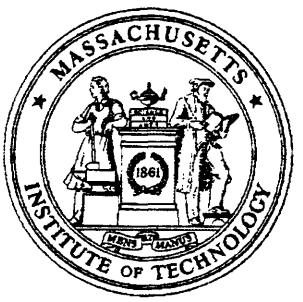


Channel Equalization and Interference Reduction Using Scrambling and Adaptive Amplitude Modulation

RLE Technical Report No. 558

June 1990

Adam S. Tom

Research Laboratory of Electronics
Massachusetts Institute of Technology
Cambridge, MA 02139 USA

This work was supported in part by the Advanced Television Research Program.

Channel Equalization and Interference Reduction Using Scrambling and Adaptive Amplitude Modulation

by

Adam S. Tom

Submitted to the
Department of Electrical Engineering and Computer Science
on May 11, 1990 in partial fulfillment of the requirements for the Degree of
Doctor of Philosophy in Electrical Engineering

Abstract. In order to appreciate the increased resolution of high-definition images, a means of reducing the degradation due to channel defects needs to be employed. Conventional methods of channel equalization use adaptive filters. These methods require long convergence times, are limited by the length of the filters, and require high computational complexity at the receiver. These are undesirable for television receivers which should be available to the consumer at a resonable cost. In addition, these methods do not reduce additive channel noise, nor do they reduce interference from other signal sources.

Presented in this thesis is a new method of channel equalization, interference reduction, and noise reduction based upon Adaptive Amplitude Modulation and Scrambling. Adaptive Amplitude Modulation is a noise reduction technique, and Scrambling is a scanning method that decorrelates any channel defects and makes them look like random noise to the desired signal. Also presented is a robust transmission system termed *Adaptively Modulated and Scrambled Pulse Amplitude Modulation* (AMSC-PAM) which incorporates this new channel equalization and interference reduction method.

This new method of channel equalization and interference reduction has several advantages. Unlike conventional equalization schemes, which employ adaptive filters, our equalization and interference-reduction scheme does not require a long convergence time to find the filter coefficients. It only requires simple computations at the receiver and is not limited to a maximum length of the channel impulse response that it can equalize. Our scheme can equalize a channel that has an impulse response of any length and does not require transmission of a training sequence. Our equalization and interference reduction method is limited, however, by the power of the channel defects and is more sensitive to mistiming errors than transmission methods that do not use Scrambling.

AMSC-PAM is a robust transmission system that codes a signal so that it can better withstand channel defects and delivers greatly improved images to the receiver without having to increase the peak transmission power. AMSC-PAM can make channel degradations of up to a 25dB-CNR power level imperceptible to just perceptible in the demodulated image. In addition, AMSC-PAM can be used in conjunction with conventional equalization methods

The Adaptive Amplitude Modulation algorithm is optimized, and its ability to reduce noise is increased over existing methods. Scrambling is discussed, and the motivation behind Scrambling is given. The type of transmission defects that are considered are mistiming errors, additive channel noise, co-channel and adjacent-channel interference, multipath, and frequency distortion. The AMSC-PAM transmission system is described,

and its ability to equalize a channel and to reduce interference and noise at radio frequencies (RF) is investigated.

Thesis Supervisor: Dr. William F. Schreiber

Title: Professor of Electrical Engineering

To
Mom and Dad

Acknowledgments

I would like to thank my advisor Dr. William F. Schreiber for his support and guidance. He made my stay at MIT a pleasant one. Not only did he provide technical insights but tried to instill in all of us a social responsibility and awareness that, hopefully, we shall all carry with us in our careers. I would also like to thank my readers Dr. Jae S. Lim and Dr. Arun N. Netravali for their support and viewpoints.

I am grateful to John Wang, Chong Lee, and Dave Kuo for maintaining our computing and display facilities and to Julien Piot for helpful discussions and suggestions. I would also like to thank Cindy LeBlanc for making everything go smoothly.

Many friends and family have made my stay at MIT enjoyable; I wish to thank them all. Sometimes during one's studies, one can be overwhelmed with the trials of school. It is these people that help one carry on. Thank you.

This thesis was supported by the Advanced Television Research Program.

Contents

List of Figures

Chapter 1

Introduction

In many communications applications, a transmitter sends a desired signal through a nonideal channel, which degrades the signal in some unknown manner, and a receiver observes this degraded signal. The types of degradation that may occur in the channel are interference from other signal sources, intersymbol interference (including multipath, a nonideal frequency response, and mistiming errors in sampling), and additive noise. Any signal processing technique or algorithm used to reduce intersymbol interference falls under the term channel equalization. Common methods of channel equalization and of reducing additive noise employ the use of adaptive filters [Luc65] [Luc66] [LR67] [Gia78]. Proper bandlimiting, band separation, and in extreme cases, spread-spectrum techniques are examples of methods for reducing interference from other signals [Dix84].

Unfortunately, the causes of intersymbol interference are not stationary, and hence, adaptive filters are necessary to reduce the amount of the resulting degradation. If the channel were stationary, then the channel characteristics would only need to be determined once, and we could fix the filter in each receiver. This would require very little signal processing while the signal was being received because the filter had already been determined. Adaptive filters make an estimate, $\hat{i}[n]$, of the desired signal, $i[n]$, given the observations, $o[n], o[n-1], \ldots, o[0]$. The estimate, $\hat{i}[n]$, is formed by applying a time-varying FIR filter, $h[n]$, with M coefficients to a subset of the available observations. The filter is allowed to vary in time in order to adapt to the changing characteristics of the observation. Upon each new observation, $o[n]$, the filter coefficients chosen are those

that minimze a function of the error between the desired signal, $i[n]$, and the estimated response, $\hat{i}[n]$, where the error is,

$$e[n] = i[n] - \hat{i}[n]. \tag{1.1}$$

Two common algorithms used to calculate the filter coefficients are the Least Mean Squared Error (LMS) and the Recursive Least Squares (RLS) algorithms [Orf85]. Finding the optimal filter which minimizes the mean-squared error requires knowledge of the statistics of the observations and of the error, which we do not have. Solving the optimal filter by means of the RLS algorithm requires inverting an nxn matrix when observation $o[n]$ is made. In order not to invert directly the matrix, the algorithms do an iterative search for the filter coefficients.

These two channel equalization schemes calculate the optimal adaptive filter iteratively either by a gradient search or by recursively finding the inverse of an nxn matrix. Gradient search methods can have a long convergence time that depends upon the spread of the eigenvalues of the autocorrelation matrix of the observation samples, $y[n]$. The convergence time is an indication of the algorithm's ability to track the nonstationarities of the channel. Methods that invert a matrix have shorter convergence times; however the complexity of the calculations is high. Neither of these methods actually reaches the optimal filter coefficients, but fluctuates about the optimal values. The ability of these methods to equalize the channel is limited by the length of the filter. For example, an adaptive transversal filter of length M cannot equalize an echo that has a delay longer than M samples. In addition, a training sequence must be transmitted periodically as part of the signal in order to facilitate the adaptive filter algorithms.

Described herein is a novel method of channel equalization, noise reduction, and interference reduction based upon the ideas of adaptive amplitude modulation and pseudo-random scanning (i.e., scrambling). Adaptive amplitude modulation is a noise-reduction method that advantageously exploits the visual masking phenomenon. Because changes in luminance mask noise in close proximity to the luminance changes, noise is less visible in busy areas than in blank or slowly varying areas of an image. This is a consequence of visual masking. Adaptive amplitude modulation makes use of this property of the

human visual system by adaptively varying the modulation indices of the transmitted video signal in such a manner that noise is significantly reduced in the blank and slowly varying areas of the image and reduced to a lesser extent in the busy areas. The end result is a greatly improved image at the receiver.

Implementation of adaptive amplitude modulation requires that the original input image be decomposed into a lowpass and a highpass frequency component. The decimated lowpass signal is sent digitally in some noise-free manner, and the highpass component is adaptively modulated by multiplying it with a time-varying set of modulation indices, or adaptation factors. It is important to note that the signal which has been adaptively modulated has the same peak power as the original signal. The cost of using adaptive amplitude modulation is side information that must be sent to the receiveer; however, the amount of side information is small.

Scrambling is a method of decorrelating channel defects and makes degradations appear as random noise. The idea of scrambling is an extension of the idea behind Roberts' pseudorandom noise technique, or dithering, where quantization noise due to coarse quantization is converted to uncorrelated noise, which is less subjectively objectionable than correlated noise. Scrambling is a method of pseudorandomly scanning through an image; thus, the transmitted signal looks like random noise to any other signal. Conversely, after descrambling at the receiver, any other signal (even the transmitted signal's own reflection) looks like random noise to the received, desired signal.

The combination of adaptive amplitude modulation and scrambling equalizes the channel and reduces interference and noise by first adaptively modulating the video signal and then scrambling it at the transmitter. At the receiver, the signal is first unscrambled and then adaptively demodulated. Scrambling converts any intersymbol intersymbol interference, frequency distortion, or interference from other signals into pseudorandom noise, and adaptive amplitude modulation reduces any pseudorandom degradations in the image and additive channel noise. In this manner, interference and noise are reduced, and the channel is equalized. This method requires no convergence time, no complicated mathematical calculations, and no training sequences. In addition, adaptive amplitude modulation and scrambling are not limited by the length of a filter

or the associated fluctuations of the coefficients; adaptive amplitude modulation and scrambling can equalize channels with an impulse response of any duration. The use of adaptive amplitude modulation and scrambling is limited, however, by the power of the resulting noise-like degradation due to scrambling and is sensitive to the synchronization of the sampling at the receiver as are the conventional equalization algorithms.

Because of the recent interest in the efficient transmission of high-definition images or television (HDTV), this research shall use images as the primary source of signal data. This thesis consists of four parts. A basic method of adaptive amplitude modulation exists; however, its ability to reduce noise can be improved. In Chapter 2, we investigate in depth the adaptive amplitude modulation algorithm. In this portion of the research we optimize and improve the performance of the adaptive amplitude modulation algorithm by determining the minimum amount of low-frequency information that we must transmit digitally, investigating different methods of calculating the adaptation factors and of interpolation, and determining the best tradeoff between noise reduction and distortion. In Chapter 3, we describe and give the motivation behind scrambling. In Chapter 4, we explore the interaction of scrambling and adaptive amplitude modulation to determine how much noise reduction should be performed in the presence of scrambling. In Chapter 5, we apply adaptive amplitude modulation and scrambling to reduce interference in and equalize a time-varying channel at radio frequencies. We will look at mistiming errors, additive channel noise, co-channel and adjacent-channel interference, frequency distortion, and multipath. This will entail finding the degree of intersymbol interference and interference from other signals that adaptive amplitude modulation and scrambling can make imperceptible to just perceptible.

Throughout this thesis we have processed 256x256 pictures; however, all processing is done on the assumption of 512x512 sized pictures. This means that the pictures in this thesis should be viewed at 8 times the 256x256 picture height, and not the usual 4 times the picture height.

Before we begin, we should define two measures that will be used throughout this thesis. They are the signal-to-noise ratio (SNR) and the carrier-to-noise ratio (CNR).

The signal-to-noise ratio is defined as follows,

$$SNR = 20\log_{10}(\frac{\text{peak-to-peak signal}}{\text{rms error}}). \tag{1.2}$$

Similarly, the carrier-to-noise ratio is defined as follows,

$$CNR = 20\log_{10}(\frac{\text{peak-to-peak transmitted signal}}{\text{rms noise in channel}}). \tag{1.3}$$

The signal-to-noise ratio is a measure of the relative strengths of the desired signal and any differences between the desired signal and the demodulated signal. The carrier-to-noise ratio, however, is a measure of the relative strengths of the modulated signal that is emitted from the transmitter and any differences between this modulated signal and the signal that appears at the antenna of the receiver.

Chapter 2

Adaptive Amplitude Modulation

In ordinary double sideband amplitude modulation (AM) the transmitted signal, $s(t)$, is formed by adding a constant to a scaled version of the baseband signal, $i(t)$, and multiplying the sum by a sine wave. Assuming that the baseband signal is normalized with $\mid i(t) \mid \leq 1$, then we can write the AM signal as

$$s(t) = A_c[1 + mi(t)] \cos(2\pi f_c t + \phi), \tag{2.1}$$

where A_c is the carrier amplitude, f_c is the carrier frequency, ϕ is a constant phase term that can be set to zero by an appropriate choice of the time origin, and m is the time invariant modulation index. We can recover the baseband signal using an envelope detector or a synchronous demodulator. If a synchronous demodulator is used, the AM signal, $s(t)$, is multiplied by the carrier to give

$$s(t) \cos 2\pi f_c t = \frac{1}{2} A_c[1 + mi(t)] \cos 4\pi f_c t + \frac{A_c}{2} + \frac{1}{2} A_c mi(t). \tag{2.2}$$

A lowpass filter and a dc blocking circuit will allow the recovery of the baseband signal, $i(t)$, or if we wish $mi(t)$. When synchronous demodulation is used, the dc offset (i.e., the '1') is not needed in eqn. 2.1. In adaptive amplitude modulation, the modulation index is allowed to vary with time, so that

$$s(t) = A_c[1 + m(t)i(t)] \cos(2\pi f_c t + \phi), \tag{2.3}$$

where $m(t)$ is the time-varying modulation index. What do we gain by this?

