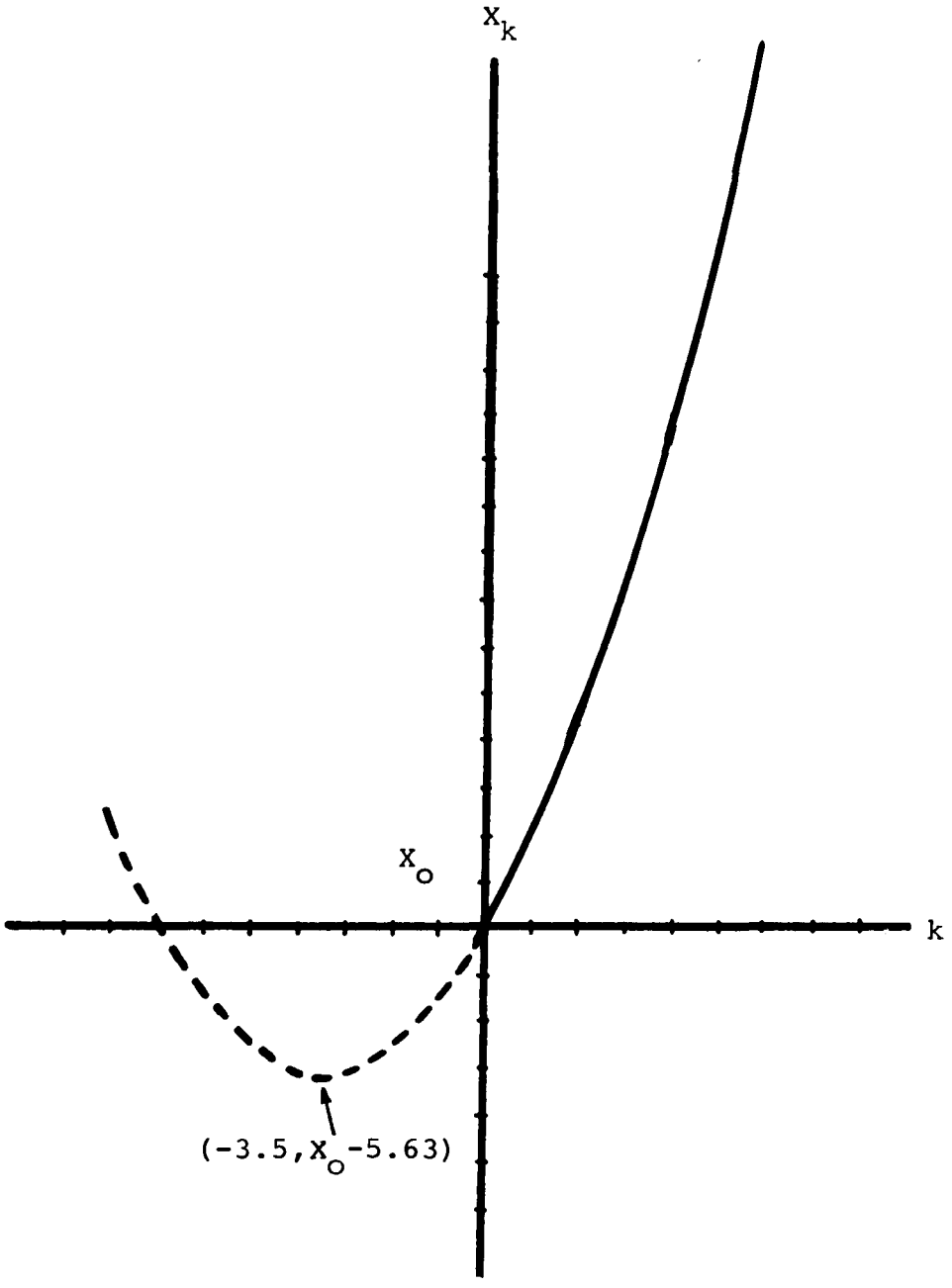


Fig. A.1 The Abate mode ADM response to a step input is a parabolic curve. In this diagram, S_o=1, S_i=4.