Adaptive Amplitude Modulation is a transmission method that adjusts the modulation index of a video signal according to local image characteristics and, thereby, reduces the noise added during transmission [SB81]. The time-varying modulation index must be sent along with the video signal; however, this signal need not have high resolution and, thus, requires a small amount of bandwidth. Prior to transmission, the video signal is multiplied by its locally adapted modulation index, and, after reception at the receiver, the video signal is divided by the same time-varying modulation index. Because additive channel noise is much more noticeable in the stationary and slowly varying portions of an image than in the quickly varying portions (a consequence of visual masking), the modulation index should be adjusted such that the SNR of the demodulated signal is larger in the stationary and slowly varying portions than in the quickly varying portions. This is done by making the modulation index larger in the stationary and slowly varying portions of the picture and smaller in the quickly varying portions of the picture. We shall call the process of finding the modulation index and multiplying it with the image data *Adaptive Modulation.*

Since more noise reduction should be performed in the stationary and slowly varying portions of a sequence and in order not to violate the peak power constraint on the transmitted signal (i.e., $\mid m(t)i(t) \mid \leq 1$), adaptive modulation should be applied to those components of the sequence that have small amplitudes when the sequence is stationary and slowly varying and large amplitudes when the sequence is busy. This means that adaptive modulation should be performed on the highpass frequency components of the sequence, $i(t)$. We shall denote this highpass frequency component as $i_h(t)$.

Adaptive Amplitude Modulation reduces noise added in the channel by first multiplying the signal, $i_h(t)$, by the index, $m(t)$ at the transmitter, and then dividing by $m(t)$ at the receiver. Since the modulation indices are the same at both transmitter and receiver, the baseband signal is unchanged; whereas, the noise added in the channel is reduced by the factor $m(t)$. The modulation index, $m(t)$, is derived from some measure of "business" of the signal, $i(t)$. Schreiber and Buckley [SB81] use the inverse of the maximum absolute value in a two-dimensional, spatial block of a three-dimensional, highpass video signal. The three-dimensional signal is in (x, y, t) space and the one-dimensional signal

$i(t)$ used for transmission may be derived from three-dimensional space by scanning in a raster fashion, although not limited to a raster scan as we shall see. Other measures for calculating the modulation index based upon the power or variance of $i(t)$ can be used; however, they add complexity to the adaptive modulation algorithm without significantly increasing the performance [Zar78]. The modulation index, $m(t)$ is chosen so that the modulated signal makes maximum use of the peak power capacity of the channel. The modulation index is larger than one, but not so large as to cause noticeable distortion in the demodulated signal. In those regions where the signal is of small amplitude, the modulation index is largest, thus performing the most noise reduction, and in those regions where the signal is large the index is smallest, thus performing little noise reduction.

The adaptive modulation algorithm that we shall use is shown in Figure 2.1. The demodulation process is shown in Figure 2.2. We will be working in the digital domain and, hence, with sampled versions of both the input signal $i(t)$ and the modulation index $m(t)$. The three-dimensional sequence of images, $i[n_1, n_2, n_3]$, will be filtered into a 3-D lowpass signal and a 3-D highpass signal, $i_h[n_1, n_2, n_3]$. The lowpass signal will be sent noise-free using digital transmission, and the highpass signal will be adaptively modulated, companded by the nonlinear amplifier (NLA), and clipped. Buckley [Buc81] used various power law companders for the nonlinear amplifier and found that an exponent of $n = 0.82$ in the nonlinear amplifier, $g(y) = y^{1/n}$, is nearly optimal when used in conjunction with adaptive companding. We shall use this value for n in the nonlinear amplifier.

This 3-D digital adaptively modulated signal is then scanned into a 1-D signal and transmitted via pulse amplitude modulation (PAM). From this point onward, we prefer to call the modulation indices "adaptation factors" and denote them as $m[n_1, n_2, n_3]$. One adaptation factor is calculated for each block of $N_1 \text{x} N_2 \text{x} N_3$ picture elements, where the subscripts correspond to the horizontal, vertical, and temporal directions, respectively. Since the adaptation factors are coarsely sampled, they need to be interpolated back to full resolution for multiplication with the input highs, $i_h[n_1, n_2, n_3]$. Each input element will then have its own adaptation factor. Note that the product of the adaptation factor and the input highs will not always be less than or equal to one for a normalized input

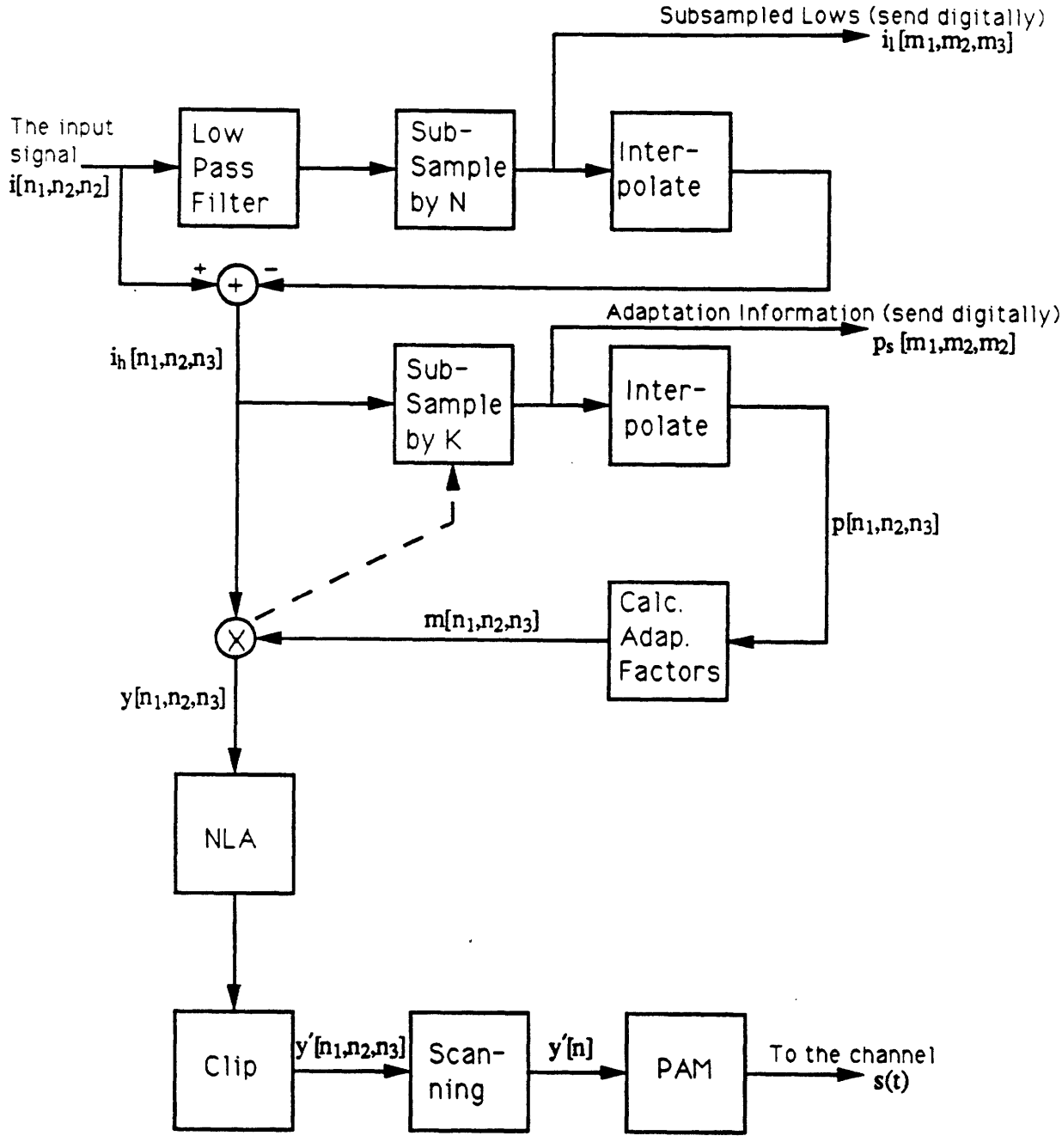


Figure 2.1: Block diagram of adaptive modulation. The dashed line indicates the use of an iterative method to calculate the modulation indices. This will be discussed in a later section.

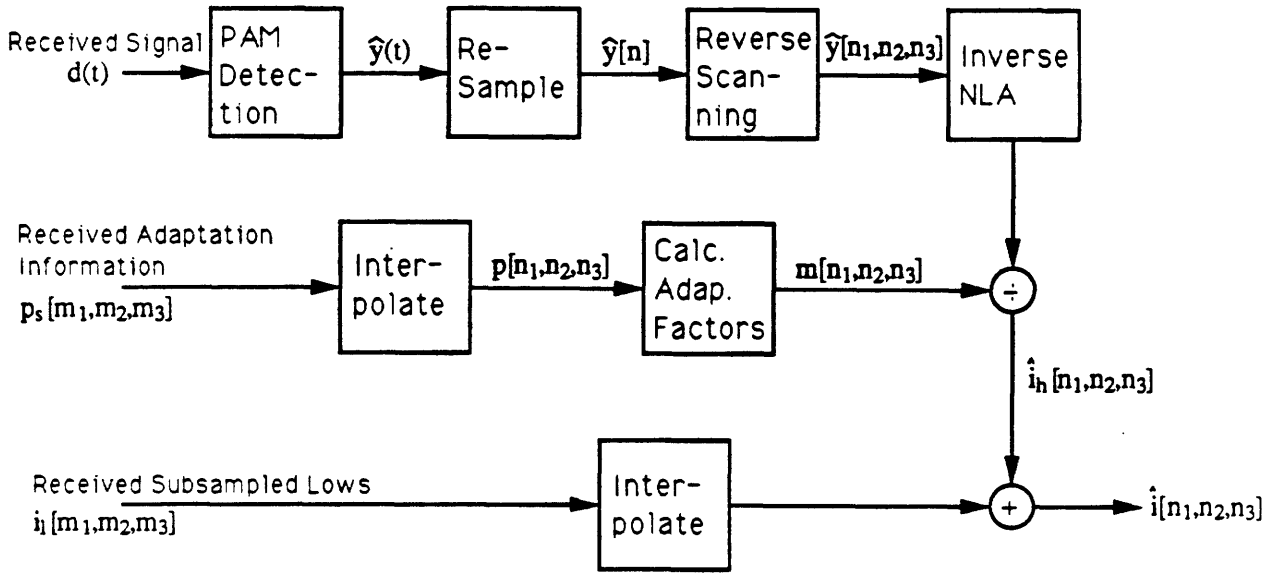


Figure 2.2: Block diagram of adaptive demodulation.

signal due to the subsampling and interpolation of the adaptation factors. In order not to violate the peak power constraint, the product will have to be "clipped" to one. Under noiseless conditions, this operation will produce some distortion in the demodulated output, $\hat{i}[n_1, n_2, n_3]$. At the receiver, a degraded PAM signal, $r(t)$, is detected, sampled, and converted to a 3-D sequence by the scanner to produce a degraded $\hat{y}[n_1, n_2, n_3]$. The coarsely sampled, or in other words subsampled, data, $p_s[n_1, n_2, n_3]$, is interpolated and converted to adaptation factors. The output of the inverse nonlinear amplifier is divided by the full resolution adaptation factors, and the result is a degraded version of the input highs. The subsampled lows are interpolated and added to the degraded highs to produce the demodulated output, $\hat{i}[n_1, n_2, n_3]$.

This chapter will investigate those areas of the adaptive modulation algorithm that have been inadequately studied in the literature [SB81] [Buc81] [Zar78] and that are pertinent to the transmission of high-definition sequences. Because adaptive modulation can only be performed on the highpass component, it is necessary to send the lowpass

component in some noise-free manner, i.e., digitally with enough error protection. In order to keep the amount of digital lowpass information to a minimum, we need to determine how small we can make the lowpass cutoff frequencies and still have adequate noise reducing capabilities near edges. This is the topic of the first section. As mentioned earlier, Schreiber and Buckely used the maximum absolute value of the input signal in a block to calculate the adaptation factors. A bilinear filter was used to interpolate these "subsampled" maximum values, and an inverse operation in conjunction with a scaling operation was used to form the time-varying modulation index, or adaptation factors. This is a very simple method and is not efficient in that it does not optimally find the adaptation factors by trading distortion due to clipping of the transmitted signal with the amount of noise reduction gained by using larger adaptation factors. The calculation of the adaptation factors is an area that requires more investigation. Included in this is the means of subsampling and interpolation of the adaptation factors.

2.1 Lowest Frequency Component

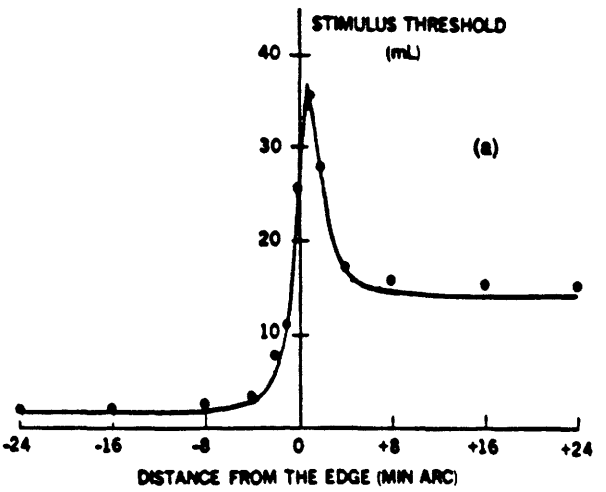
The first step in adaptive modulation is to separate the input signal into low- and high-frequency components. It has been proposed in recent literature to use Quadrature Mirror Filters (QMF) to perform the band separation when transmitting high-resolution images [Sch88]. This results in one DC component and possibly numerous high-frequency components. In order to isolate the effects of our experiments upon the subjective quality of images, we have chosen to decompose our input signal into one low-frequency component, which contains DC, and one high-frequency component as Troxel, et. al. [Tro81] and Schreiber and Buckley [SB81] have done in their two-channel picture coding systems. The band separation is performed as in Figure 2.1 by lowpass filtering the input signal, $i[n_1, n_2, n_3]$, subsampling, interpolating, and subtracting from the input signal to form the highpass signal. Even though this research was done on a two-channel system, the results are expected to be applicable to the high-frequency components obtained by quadrature mirror filtering.

Buckley [Buc81] based his choice of spatial cutoff frequency on the minimum color

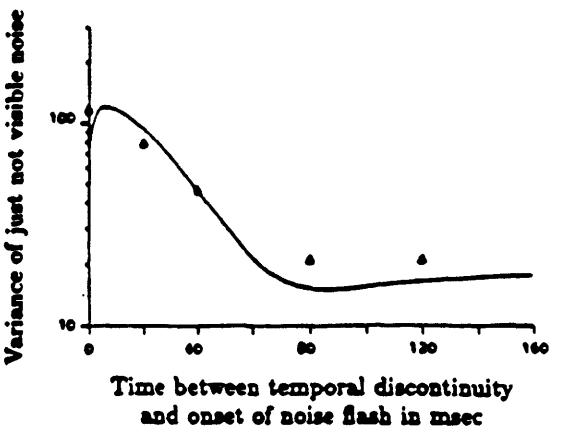
resolution needed to render good-looking color images. He concluded that a cutoff frequency of $\pi/5$ gave adequate color rendition and used this cutoff frequency for both chrominance and luminance. Schreiber and Buckley and Troxel, et. al. chose the same cutoff frequency. In his work on varying the quantization of the information content of different frequency bands of a signal, Kretzmer [Kre56] used a cutoff frequency of 0.5 MHz for a 4.0 MHz signal, but gives no explanation for this choice. This corresponds to a cutoff frequency of $\pi/8$. These cutoff frequencies are suboptimal in terms of the perceptual gains offered by the visual masking effect.

As mentioned previously, the amount of low-frequency data that we send digitally depends upon the cutoff frequency of the three-dimensional lowpass filter. The input data, $i[n_1, n_2, n_3]$, is lowpass filtered with a filter that has cutoff frequencies of $\pi/L_1, \pi/L_2$, and π/L_3 in the horizontal, vertical, and temporal frequency spaces, respectively. This signal is decimated by L_1, L_2, and L_3 in the respective dimensions. The decimated lowpass signal is interpolated and then subtracted from the input sequence to produce the highpass signal, $i_h[n_1, n_2, n_3]$. The decimated lowpass signal, $i_l[m_1, m_2, m_3]$, is sent digitally with assumed very little error. The larger L_1, L_2, and L_3 are, the less data we need to send digitally.

The permissible cutoff frequencies depend upon the spatial and temporal extent of visual masking. Visual masking was first studied by B. H. Crawford [Cra47], who, in 1946, studied the duration and amplitude of the impairment upon visual acuity a soldier may suffer due to flashes of military gunfire and explosions. Today, visual masking is defined as the destructive interference by an abrupt change in luminance levels upon stimuli in close proximity, both spatialy and temporaly, to the luminance gradient [Fox]. The threshold of detection of a small test stimulus near a luminance edge is a decreasing function of the absolute distance from the edge; the masking effect reaches its height at the edge and decreases in effect after several minutes of arc on either side of the edge, as shown in Figure 2.3(a). Similarly, the temporal masking effect reaches its height at the luminance change and decreases after many milliseconds, as shown in Figure 2.3(b). In addition, the higher the contrast of the luminance edge, the greater is the diminishing effect upon stimuli near the luminance change [FJdF55] [Spe65].



(a)



(b)

Figure 2.3: The spatial masking effect (a) decreases on either side of a luminance change (from Lukas et. al. [LTKL80].) Shown is the threshold of visibility for a test stimulus versus distance from the edge for a luminace edge having a 1:20 luminance ratio. Likewise, temporal masking (b) is a decreasing function of the time since a temporal luminance change (from Girod [Gir89]). Shown is the variance of just not visible noise versus time for a temporal luminance change of from 50 to 180.

If we consider the absolute value of the high-frequency, unit-step response resulting from the two-channel frequency decomposition (Figure 2.4(d)), we see that the response has a peak at the luminance change and decreases on each side. Since the adaptation factors are inversely proportional to the absolute value of the highs component, the noise-reducing capabilities of adaptive modulation increases as one moves away from the luminance edge. Smaller cutoff frequencies (which mean less digital data) spread an edge over more picture elements and frames thereby diminishing the noise-reducing capabilities of adaptive modulation in a larger area about the edge. We would like to match the noise reducing capabilities of adaptive modulation with the masking effect. This means that the extent of the filtered edge should roughly correspond to the extent of visual masking.

Lukas, Tulunay-Keesey, and Limb [LTKL80] and Girod [Gir89] have shown that spatial masking extends to between 8 min. and 16 min. of arc subtended by the eye, and temporal masking lasts for approximately 120 msec. For a 512x512, 60 frames-per-second sequence of images viewed at 4 times the picture width, this corresponds to between 5 to 10 picture elements spatially and 8 frames temporally; therefore, we would like to choose cutoff filters whose impulse responses approximately extend over this area and which produce filtered edges that have the approximate shape of the visual masking function.

Gaussian filters produce pleasant looking pictures when used for decimation and interpolation [ST85] [Rat80]. In addition, Gaussian filters are separable, are circularly symmetric, and are optimal in their space-width/frequency-width tradeoff. This last property is important because we would like to choose filters that, for a given cutoff frequency, produce impulse responses that are as short as possible. In our set of experiments, we bandlimit with Gaussian filters prior to decimation and interpolate with sharpened Gaussian filters, while paying attention to the decay rate of the unit-step response, the amount of overshoot, and the magnitude of the sampling structure. These factors will affect the performance of the adaptive modulation algorithm. Since we are trying to have the extent of the highpass unit-step response decay within approximately 8 pixels and 8 frames for a 512x512, 60 frames-per-second sequence, we shall consider decimation factors of 6, 8, and 10.

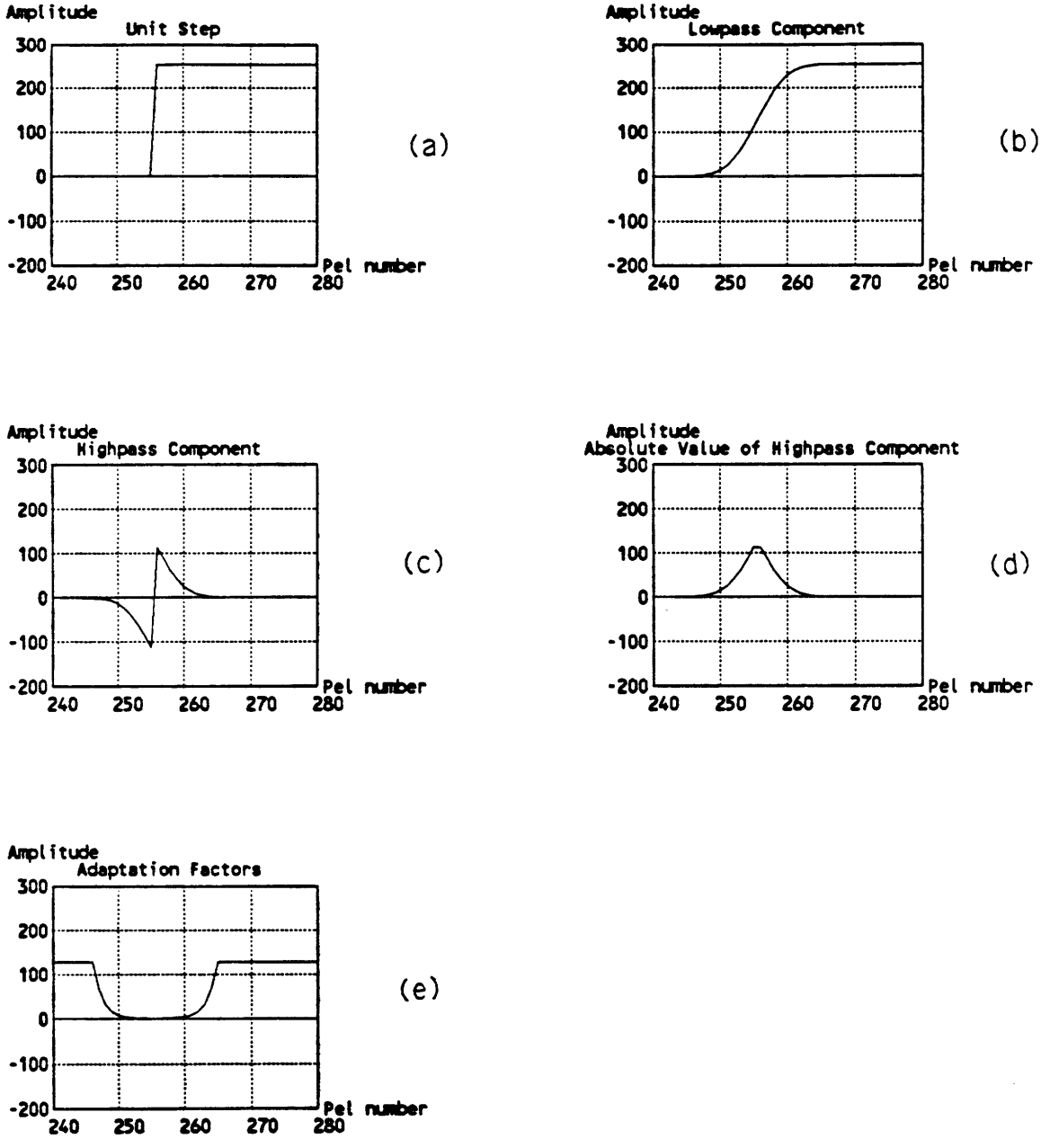


Figure 2.4: The unit-step response of a two-channel system and the ideal adaptation factors for this signal.

σ_{post}	N=6	N=8	N=10
1.5	144.9	224.0	--
2.0	56.8	144.9	--
2.5	16.6	74.0	--
3.0	3.7	31.8	--
3.5	0.6	11.7	45.4
4.0	0.07	3.7	21.7
4.5	0.06	0.94	9.3
5.0	--	0.10	3.4
5.5	--	--	1.3

Table 2.1: Amplitude of sampling structure for various postfilter standard deviations.