Appendix II

The Exponential Characteristics Of The Song Mode ADM

Starting with an initial DM estimate X_o, and an initial step size S_i, the DM response of the Song mode ADM to a step input is derived as follows:

$$X_1 = X_o + S_i$$

$$X_2 = X_1 + S_{i+1}$$

$$= X_o + S_i + S_{i+1}$$

$$= X_o + S_i + 1.5S_i$$

$$X_3 = X_2 + S_{i+2}$$

$$= X_1 + S_{i+1} + S_{i+2}$$

$$= (X_o + S_i + 1.5S_i) + (1.5)^2 S_i$$

$$\cdots$$

$$X_k = X_o + S_i[1 + 1.5 + (1.5)^2 + \ldots + (1.5)^{k-1}]$$

$$= X_o + 2[(1.5)^k - 1]S_i \ . \qquad \text{(A.2.1)}$$

$(1.5)^k$ can be written as

$$(1.5)^k = e^{k\ell n(1.5)} . \qquad \text{(A.2.2)}$$

Hence, from Eqs. (A.2.1) and (A.2.2) we obtain

$$X_k = X_o + 2S_i e^{k\ell n(1.5)} - 2S_i$$

$$= X_o + 2S_i e^{0.405k} - 2S_i \ . \qquad \text{(A.2.3)}$$

Equation (A.2.3) indicates that the Song mode ADM's response to a step input is an exponentially rising function with a time constant 0.405k and an intial slope $0.91S_i$.

Appendix III

Comparison Of The A-Mode ADM Algorithm And The CVSD Algorithm

Although the A-mode ADM algorithm and the CVSD algorithm are different in many aspects, e.g. the steady state response, the settling characteristic etc., they have the same rising characteristics in response to a step input.

When responding to a step input, the step size equation of the CVSD algorithm is

$$S_k = [\alpha|S_{k-1}| + S_o]e_k \qquad (A.3.1)$$

where α is a constant very close to 1, and S_o is an arbitrary constant. From Eq. (2.8a) we have, for the A-mode ADM,

$$S_k = [(1 - \frac{1}{c})|S_{k-1}| + \frac{S_{max}}{c}]\ e_k. \qquad (A.3.2)$$

From Eqs. (A.3.1) and (A.3.2) we can see that they are identical if

$$\alpha = 1 - \frac{1}{c} \qquad (A.3.3)$$

and

$$S_o = \frac{S_{max}}{c}. \qquad (A.3.4)$$

c is a constant of large value. Hence $(1 - \frac{1}{c})$ has a value close to 1. Eq. (A.3.3) is valid because α is also a constant close to 1. α and $(1 - \frac{1}{c})$ are of the same order

of magnitude. From Eq. (2.11), we can see that the largest increment in step size is generated when $S_i = 0$ and $k = 1$, therefore,

$$\Delta S_{k_{max}} = \frac{S_{max}}{c} . \qquad (A.3.5)$$

Comparing Eq. (A.3.5) with Eq. (A.3.4) we can see that in order to be able to conclude that the CVSD algorithm is similar to the A-mode ADM algorithm, S_o in Eq. (A.3.1) has to be the maximum increment in step size in the CVSD algorithm. This is true, because in response to a step input, the step size in the CVSD algorithm is increased by the amount S_o at every sample and then, subtracted by a "leak factor" that was resulted from multiplying the previous step size by a constant which is less than 1. The increment in step size is equal to or less than S_o, thus, the largest value of increment is S_o.

The A-mode ADM algorithm and the CVSD algorithm started from different points of view. They were analyzed differently. But their final effects are the same. They both have the same final goal, to reduce the amount of overshoot thereby reduces the amount of edge busyness.