We would like to have the fluctuations of the sampling structure that are produced by the sharpened-Gaussian postfilter in uniform regions to be such that after adaptive demodulation, there are no fluctuations in the noise that is passed in these regions. Sampling structure with an amplitude of between ± 1 will produce no fluctuations in the adaptation factors in a uniform region when the maximum adaptation factor is 128. This means that the noise will be uniform over that region. Since the postfilter determines the amplitude of the sampling structure, our first step is to fix the standard deviation of the prefilter to some reasonable value ($\sigma = 3.0$ for $N = 6$, $\sigma = 4.0$ for $N = 8$, $\sigma = 5.0$ for $N = 10$), set the standard deviation of the postfilter, and observe the amplitude of the sampling structure in a uniform region of largest magnitude. A unit-step function between 0 and 255 is used for these experiments. The results of this experiment are shown in Table 2.1. For each decimation factor, the amplitude of the sampling structure decreases as the standard deviation of the postfilter increases. For each decimation factor, we should choose the smallest postfilter standard deviation that produces sampling structure with an amplitude of between ± 1, because a smaller standard deviation will produce a smaller initial overshoot in the highs component, and a smaller overshoot here will allow a faster decay rate when the choice of the standard deviation of the prefilter is made. For decimation factors of $N = 6$, 8, and 10, we choose a postfilter standard

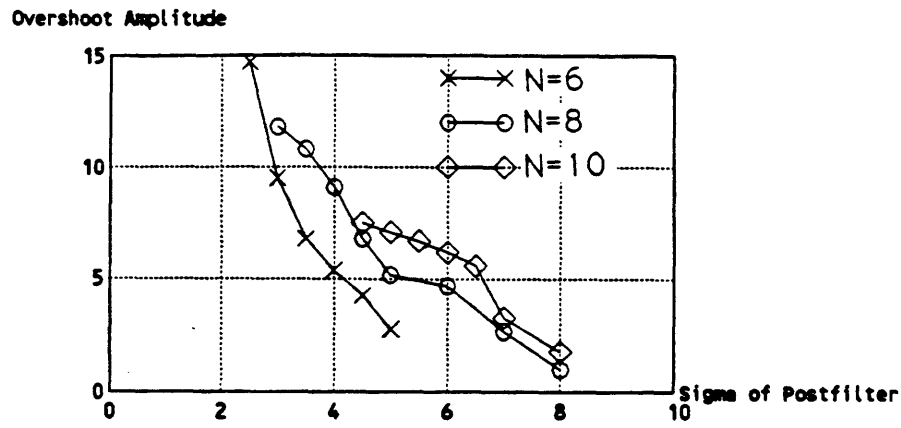


Figure 2.5: Overshoot versus prefilter standard deviation.

deviation of 3.5, 4.5, and 5.5, respectively.

Now that we have selected a postfilter based upon the resulting sampling structure, we need to choose a prefilter, which will be based upon the amount of overshoot and upon the first zero-intercept of the highpass unit-step response. Plotted in Figure 2.5 is the amount of overshoot versus the standard deviation of the prefilter for the postfitlers selected above. Too large an overshoot means that a "halo" of noise will appear several pels away from an edge; however, the smaller the allowed overshoot, the longer will be the extent of the unit-step response. The largest overshoot that produces no noticeable "halo" of noise is approximately 4 to 5. This magnitude of overshoot corresponds to prefiter standard deviations of 4.0, 5.0, and 6.0 for $N = 6$, 8, and 10, respectively.

For each decimation factor, we have a prefilter and postfilter pair whose unit-step response satisfies the sampling structure and overshoot constraints. Which filter we choose is based upon the spatial/temporal extent of this unit-step response, and we should choose the largest decimation factor whose corresponding unit-step response has an extent of approximately 8 pixels and 8 frames for a 512x512, 60 fps sequence. The location where the unit-step response falls to zero is a measure of the extent of the unit-step response. For $N = 6$, 8, and 10 and the above corresponding standard deviations, the zero-intercepts are 7, 8, and 12 pels respectively. For a 512x512, 60 fps sequence

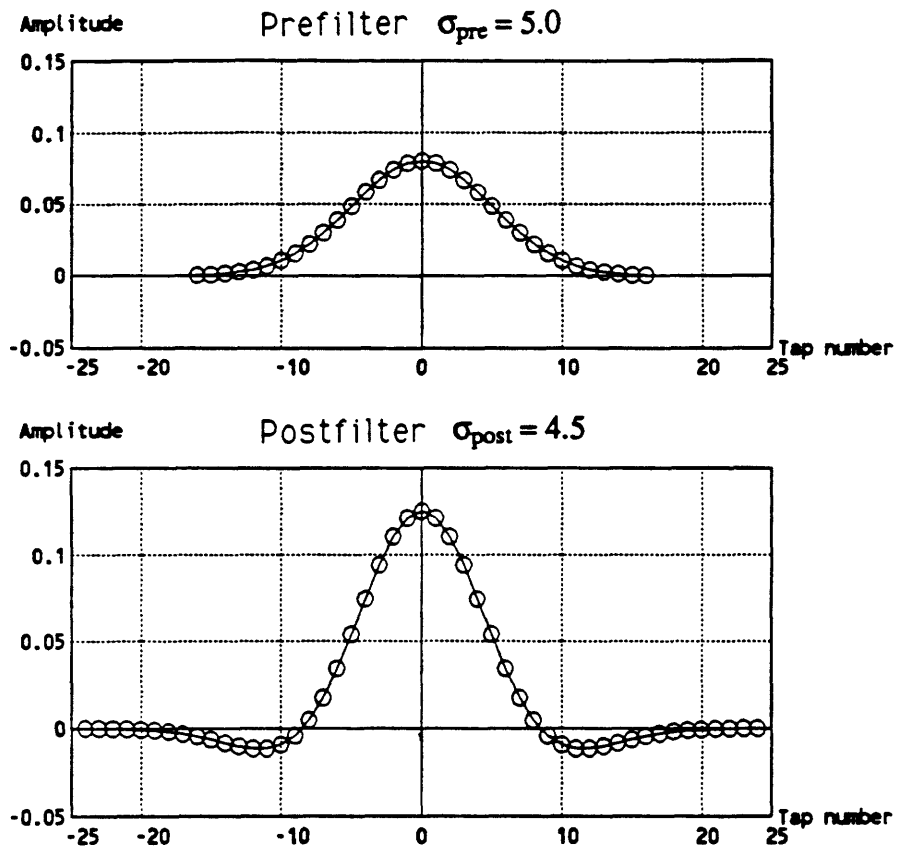


Figure 2.6: The optimal band-separation Gaussian filter-pair. The standard deviation of the Gaussian prefilter is 5.0, and the standard deviation of the sharpened-Gaussian postfilter is 4.5.

the lowpass component should be calculated by using a Gaussian prefitler of standard deviation 5.0, a sharpened Gaussian postfilter of standard deviation 4.5, and a decimation factor of 8. This filter and its unit-step response are shown in Figure 2.6. Figure 2.7 shows the original CMAN and GIRL pictures, and Figure 2.8 shows two 2-D highpass and two decimated 2-D lowpass images produced by this filter-pair.

Because of the long length of this filter pair, its use in the temporal direction is impractical. The shortest filter pair that produces the desirable decay rate is the averaging prefilter and the linear postfilter. This filter pair produces a unit-step response that has linear decay and a zero-intercept at 8 pels. The decay rate corresponding to the Gaussian filter pair is faster than that corresponding to the linear filter pair.; however, the length

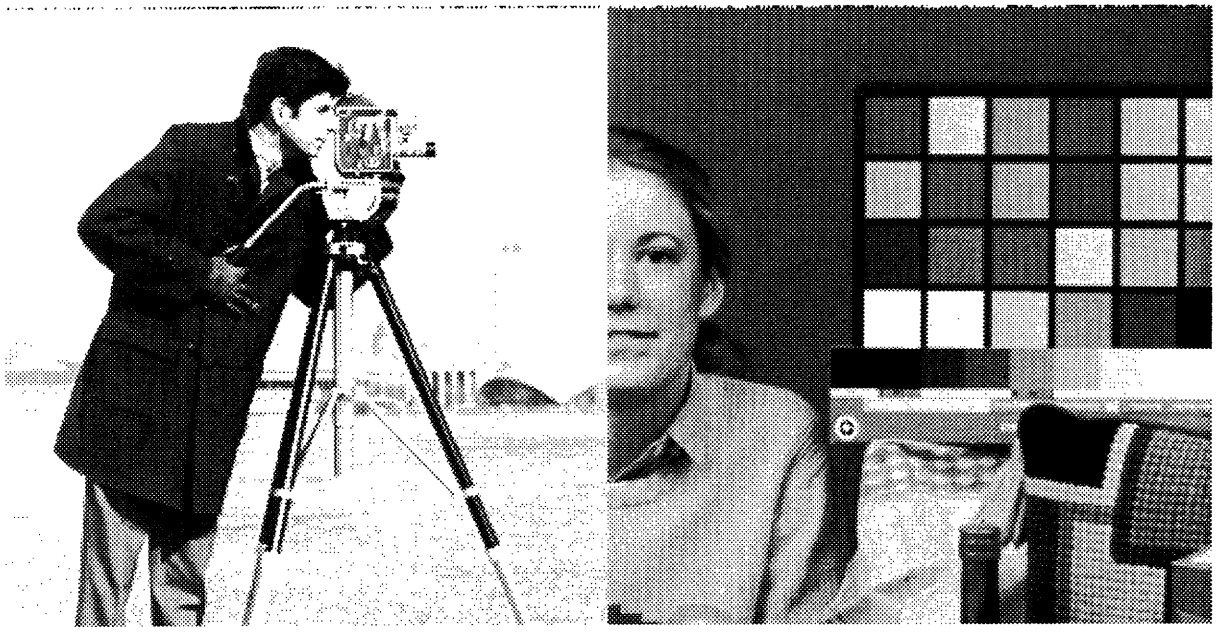


Figure 2.7: The original CMAN and GIRL pictures.

of the linear postfilter is much shorter and can be implemented in a cascade structure to reduce the required number of frame stores. Figure 2.9 shows the 2-D lowpass and highpass frequency components produced by this filter pair.

2.2 Calculating the Adaptation Factors

Now that we have settled upon a filter pair to perform the band separation, let us take a closer look at how the adaptive modulation is performed. Since one adaptation factor (or modulation index) is calculated for every 3-D block of image data, the sequence $p_s[m_1, m_2, m_3]$ can be considered as a type of subsampled, or decimated, sequence of the input highs, $i_h[n_1, n_2, n_3]$. The subsampled data can be written as,

$$p_s[m_1, m_2, m_3] = S[i_h[n_1, n_2, n_3]]. \tag{2.4}$$

The subsampled indices, $m_i = \lceil \frac{n_i}{N_i} \rceil$, denote the integer part of the fraction. The operator S calculates a single sample for an N_1xN_2xN_3 block of data in the horizontal, vertical,

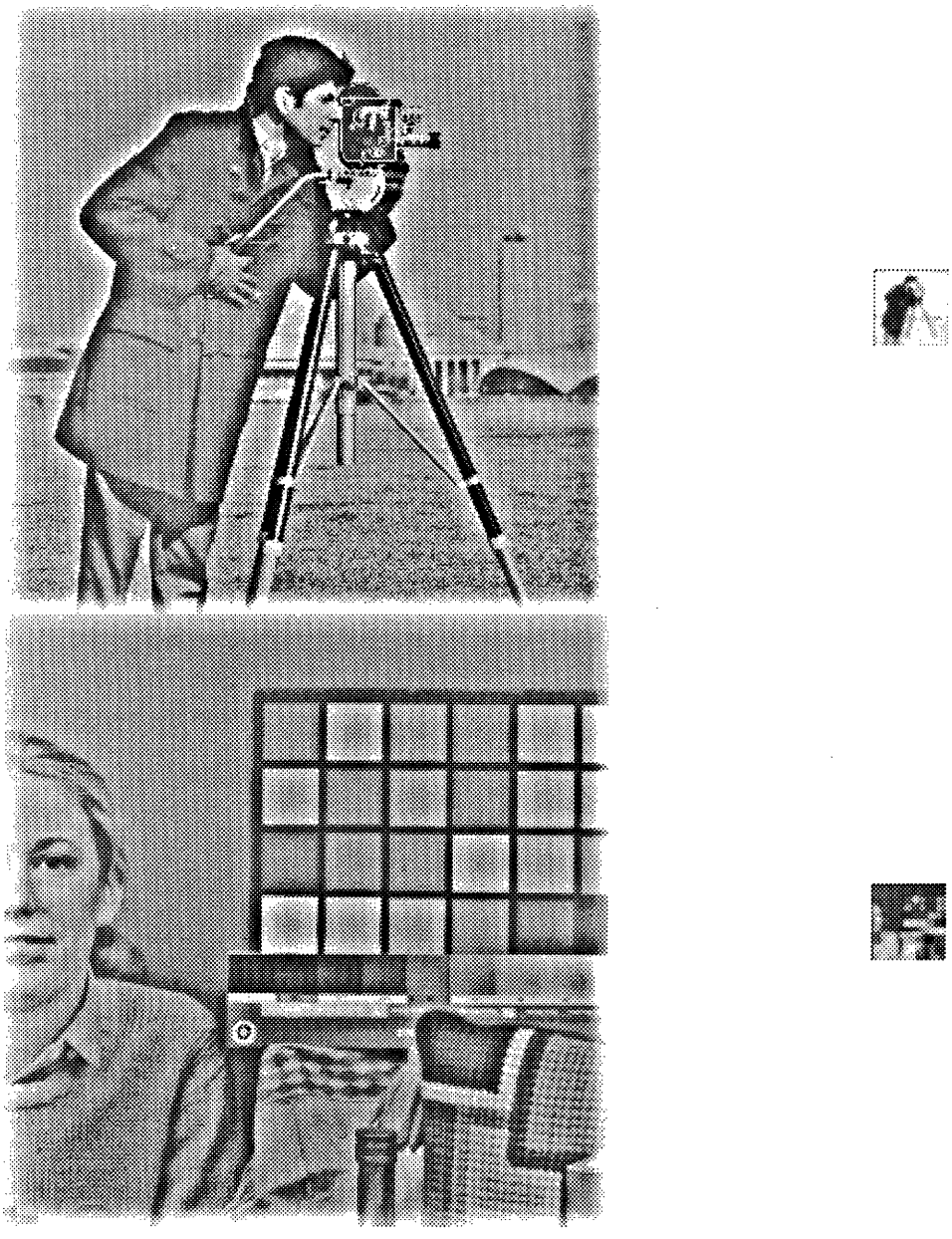


Figure 2.8: The highpass response of CMAN and GIRL using the final Gaussian filter-pair. The decimated lowpass images are also shown. The highpass components have an offset of 128.

Figure 2.9: The highpass and decimated lowpass components of CMAN using an averaging prefilter and a linear postfilter. Notice the aliasing. The highpass component has an offset of 128.

and temporal directions, respectively, and may include prefiltering prior to subsampling. The sequence, $p_s[m_1, m_2, m_3]$ is interpolated by N_1, N_2, and N_3 to form the sequence $p[n_1, n_2, n_3]$, and the adaptation factors are calculated as

$$m[n_1, n_2, n_3] = f(p[n_1, n_2, n_3]) \tag{2.5}$$

for some function f. The sequence of adaptation factors $m[n_1, n_2, n_3]$ is multiplied with $i_h[n_1, n_2, n_3]$ to form $y[n_1, n_2, n_3]$. This new sequence is expanded and then clipped to $\pm i_{h\max}$ if the signal exceeds this range, where $i_{h\max}$ is the absolute maximum value that $i_h[n_1, n_2, n_3]$ is assumed to attain. For image data with a dynamic range from 0 to 255, the largest absolute value that the highpass component can attain is 255, although values over 128 for normal images are rare. The maximum value, $i_{h\max}$, is chosen so as not to have significant distortion due to clipping but yet to have a sufficient blank area signal-to-noise ratio (SNR). Normally the maximum value is set to 128. Clipping $y[n_1, n_2, n_3]$ will produce distortion in the demodulated output, $\hat{i}[n_1, n_2, n_3]$, which will show up as a softening or blurring of the picture.

Now, let us introduce a set of adaptation factors for which the sequence $y[n_1, n_2, n_3]$ is clipped only when $i_h[n_1, n_2, n_3] > i_{h\max}$, and the adaptation factors are the largest they can be given this condition. This distinction is important because the clipping arises only from our choice of $i_{h\max}$ and not from subsampling or interpolation. This set of adaptation factors is

$$m_o[n_1, n_2, n_3] = \begin{cases} m_{\max} & \text{for } \frac{i_{h\max}}{|i_h[n_1, n_2, n_3]|} \geq m_{\max} \\ \frac{i_{h\max}}{|i_h[n_1, n_2, n_3]|} & \text{otherwise,} \end{cases} \tag{2.6}$$

where the subscript, o, indicates the "ideal" adaptation factors and $m_{\max}$ is the largest value we will let the adaptation factors attain. Multiplying the input data, $i_h[n_1, n_2, n_3]$, with $m_o[n_1, n_2, n_3]$ would use the channel most efficiently; however, the amount of side information would be as much as that required to send the input data itself. This is why we need to subsample the input data and interpolate it in calculating the adaptation factors. An example of an adaptive modulation algorithm is that proposed by Schreiber and Buckley [SB81]. The subsampling operation they use is

$$p_s[m_1, m_2, m_3] = S(i_h[n_1, n_2, n_3]) = | \; i_h[n_1, n_2, n_3] \; |\max \tag{2.7}$$

for $[n_1, n_2, n_3]$ which are elements of an N_1xN_2xN_3 block. The subsampled data, $p_s[m_1, m_2, m_3]$, are linearly interpolated with the sample located in the center of the block, and the adaptation factors are calculated according to

$$m[n_1, n_2, n_3] = \begin{cases} m\text{max} & \text{for } \frac{K}{|p[n_1,n_2,n_3]|} \geq m\text{max} \\ \frac{K}{|p[n_1,n_2,n_3]|} & \text{otherwise,} \end{cases} \tag{2.8}$$

where K is set to a value that causes negligible distortion in the worst case. This value is set to 60.

Because of subsampling and interpolation, the sequence of adaptation factors, $m[n_1, n_2, n_3]$, differs from the ideal adaptation factors, $m_o[n_1, n_2, n_3]$. In some instances, $m[n_1, n_2, n_3]$ will be greater than $m_o[n_1, n_2, n_3]$ and in others, less than $m_o[n_1, n_2, n_3]$. In those cases where $m[n_1, n_2, n_3] > m_o[n_1, n_2, n_3]$ (i.e., where $\mid p[n_1, n_2, n_3] \mid < \mid i_h[n_1, n_2, n_3] \mid$), the absolute value of the product of $m[n_1, n_2, n_3]$ and $i_h[n_1, n_2, n_3]$ will be greater than $i_h\text{max}$, and distortion will occur in the demodulated output due to the subsampling and interpolation operations. Having $m[n_1, n_2, n_3] > m_o[n_1, n_2, n_3]$ is not entirely undesirable; in fact, it should be weighed against the increase in noise-reducing capability that having a larger $m[n_1, n_2, n_3]$ over the ideal case will produce. Having $m[n_1, n_2, n_3] > m_o[n_1, n_2, n_3]$ for certain picture elements will tend to raise the value of adaptation factors of neighboring picture elements without distorting those picture elements. For example, in regions near an edge, some picture elements will have adaptation factors larger, and some smaller, than the ideal adaptation factors due to the interpolation . By allowing more distortion of the picture elements with larger adaptation factors, the noise-reducing capabilities of picture elements with smaller adaptation factors near edges is increased without distorting those picture elements. If the distortion is not noticeable, then we have increased the overall noise-reducing capability without visibly degrading the demodulated output image, $\hat{i}[n_1, n_2, n_3]$. Little work has been done in finding this threshold of tolerable distortion or in using different types of subsampling operations or interpolation methods.

Before we investigate the subsampling (decimation) and interpolation processes, it is instructive to look at the signal as it evolves through the adaptive modulation process for the adaptive modulation algorithm used by Schreiber and Buckley. Figure 2.10 shows various signals of the adaptive modulation process for line 51 of CMAN. The band

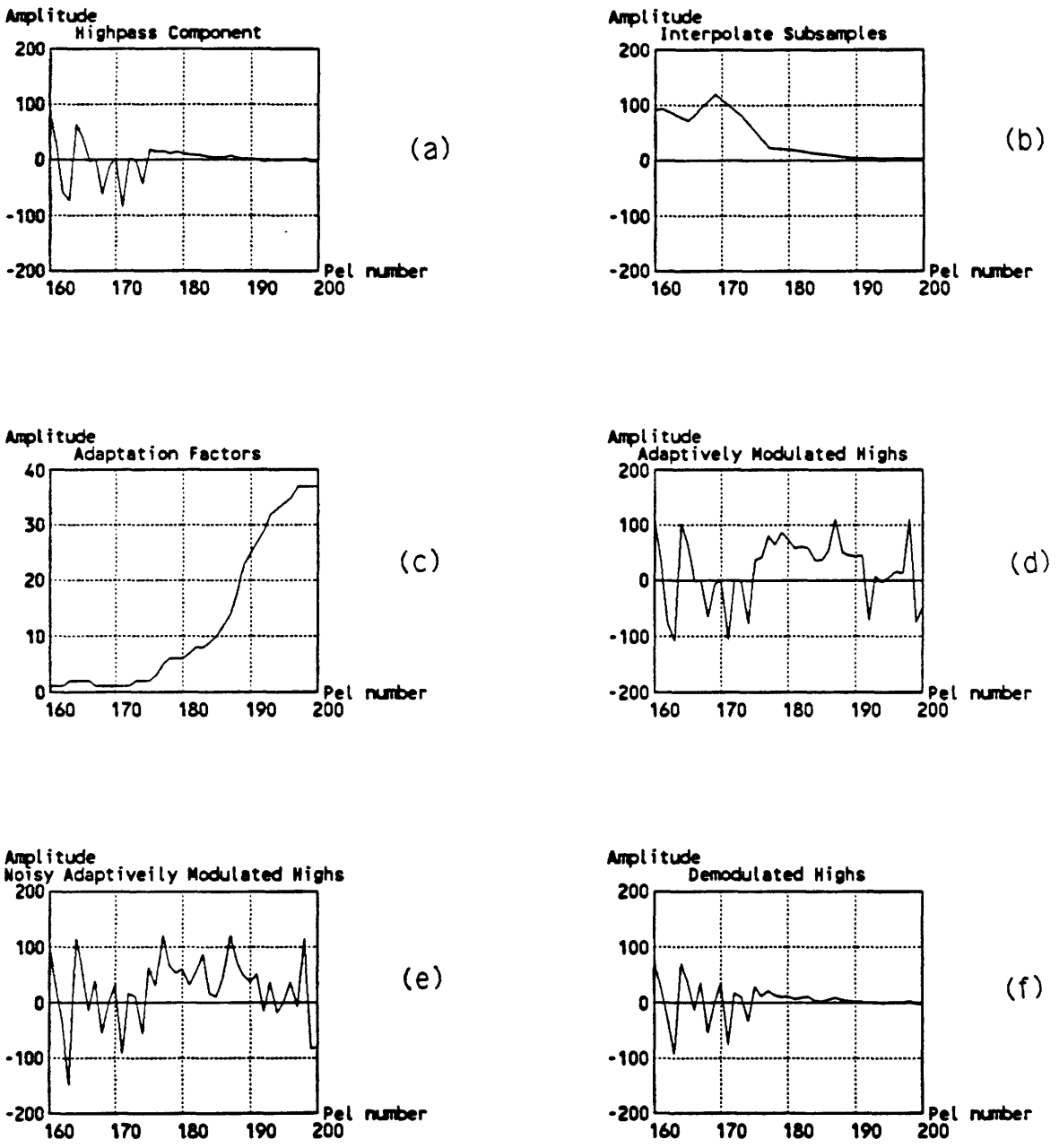


Figure 2.10: Evolution of signals in the adaptive modulation process.

separation is performed with the filter pair from the previous section, where the lowpass signal is decimated by 8 in both the vertical and horizontal directions. No filtering is performed in the temporal direction. Schreiber and Buckley subsample the highpass component, $i_h[n_1, n_2, n_3]$ (Figure 2.10(a)), by finding the maximum of the absolute values in a 4x4 spatial block of the highpass signal and interpolate using a bilinear filter to produce $p[n_1, n_2, n_3]$ (Figure 2.10(b)). The adaptation factors, $m[n_1, n_2, n_3]$, are calculated according to eqn. 2.8 with $K = 128$ and are shown in Figure 2.10(c). Here, we have allowed the maximum adaptation factor to be 128. Multiplying the highpass component, $i_h[n_1, n_2, n_3]$, with the adaptation factors, $m[n_1, n_2, n_3]$, produces the adaptively modulated signal, $y[n_1, n_2, n_3]$, as shown in Figure 2.10(d). If the noise in the channel has a 20dB CNR, then the signal in Figure 2.10(e) results. The adaptively demodulated signal, $\hat{i}_h[n_1, n_2, n_3]$, is shown in Figure 2.10(f). The entire highpass component of CMAN, its adaptation factors, and the adaptively modulated highpass component are shown in Figure 2.11, and the demodulated pictures which result when no adaptive modulation is used and when adaptive modulation is used over a 20dB CNR channel are shown in Figure 2.12.