REFERENCES

[1] B. M. Oliver, J. R. Pierce and C. E. Shannon, "The philosophy of PCM," Proc. IRE, Vol. 36, pp. 1324-1331, Nov. 1948.

[2] T. S. Huang, "PCM picture transmission," IEEE Spectrum, Vol. 2, pp. 57-63, Dec. 1965.

[3] "Waveform quantization and coding," Edited by N. S. Jayant, IEEE Press, 1976.

[4] "Redundancy reduction," Proc. IEEE (Special Issue), Vol. 55, Mar. 1967.

[5] L. C. Wilkins and P. A. Wintz, "Bibliography on data compression, picture properties, and picture coding," IEEE Trans. Inform. Theory, Vol. IT-17, pp. 180-197, Mar. 1971.

[6] E. R. Kretzmer, "Statistics of television signals," Bell Syst. Tech. J., Vol. 31, pp. 751-763, 1952.

[7] L. E. Franks, "A model for the random video process," Bell Syst. Tech. J., Vol. 45, pp. 609-630, April 1966.

[8] W. K. Pratt, J. Kane and H. C. Andrews, "Hadamard transform image coding," Proc. IEEE, Vol. 57, pp. 58-68, Jan. 1969.

[9] G. B. Anderson and T. S. Huang, "Piecewise Fourier transformation for picture bandwidth compression," IEEE Trans. Commun. Technol., Vol. COM-19, pp. 133-140, Apr. 1971.

[10] P. A. Wintz, "Transform picture coding," Proc. IEEE, Vol. 60, pp. 809-820, July 1972.

[11] M. Tasto and P. A. Wintz, "Image coding by adaptive block quantization," IEEE Trans. Commun. Technol. Vol. COM-19, pp. 957-971, Dec. 1971.

[12] A. Habibi and P. A. Wintz, "Image coding by linear transformations and block quantization," IEEE Trans. Commun. Technol. Vol. COM-19, pp. 50-62, Feb. 1971.

[13] A. Habibi and R. S. Hershel, "A unified representation of differential pulse-code modulation (DPCM) and transform coding systems," IEEE Trans. Commun., Vol. COM-22, pp. 692-696, May 1974.

[14] A. Habibi, "Comparison of nth order DPCM encoder with linear transformations and block quantization technoques." IEEE Trans. Commun. Technol., Vol. COM-19, pt. 1, pp. 948-956, Dec. 1971.

[15] ———, "Hybrid coding of pictorial data," IEEE Trans. Commun., Vol. COM-22, pp. 614-624, May 1974.

[16] G. S. Ribonson, "Orthogonal transform feasibility study," prepared for NASA/MSC by COMSAT, Contract no. NAS 9-11240.

[17] N. S. Jayant, "Digital coding of speech waveforms: PCM, DPCM and DM quantizers," Proc. IEEE, Vol. 62, pp. 611-632, May 1974.

[18] J. Max, "Quantizing for minimum distortion," IRE Trans. Inform. Theory, Vol. IT-6, pp. 7-12, Mar. 1960.

[19] J. O. Limb and F. W. Mounts, "Digital differential quantizer for television," Bell Syst. Tech. J., Vol. 48, pp. 2583-2599, Sept. 1969.

[20] D. J. Goodman, "Theory of an adaptive quantizer," IEEE Trans. Commun., Vol. COM-22, pp. 1037-1045, Aug. 1974.

[21] N. S. Jayant, "Adaptive quantization with a one-word memory," Bell Syst. Tech. J., Vol. 52, pp. 1119-1144, Sept. 1973.

[22] L. H. Goldstein and B liu, "Quantization noise in ADPCM systems," IEEE Trans. Commun., Vol. COM-25, pp. 227-238, Feb. 1977.

[23] B. S. Atal and M. R. Schroeder, "Adaptive predictive coding of speech signals," Bell Syst. Tech. J., Vol. 49, pp. 1973-1986, Oct. 1970.