How does adaptive modulation change the characteristics of the highpass signal? Three characteristics of interest are the histograms, the power spectra, and the energies of the signals. The histogram of the highs component is shown in Figure 2.13(a). As one would expect, most picture elements have very small values. Adaptive modulation, in its ideal form, should raise the value of each pel, except 0, to $i_{h\max} = 128$ but is limited by the value of the maximum adaptation factor, which we have set to 128. Figure 2.13(b) shows the histogram of the adaptively modulated highs. Figure 2.14 shows the power spectra of the highs component and of the adaptively modulated highs component. The standard deviation of the highpass component is 43.2, and the standard deviation of the adaptively modulated highpass component is 69.2. For comparison, the standard deviation of the original, fullband CMAN is 68.6. Although adaptive modulation increases the energy of the highpass component, it does not change the peak power of the signal.

During all of our experiments on adaptive modulation, we set the length of any one side of the 3-D block to be one half the decimation factor used when decomposing the

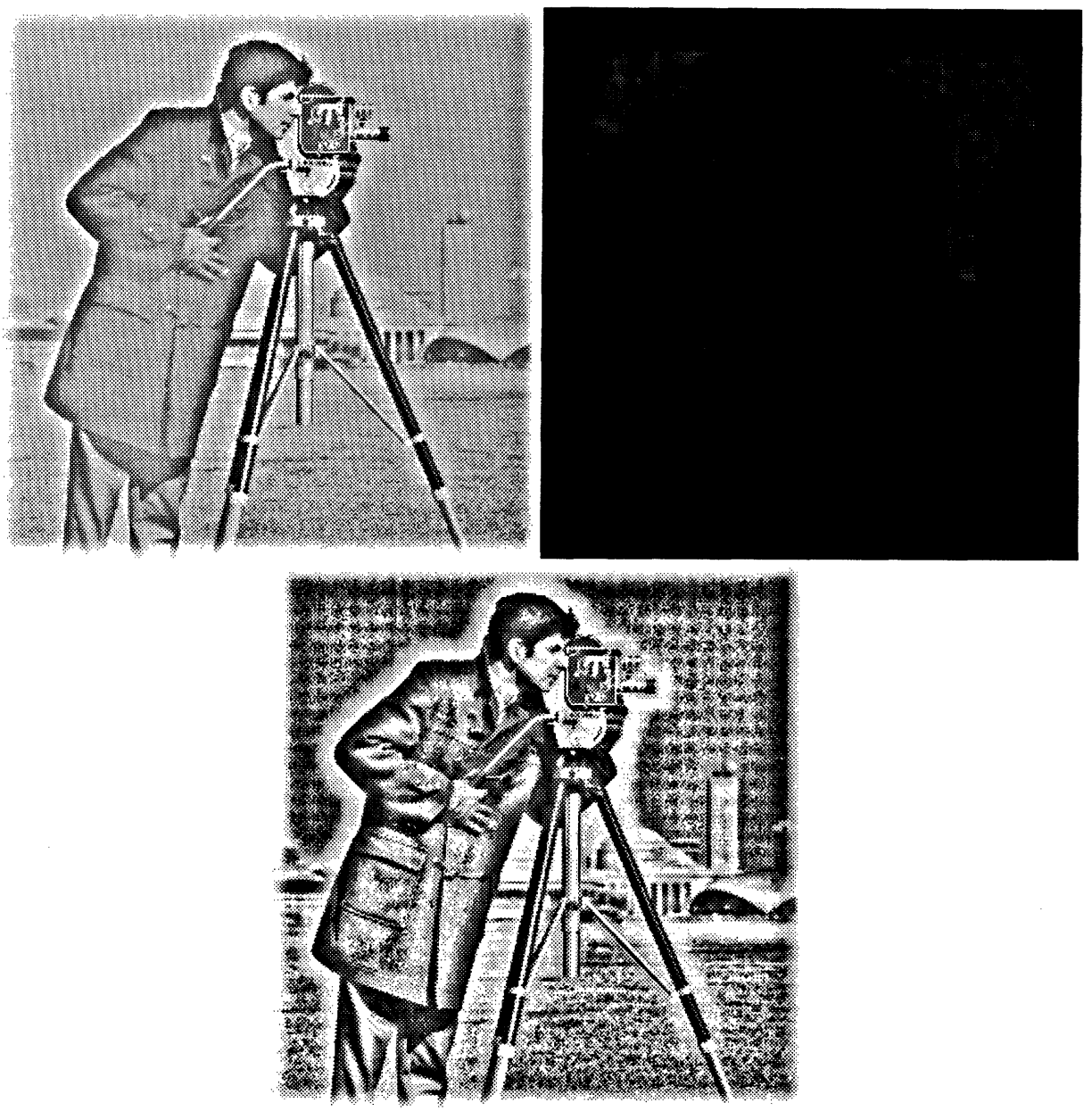


Figure 2.11: Highpass component, adaptation factors, and adaptively modulated highpass component of CMAN. An offset of 128 has been added to the highpass component and to the adaptively modulated highpass component. The adaptation factors have been scaled by 4.

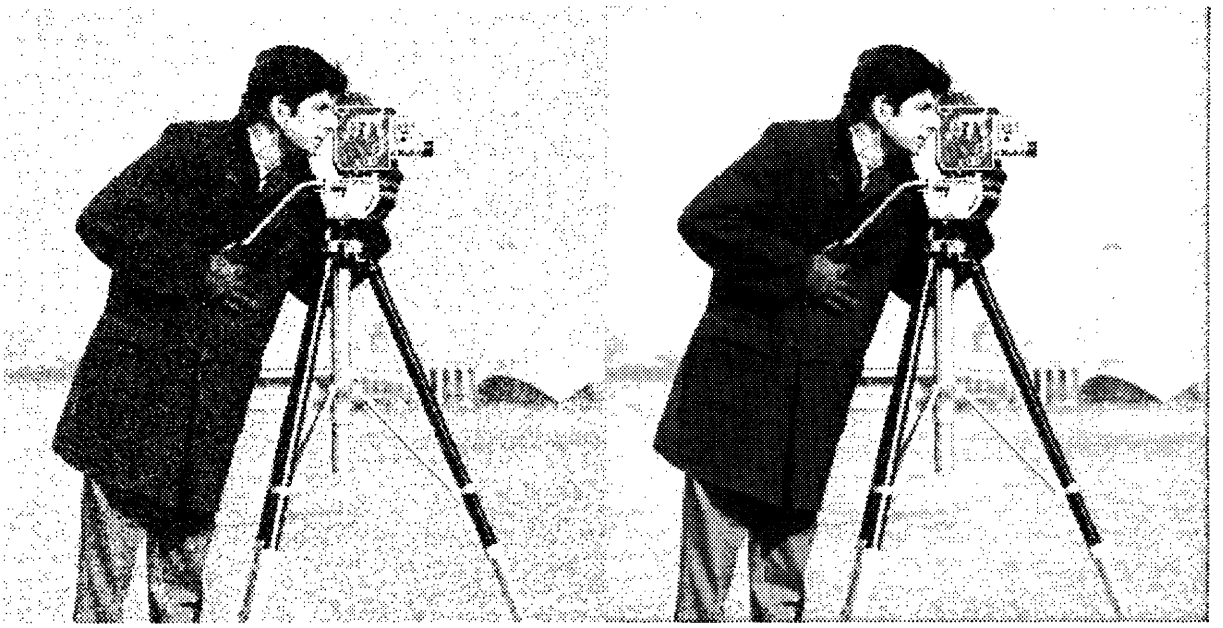


Figure 2.12: Demodulated signals when no adaptive modulation is used and when adaptive modulation is used when transmitting over a 20dB CNR channel.

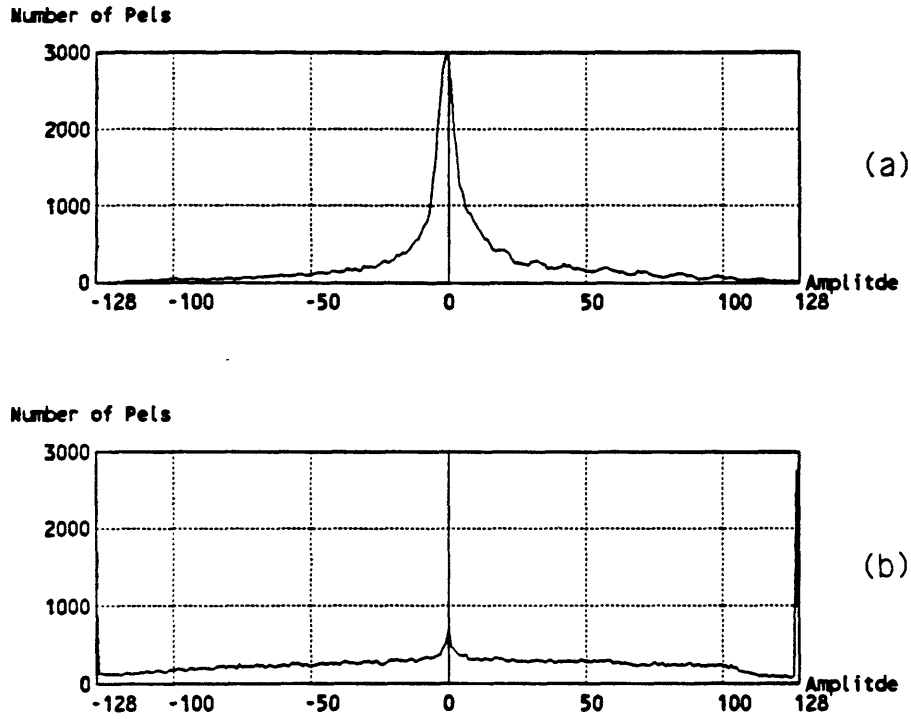


Figure 2.13: Histograms of the highpass (a) and adaptively modulated highpass (b) components of CMAN.

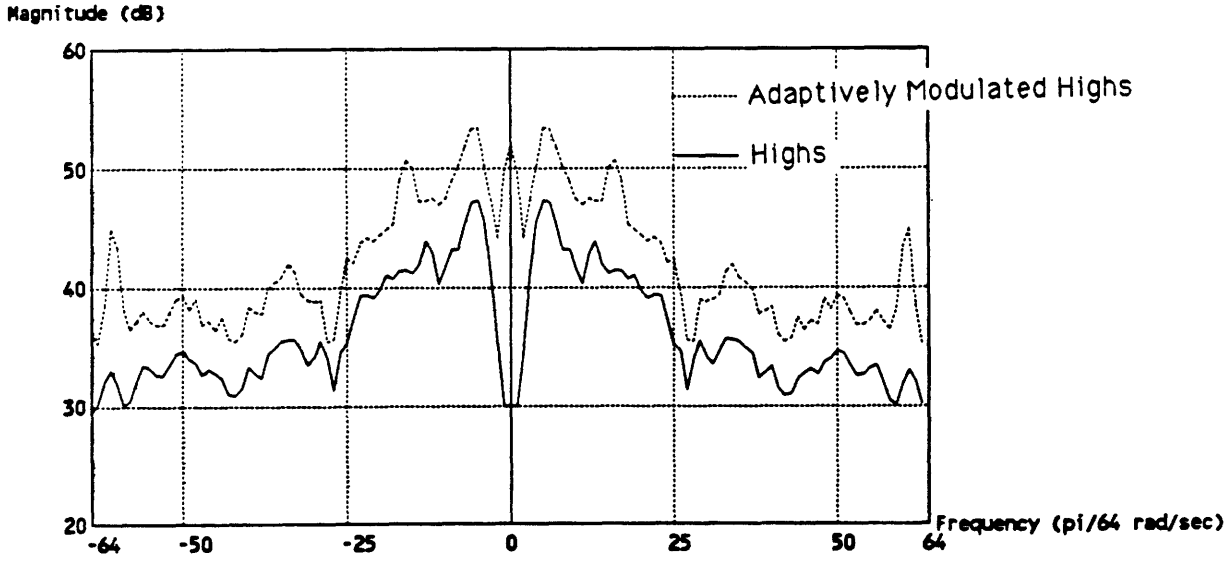


Figure 2.14: The horizontal frequency axis of the power spectra of the highpass and adaptively modulated highpass components of CMAN.

input signal into lowpass and highpass components. For a 512x512, 60 fps sequence the decimation factor is 8 in each dimension; therefore, our block size will be 4x4x4. The rate of decay of the interpolated subsamples, $p[n_1, n_2, n_3]$, determines the block size. When the peak value of the highpass component within a block occurs in the middle of the block, then, depending upon what interpolation method is used, the effect of the peak upon the rate of decay of the interpolated subsamples can extend up to twice the length of the block along each dimension. Since we would like the rate of decay of the interpolated subsamples to coincide with the effects of masking, we choose the length of each side of the block to be one half the decimation factor used during band separation. In addition, we use the function described in eqn. 2.8 with $K = i_{h\max} = 128$ to calculate the adaptation factors.