[24] J. B. O'Neal, Jr., "Predictive quantizing systems (differential pulse code modulation) for transmission of television signals," Bell Syst. Tech. J., Vol. 45, pp. 689-721, May-June 1966.

[25] ———, "A bound on signal-to-quantizing noise ratios for digital encoding systems," Proc. IEEE, Vol. 55, pp. 287-292, Mar. 1967.

[26] C. W. Harrison, "Experiments with linear prediction in television," Bell Syst. Tech. J., Vol. 31, pp. 764-783, July 1952.

[27] D. J. Connor, R. C. Brainard and J. O. Limb, "Intraframe coding for picture transmission," Proc. IEEE, Vol. 60, pp. 779-791, July 1972.

[28] D. J. Connor, R. F. W. Pease and W. G. Scholes, "Television coding using two-dimensional spatial prediction," Bell Syst. Tech. J., Vol. 50, pp. 1049-1061, Mar. 1971.

[29] R. P. Abbott, "a differential pulse-code-modulation codec for video telephony using four bits per sample," IEEE Trans. Commun. Technol., Vol. COM-19, pp. 907-912, Dec. 1971.

[30] J. B. Millard and H. I. Maunsell, "Digital encoding of the video signal," Bell Syst. Tech. J., Vol. 50, pp. 459-497, Feb. 1971.

[31] Hiroshi Inose and Mitsuru Ishizuka, "Adaptive coding using two-dimensional prediction for image signals," Electron. & Commun. in Japan, Vol. 57, pp. 50-58, July 1974.

[32] J. O. Limb, C. B. Rubinstein and K. A. Walsh, "Degital coding of color picturephone signals by element-differential quantization," IEEE Trans. Commun. Technol., Vol. COM-19, pt. 1, pp.992-1006, Dec. 1971.

[33] F. W. Mounts, "Video encoding system with conditional picture-element replenishment," Bell Syst. Tech. J., Vol. 48, pp. 2545-2554, Sept. 1969.

[34] R. C. Brainard, F. W. Mounts and B. Prasada, "Low-resolution TV: Subjective effects of frame repetition and picture replenishment," Bell Syst. Tech. J., Vol. 46, pp. 261-271, Jan. 1967.

[35] J. C. Candy, M. A. Franke, B. G. Haskell and F. W. Mounts, "Transmitting television as clusters of frame-to-frame differences," Bell Syst. Tech. J., Vol. 50, pp. 1889-1917, July-Aug. 1971.

[36] J. O. Limb and R. F. W. Pease, "A simple interframe coder for video telephony," Bell Syst. Tech. J., Vol. 50, pp. 1877-1888, July-Aug. 1971.

[37] B. G. Haskell, F. W. Mounts, and J. C. Candy, "Interframe coding of videophone pictures," Proc. IEEE, Vol. 60, pp. 792-800, July 1972.

[38] R. F. W. Pease and J. O. limb, "Exchange of spatial and temporal resolution in television coding," Bell Syst. Tech. J., Vol. 50, pp. 191-200, Jan. 1971.

[39] H. Yasuda, H. Kawanishi, F. Kanaya and H. Hashimoto, "Transmitting 4-MHz TV signals by combinational difference coding," IEEE. Trans. Commun., Vol. COM-25, pp. 508-516, May 1977.

[40] I. Dinstein, E. Yam, A. Gatfield and R. Garlow, "Evaluation of NETEC-22H-- A codec for NTSC color composite TV signals," COMSAT Labs. Tech. Memo., CL-76-76, Dec. 1976.

[41] L. S. Golding, "DITEC -- A digital television communications system for satellite links," Presented at the 2nd Int. Conf. on Digital Satellite Commun., Paris, France, Nov. 1972.

[42] F. deJager, "Delta modulation - A method of PCM transmission using a 1-unit-code," Philips Res. Rep., Vol. 7, pp. 442-466, 1952.

[43] J. B. O'Neal, Jr., "Delta Modulation quantizing noise analytical and computer simulation results for Gaussian and TV input signals," Bell Syst. Tech. J., Vol. 45, pp. 117-142, Jan. 1966.

[44] P. P. Wang, "Idle channel noise of delta modulation," IEEE Trans. Commun. Technol., Vol. COM-16, pp. 737-742, Oct. 1968.

[45] D. J. Goodman, "Delta modulation granular quantizing noise," Bell Syst. Tech. J., Vol. 48, pp. 1197-1219, May-June 1969.

[46] D. Slepian, "On delta modulation," Bell Syst. Tech. J., Vol. 51, pp. 2101-2137, Dec. 1972.

[47] L. J. Greenstein. "Slope overload noise in linear delta modulators with Gaussian inputs," Bell Syst. Tech. J., Vol. 52, pp. 387-422, Mar. 1973.

[48] W. C. Adams, Jr. and J. B. O'Neal, Jr., "Linear delta modulation quantizing noise characteristics," IEEE Trans. Commun. (corresp.), Vol. COM-24, pp. 940-944, Aug. 1976.

[49] M. R. Winkler, "Pictorial transmission with HIDM," Int. Conv. Rec., pt. 1, pp. 285-291, 1965.

[50] J. A. Greefkes and F. deJager, "Continuous delta modulation," Philips Res. Rep., Vol. 23/2, pp. 233-246, 1968

[51] A. Tomozowa and H. Kaneko, "Companded delta modulator for telephone transmission," IEEE Trans. Commun. Technol., Vol. COM-16, pp. 149-157, Feb. 1968.

[52] S. J. Brolin and J. M. Brown, "Companded delta modulator for telephony," IEEE Trans. Commun. Technol., Vol. COM-16, pp. 157-162, Feb. 1968.

[53] R. H. Bosworth and J. C. Candy, "A companded one-bit coder for television transmission," Bell Syst. Tech. J., Vol. 48, pp. 1459-1479, May-June 1969.

[54] A. H. Frei, H. R. Schindler and P. Vettiger, "An adaptive dual-mode coder/decoder for television signals," IEEE Trans. Commun. Technol., Vol. COM-19, pt. 1, pp. 933-944, Dec. 1971.

[55] P. A. Bello, R. N. Lincoln and H. Gish, "Statistical delta modulation," Proc. IEEE, Vol. 55, pp. 308-319, Mar. 1967.

[56] H. Inose and T. Aoki, "Asynchronous delta-modulation system," Electron. Commun. (Japan), pp. 34-42, Mar. 1966.

[57] T. A. Hawkes and P. A. Simonpieri, "Signal coding using Asynchronous delta modulation," IEEE Trans. Commun., Vol. COM-22, pp. 346-348, Mar. 1974.

[58] C. V. Chakravarthy and M. N. Faruqui, "Two loop adaptive delta modulation systems," IEEE Trans. Commun., Vol. COM-22, pp. 1710-1713, Oct. 1974.

[59] ———,"A multidigit adaptive delta modulation (ADM) system," IEEE. Trans. Commun., Vol. COM-24, pp. 931-935, Aug. 1976.

[60] P. Papantoni-Kazakos and G. C. Collins, "A three-level adaptive delta modulator," IEEE Trans. Commun., Vol. COM-25, pp. 532-536, May 1977.

[61] J. E. Abate, "Linear and adaptive delta modulation," Proc. IEEE, Vol. 55, pp. 298-308, Mar. 1967.

[62] C. L. Song, J. Garodnick and D. L. Schilling, "A variable-step-size robust delta modulator," IEEE Trans. Commun., Vol. COM-19, pp. 1033-1044, Dec. 1971.

[63] C. L. Song, "Adaptive delta modulation," Ph.D. dissertation, Polytechnic Inst. Brooklyn, New York, 1971.

[64] D. L. Schilling, et al., "Delta Modulation," Jan.-Dec. 1, 1972, Final Report for NASA Grant NGR 33-013-063.