2.3 Origins of Distortion

Before we embark on the discussion of our experiments on subsampling, interpolation, and distortion, let us explore the origins of distortion. The transmitter must clip the adaptively modulated signal prior to transmission in order to conform to the peak transmission power constraint for broadcasting television signals; therefore, distortion occurs when the adaptively modulated highpass component, $y[n_1, n_2, n_3]$, is above $i_{h\max}$, which happens either when the highpass signal, $i_h[n_1, n_2, n_3]$, is above $i_{h\max}$ or when the adaptation factors, $m[n_1, n_2, n_3]$, are larger than the ideal adaptation factors, $m_o[n_1, n_2, n_3]$. This latter condition occurs when the absolute value of the interpolated subsamples, $\mid p[n_1, n_2, n_3] \mid$, are smaller than the corresponding absolute value of the highpass samples, $\mid i_h[n_1, n_2, n_3] \mid$. Without loss of generality, let us consider only positive values and suppose $p[n_1, n_2, n_3] < i_h[n_1, n_2, n_3]$ and

$$p[n_1, n_2, n_3] = i_h[n_1, n_2, n_3] - \Delta i_h[n_1, n_2, n_3] \tag{2.9}$$

for some n_1, n_2, and n_3, and where $\Delta i_h[n_1, n_2, n_3] > 0$. Under these conditions, distortion occurs. The adaptation factor is

$$m[n_1, n_2, n_3] = \frac{128}{p[n_1, n_2, n_3]}$$

$$= \frac{128}{i_h[n_1,n_2,n_3] - \Delta i_h[n_1,n_2,n_3]} \tag{2.10}$$

where we assume that $1 \leq p[n_1,n_2,n_3] \leq 128$. The adaptively modulated sample is

$$\begin{aligned} y[n_1,n_2,n_3] &= i_h[n_1,n_2,n_3]m[n_1,n_2,n_3] \\ &= i_h[n_1,n_2,n_3](\frac{128}{i_h[n_1,n_2,n_3]} + \frac{128\,\Delta i_h[n_1,n_2,n_3]}{i_h[n_1,n_2,n_3](i_h[n_1,n_2,n_3] - \Delta i_h[n_1,n_2,n_3])}) \\ &= 128 + \frac{128\,\Delta i_h[n_1,n_2,n_3]}{i_h[n_1,n_2,n_3] - \Delta i_h[n_1,n_2,n_3]}, \end{aligned} \tag{2.11}$$

which is > 128. The second term is the amount of the signal which gets "clipped off."

The received sample is

$$\hat{y}[n_1,n_2,n_3] = 128 + v[n_1,n_2,n_3] \tag{2.12}$$

where $v[n_1,n_2,n_3]$ is additive noise. The demodulated sample is

$$\begin{aligned} i_h[n_1,n_2,n_3] &= \frac{\hat{y}[n_1,n_2,n_3]}{m[n_1,n_2,n_3]} \\ &= \frac{128 + v[n_1,n_2,n_3]}{128}(i_h[n_1,n_2,n_3] - \Delta i_h[n_1,n_2,n_3]) \\ &= (i_h[n_1,n_2,n_3] - \Delta i_h[n_1,n_2,n_3]) \\ &\quad + \frac{v[n_1,n_2,n_3](i_h[n_1,n_2,n_3] - \Delta i_h[n_1,n_2,n_3])}{128}. \end{aligned} \tag{2.13}$$

The first term is the distorted highpass signal and the second term is the reduced noise. The distortion is $\Delta i_h[n_1,n_2,n_3]$, which is just the difference between the highpass sample, $i_h[n_1,n_2,n_3]$, and the interpolated subsample, $p[n_1,n_2,n_3]$, when $\mid p[n_1,n_2,n_3] \mid < \mid i_h[n_1,n_2,n_3] \mid$. When $\mid p[n_1,n_2,n_3] \mid \geq \mid i_h[n_1,n_2,n_3] \mid$, no distortion occurs. This is shown pictorially in Figure 2.15. To increase noise reduction, we want $\Delta i_h[n_1,n_2,n_3]$ to be as large as possible; however, to reduce distortion, we want $\Delta i_h[n_1,n_2,n_3]$ to be as small as possible.

2.4 Decimation and Interpolation of the Highpass Component

The existing adaptive modulation algorithm, which was put forth by Schreiber and Buckley, allows negligible distortion due to clipping to occur. Under bad (< 25dB CNR)

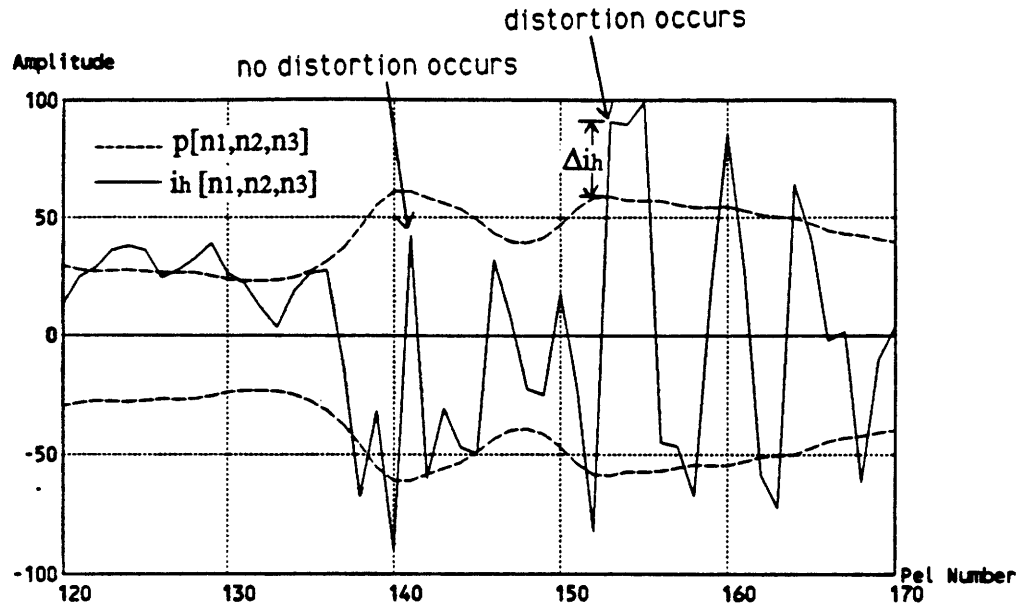


Figure 2.15: Distortion occurs when the absolute value of the interpolated subsamples is less than the absolute value of the highpass samples.

noise conditions, this algorithm does not sufficiently suppress the noise around edges (see Figure 2.12.) Two ways to improve the performance of adaptive modulation is to increase the rate of decay of the interpolated subsamples by using a sharpened interpolating filter and to allow more distortion to occur.

In our first set of experiments, we subsampled by finding the maximum of the absolute values of the highpass component in a block (Schreiber and Buckley's method) and interpolated with a sharpened-Gaussian filter. Sharpened-Gaussian filters with standard deviations from 1.5 to 2.5 were tried (see [Rat80] for formula of a sharpened-Gaussian filter). No improvement over using a bilinear interpolating filter was found. In fact, the pictures looked worse because using sharpened Gaussian filters resulted in large overshoots and sampling structure, which manifested themselves as significant distortion in the demodulated output. Using the nonlinear mechanism of finding the maximum absolute value in a block as the subsampling operation does not sufficiently bandlimit the signal, so that significant aliasing occurs when interpolating with a sharpened filter. The aliasing shows up as strong ringing. One way to circumvent this problem is to

use a subsampling operation that better bandlimits the signal. Buckley also performed experiments where he varied the interpolating filter; however, none of his filters were sharpening filters. He similarly found that when using the maximum absolute value in a block as the subsampling operation, a linear interpolating filter gives the best looking pictures.

Prefiltering prior to using the conventional subsampling operation,

$$p_s[m_1, m_2, m_3] = S(i_h[n_1, n_2, n_3]) = i_h[m_1 N_1, m_2 N_2, m_3 N_3] \tag{2.14}$$

where $N_1 \text{x} N_2 \text{x} N_3$ is the block size, should allow the use of sharpened-Gaussian filters to increase the decay rates of the interpolated subsamples without the large ringing and overshoots. We limit ourselves to Gaussian prefilters and decimate and interpolate only the absolute value of the highpass component to find $p[n_1, n_2, n_3]$. Using the maximum absolute value in a block and linear interpolating produces negligible distortion; whereas, using conventional decimation and interpolation operations can produce significant amounts of distortion. This in itself is not bad. To reduce the distortion one only needs to scale $p[n_1, n_2, n_3]$ by some factor to produce any desired level of distortion. This is done in latter experiments. What is important is that the $p[n_1, n_2, n_3]$ that results when using Gaussian prefilters and sharpened-Gaussian postfilters somehow better follow the fluctuations of $i_h[n_1, n_2, n_3]$ compared to the $p[n_1, n_2, n_3]$ that result when using linear interpolation of the maximum absolute values. This statement is hard to quantify, and a good filter-pair can only be found by subjective judgment.

The prefilter controls the amount of high-frequency information that gets passed. Too small a standard deviation of the Gaussian prefilter will allow too much aliasing and too much overshoot after interpolation. Too large a standard deviation of the prefilter will excessively blur the edges and will decrease the rate of decay of the impulse response of the interpolated subsamples. The postfilter controls the band separation properties, visibility of sampling structure, and the sharpness of edges. A large standard deviation of the postfilter will produce fast transitions and large overshoots, which are undesirable because they produce a "halo" of noise around sharp edges. Small standard deviations of the postfilter have poor band-separation properties, and sampling structure becomes

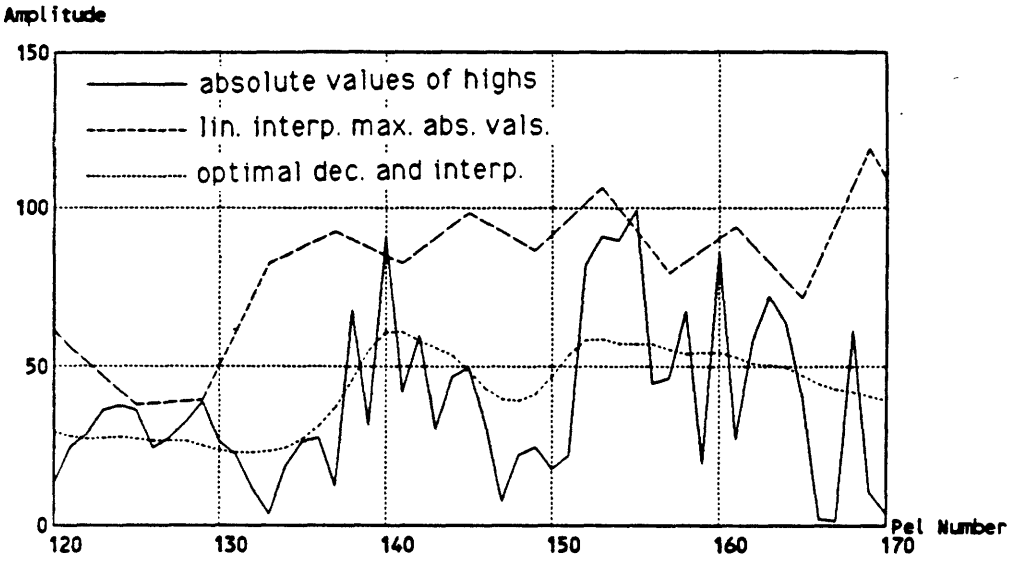


Figure 2.16: Two different methods of subsampling and interpolation: optimal decimation-and-interpolation and maximum-linear.

noticeable.

Using a block size of 4x4x4, we varied the standard deviation of the prefilter from 1.0 to 5.0 and varied the standard deviation of the postfilter form 1.0 to 3.0. The filter-pair which produces the most pleasing result had a prefilter standard deviation of 1.5 and a postfilter standard deviation of 2.0 with sharpening factor of 0.135. We shall call this the optimal decimation-and-interpolation method for calculating the adaptation factors. For one section of line 51 of CMAN, the decimated-and-interpolated samples generated using this filter-pair are shown in Figure 2.16 along with the absolute values of the highpass component and the interpolated subsamples obtained from linearly interpolating the maximum absolute values in the blocks.