[65] N. S. Jayant, "Adaptive delta modulation with a one-bit memory," Bell Syst. Tech. J., Vol. 49, pp. 321-342, Mar. 1970.

[66] D. L. Schilling, et al., "Investigation of television transmission using adaptive delta modulation principles," April 15, 1975 - April 14, 1976, Final Report for NASA Contract NAS9-13940.

[67] L. Weiss, I. M. Paz and D. L. Schilling, "Video encoding using an adaptive digital delta modulator with overshoot suppression," IEEE Trans. Commun., Vol. COM-23, pp. 905-920, Sept. 1975.

[68] T. Oshima and T. Ishiguro, "Reduction of edge busyness in delta modulation," IEEE Trans. Commun., Vol. COM-23, pp. 550-554, May 1975.

[69] C. C. Cutler, "Delayed-encoding: Stabilizer for adaptive coders," IEEE Trans. Commun. Technol., Vol. COM-19, pp. 898-907, Dec. 1971.

[70] L. H. Zetterberg and J. Uddenfeldt, "Adaptive delta modulation with delayed decision," IEEE Trans. Commun., Vol. COM-22, pp. 1195-1198, Sept. 1974.

[71] ———, "Algorithms for delayed encoding in delta modulation with speech-like signals," IEEE Trans. Commun., Vol. COM-24, pp. 652-658, June 1976.

[72] T. R. Lei, N. Scheinberg and D. L. Schilling, "Two dimensional delta modulation for picture encoding," Int. Symp. on Inform. Theory, Ronneby, Sweden, June, 1976.

[73] ———, "Adaptive delta modulation system for video encoding," IEEE Trans. Commun., Vol. COM-25, pp. 1302-1314, Nov. 1977.

[74] ———, "Delta modulation of video signals," Nat. Telecommun. Conf., Los Angeles, Dec. 1977.

[75] H. R. Schindler, "Delta modulation," IEEE Spectrum, Vol. 7, pp. 69-78, Oct. 1970.

[76] H. R. Schindler, "Linear, nonlinear, and adaptive delta modulation," IEEE Trans. Commun., Vol. COM-22, pp. 1807-1823, Nov. 1974.

[77] J. B. O'Neal, Jr., "Delta modulation of data signals," IEEE Trans. Commun., Vol COM-22, pp. 334-339, Mar. 1974.

[78] S. Tazaki, H. Osawa and Y. Shigematsu, "A useful analytical method for discrete adaptive delta modulation," IEEE Trans. Commun., Vol. COM-25, pp. 193-199, Feb. 1977.

[79] S. A. Cutts, "design of a television interframe delta modulation system," prepared for NASA/Johnson Space Center by Lockheed, Nov. 1976.

[80] J. Garodnick, "Digital processing in communications systems" PH.D. dissertation, The City University of New York, New York, New York, June 1972.

[81] E. F. Brown, "Low-resolution TV: Subjective Comparison of interlaced and noninterlaced pictures," Bell Syst. Tech. J., Vol. 46, pp. 199-232, Jan. 1967.

[82] E. G. Bowen and J. O. Limb, "Subjective effect of substituting lines in a video-telephone signal," IEEE Trans. Commun. (Corresp.), Vol. COM-24, pp. 1208-1212, Oct. 1976.

[83] F. W. Campbell, "The human eye as an optical filter," Proc. IEEE, Vol. 56, pp. 1009-1014, June 1968.

[84] W. C. Adams, Jr., "Entropy measurements for three adaptive source encoders," IEEE Trans. Commun. (Corresp.), Vol. COM-24, pp. 131-133, Jan. 1976.

[85] S. K. Goyal and J. B. O'Neal, Jr., "Entropy coded differential pulse-code modulation systems for television," IEEE Trans. Commun., Vol. COM-23, pp. 660-666, June 1975.

[86] B. G. Haskell, "Entropy measurements for nonadaptive and adaptive frame-to-frame linear-predictive coding of Videophone signals," Bell Syst. Tech. J., p. 1175, July-Aug. 1975.

[87] K. Virupaksha and J. B. O'Neal, Jr., "Entropy-coded adaptive differential pulse-code modulation (DPCM) for speech," IEEE Trans. Commun., Vol. COM-22, pp. 777-787, June 1974.

[88] J. B. O'Neal, Jr., "Differential pulse-code modulation (PCM) with entropy coding," IEEE Trans. Inform. Theory, Vol. IT-22, pp. 169-174, Mar. 1976.

[89] D. A. Huffman, "A method for the construction of minimum redundancy codes," Proc. IRE, Vol. 40, pp. 1098-1101, Sept. 1952.

[90] J. Capon, "A probabilistic model for run-length coding of pictures," IRE Trans. Inform. Theory, Vol. IT-5, pp. 157-163, Dec. 1959.

[91] J. I. Molinder, "Optimal coding with a single standard run length," IEEE Trans. Inform. Theory, Vol. IT-20, pp. 336-343, May 1974.

[92] R. G. Gallager and D. C. Van Voorhis, "Optimal source codes for geometrically distributed integer alphabets," IEEE Trans. Inform. Theory, Vol. IT-21, pp. 228-230, March 1975.

[93] B. Arazi, "An optimal coding scheme for a certain type of Ensemble," IEEE Trans. Commun., Vol. COM-23, pp. 741-743, July 1975.

[94] T. S. Huang, "An upper bound on the entropy of run-length coding," IEEE Trans, Inform. Theory, Vol. IT-20, pp. 675-676, Sept. 1974.

[95] T. S. Huang and A. B. Shahid Hussain, "Facsimile coding by skipping white," IEEE Trans. Commun., Vol. COM-23, pp. 1452-1460, Dec. 1975.

[96] H. Meyr, H. G. Rosdolsky and T. S. Huang, "Optimum run length codes," IEEE Trans. Commun., Vol. COM-22, pp. 826-835, June 1974.

[97] S. A. Smith, "A generalization of Huffman coding for messages with relative frequencies given by upper and lower bounds," IEEE Trans. Inform. Theory, (Corresp.) Vol. IT-20, pp. 124-125, Jan. 1974.

[98] J. C. Lawrence, "Application of Schalkwijk source coding techniques to pictorial sources," Report NELC Z179, Naval Electron. Lab. Center, San Diego, Calif., March 1973.

[99] N. Scheinberg and D. L. Schilling, "Techniques for correcting transmission errors in video adaptive delta modulation channels," IEEE Trans. Commun., Vol. COM-24, pp. 1064-1069, Sept. 1976.

[100] K. Chang and R. W. Donaldson, "Nonadaptive DPCM transmission of monochrome pictures over noisy communication channels," IEEE Trans. Commun., Vol. COM-24, pp. 173-183, Feb. 1976.

[101] D. J. Connor, "Techniques for reducing the visibility of transmission errors in digitally encoded video signals," IEEE Trans. Commun., Vol. COM-21, pp. 695-706, June 1973.

[102] R. Lippmann, "A technique for channel error correction in differential PCM picture transmission," IEEE Int. Conf. Commun., Conf. Rec., Vol 11, June 11-13, 1973.

[103] J. C. Candy, "Limiting the propagation of errors in one-bit differential codecs," Bell Syst. Tech. J., Vol. 53, pp. 1667-1676, Oct. 1974.

[104] R. J. Arguello, H. R. Sellner and J. A. Stuller, "The effect of channel errors in the differential pulse-code-modulation transmission of sampled imagery," IEEE Trans. Commun. Technol., Vol. COM-19, pp. 926-933, Dec. 1971.

[105] P. Jung, VDE/NTG and R. Lippmann, VDE/NTG, "Error response of DPCM decoders," Communication from the German Aerospace Research and Experimental Establishment (DFVLR), Research Centre Braunschweig, Institut für Flugführung, manuscript May 1975.