This set of Gaussian interpolated subsamples can be scaled by a factor to obtain any desired amount of distortion. The same scaled subsamples, $p[n_1, n_2, n_3]$, are used at both the transmitter and the receiver. For example, the interpolated subsamples can be scaled by 1.6 so that the amount of visible distortion is comparable to that produced by linearly interpolating the maximum absolute values. Using this scaled signal

to calculate the adaptation factors approximately produces a 1.0 dB SNR improvement when compared to using the linearly interpolated maximum absolute values to calculate the adaptation factors. We can also scale the linearly interpolated maximum absolute values to vary the amount of distortion. If we scale both the optimal decimation-and-interpolation subsamples and the linearly interpolated maximum absolute values (i.e., in order to produce other comparable levels of distortion), we find that this gain in SNR is consistent.

One useful application of varying the scaling factor of the interpolated subsamples is to match the distortion level to the anticipated CNR of the channel in order to produce a picture of the best subjective quality at the receiver. We carried out subjective tests to measure the subjective picture quality for various combinations of scaling factors and CNR levels. Ten subjects, experienced in viewing pictures, sat at 4 picture heights to judge the images by rating the pictures on a scale of 1 to 10 where 10 indicated excellent picture quality. The results are summarized in Table 2.2, where the mean quality rating is shown. In the last row is the subjective picture quality if no noise reduction is performed. For each CNR level, there is a maximum in subjective picture quality (the circled values) that indicates the optimal tradeoff between distortion and noise reduction. A smaller scaling factor means that more distortion occurs, where a scaling factor of 1.6 produces negligible distortion. Plotted in Figure 2.17 is the optimal scaling factor versus CNR. This graph tells us what scaling factor (or distortion level) will produce the best looking images (in terms of subjective picture quality) for a given CNR. At a 20dB CNR, the optimal scaling factor is 1.2, and the demodulated pictures when using this scaling factor of 1.2 under noiseless and noisy conditions are shown in Figure 2.18.

2.5 Subsampling and Distortion Level

One sample of data, $p_s[m_1, m_2, m_3]$, is transmitted for each block of highpass data, $i_h[n_1, n_2, n_3]$. We call this the subsampling operation. In the previous section, we limited ourselves to using the conventional subsampling operation of eqn. 2.14 to calculate the transmitted subsamples, $p_s[m_1, m_2, m_3]$. The value of the subsample was determined

Scaling Factor	Signal-to-Noise Ratio in Decibels					
	40 dB	30 dB	25 dB	20 dB	15 dB	10 dB
1.6	(9.0)	(9.0)	7.6	5.7	4.3	3.1
1.4	8.6	8.4	(7.9)	6.4	4.9	3.4
1.2	7.4	7.6	7.7	(6.4)	(5.3)	3.6
1.0	6.1	6.0	6.4	5.7	4.9	(4.0)
0.8	5.3	5.4	5.3	4.6	4.6	3.9
0.6	3.6	3.3	3.3	3.1	3.0	2.9
0.4	2.1	2.0	1.9	1.9	1.7	1.7
Just Noise	8.9	6.1	5.0	4.0	2.7	1.7

Table 2.2: Table of the results of the subjective test. The numbers represent subjective picture quality ratings, and the circled numbers are the maximum in their column.

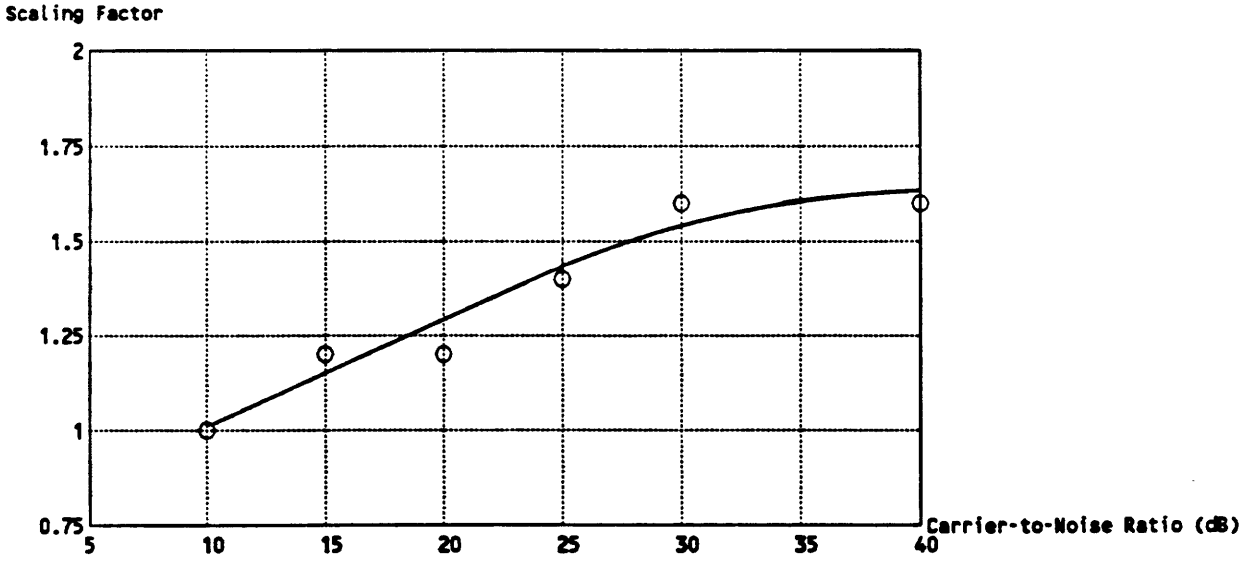


Figure 2.17: Optimal scaling factor versus CNR.

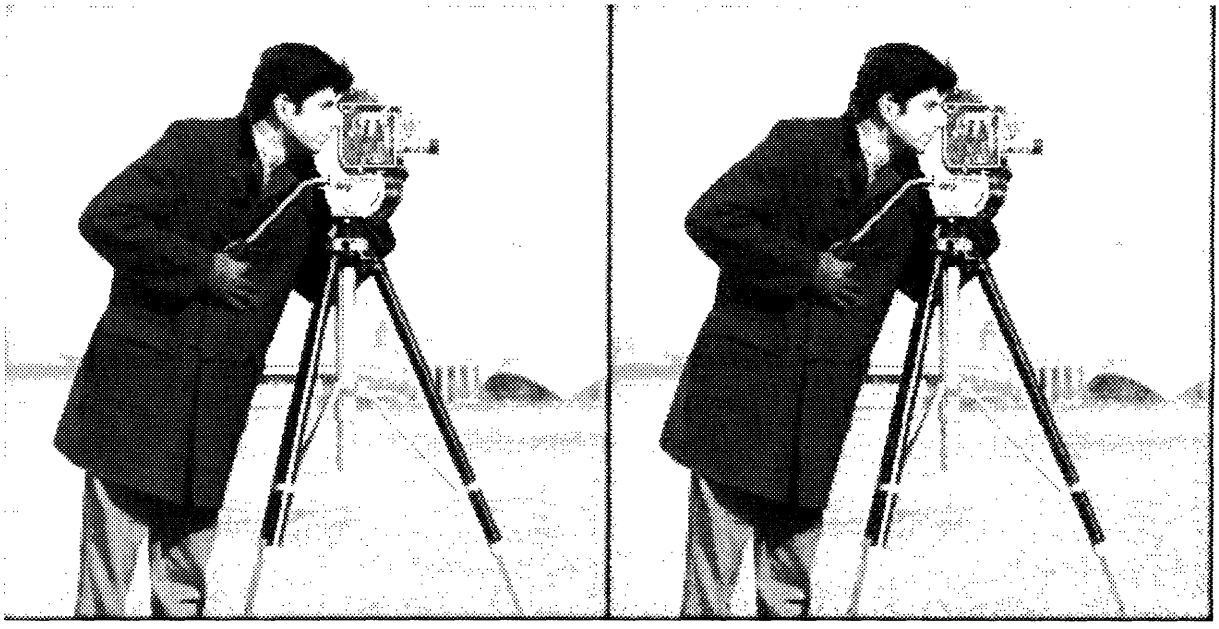


Figure 2.18: Demodulated pictures using the optimal decimation-and-interpolation method of calculating the adaptation factors with a scaling factor of 1.2 under noiseless and noisy (20dB CNR) channel conditions.

solely by the prefiltering operation. In this section, we try to determine the "best" value of the subsample to transmit. This criterion of "best" is based upon trying to maximize noise reduction near edges without causing perceptible distortion of the edges. We shall attack this problem by starting with a set of subsamples, $p_s^o[m_1, m_2, m_3]$, which cause negligible distortion (i.e., those obtained by finding the maximum absolute value in a block of highpass data) and then doing an iterative search over all blocks to find the value of each subsample that produces a specified distortion level (in percent) in each block. Various distortion measures will be tried, and for each measure, the distortion level which results in imperceptible to just perceptible distortion will be found. Once the distortion level for each measure is determined, that measure which provides the greatest noise reduction will be selected.

2.5.1 The Search Method

The distortion in one block is a function of its own subsample and the subsamples of its neighboring blocks. The dependence upon its neighboring subsamples is determined by the interpolation process used. To find the set of subsamples, $p_s[m_1, m_2, m_3]$, which produce the desired distortion level in each block, one can use Newton's method of finding the zeros of a function [Str86]. For a function, f, of one variable, p_s, we find the solution of $f(p_s) = T$ by iteratively solving

$$f'(p_s^k)(p_s^{k+1} - p_s^k) = T - f(p_s^k) \tag{2.15}$$

for p_s^{k+1}, where p_s^k is the current approximation and f' is the derivative of f. For a convex function, f, the $\{p_s^k\}$ are guaranteed to converge to the solution. For our purposes, a direct implementation of Newton's method requires us to solve the set of n linear equations

$$\mathbf{J}^k(\mathbf{p}_s^{k+1} - \mathbf{p}_s^k) = \mathbf{T} - \mathbf{f}^k(\mathbf{p}_s^k) \tag{2.16}$$

corresponding to the n blocks, where $\mathbf{p}_s^k$ is the n-vector of all subsamples in a set of frames, $\mathbf{f}^k$ is an n-vector of the corresponding block distortion evaluated at $\mathbf{p}_s^k$, and $\mathbf{J}^k$ is the nxn Jacobian matrix evaluated at $\mathbf{p}_s^k$ and contains the first partial derivatives $\partial f_i / \partial p_{s_i}$ of the n distortion functions f_i with respect to the n subsamples p_{s_i}. The nondiagonal,

nonzero entries in the Jacobian arise because of the dependency of the distortion in a block upon the subsamples of its neighbors. Solving this set of linear equations requires a matrix inversion, which can be a problem if the Jacobian is ill conditioned. One way to circumvent this problem is to use the idea behind minimization by coordinate descent [Lue84] [AO82].

Coordinate descent algorithms minimize a function of several variables by successively minimizing the function with respect to one variable at a time and then repeating the process until the minimum is found. If we were to apply this idea to the problem at hand, then we would hold all subsamples constant except for one, and, using Newton's method for one variable, solve for the value which would produce the desired distortion level in that block. This process would be done for all subsamples, and the process would be repeated until the distortion in each of the blocks had attained the desired level. This method does not require computation of the Jacobian or its inversion; however, this method needs n times more steps to converge!

We can eliminate the matrix inversion and the increased convergence time by ignoring the dependence of the distortion in a block on the subsamples of the neighboring blocks. This makes the Jacobian a diagonal matrix and allows us to perform n independent, single-variable searches. Since the distortion function in one block should be very similar to its neighboring blocks, we expect the solution to be very close to that obtained from a direct implementation of Newton's method; however, oscillations in the search may occur.

The single-variable search method that we use is a modified version of Newton's method called the Secant method [Avr76]. In the Secant method, the derivative of f is approximated by a quotient of differences, so that we find the solution to $f(p_s) = T$ by iteratively solving

$$\frac{(f(p_s^k) - f(p_s^{k-1}))}{(p_s^k - p_s^{k-1})}(p_s^{k+1} - p_s^k) = T - f(p_s^k) \tag{2.17}$$

for p_s^{k+1}. This method is shown pictorially in Figure 2.19. As with Newton's method, the Secant method converges if f is a convex function. The distortion for a group of horizontally adjacent blocks as a function of the decrease in the subsample is shown in Figure 2.20. We see that the distortion is essentially a convex, monotonically increasing

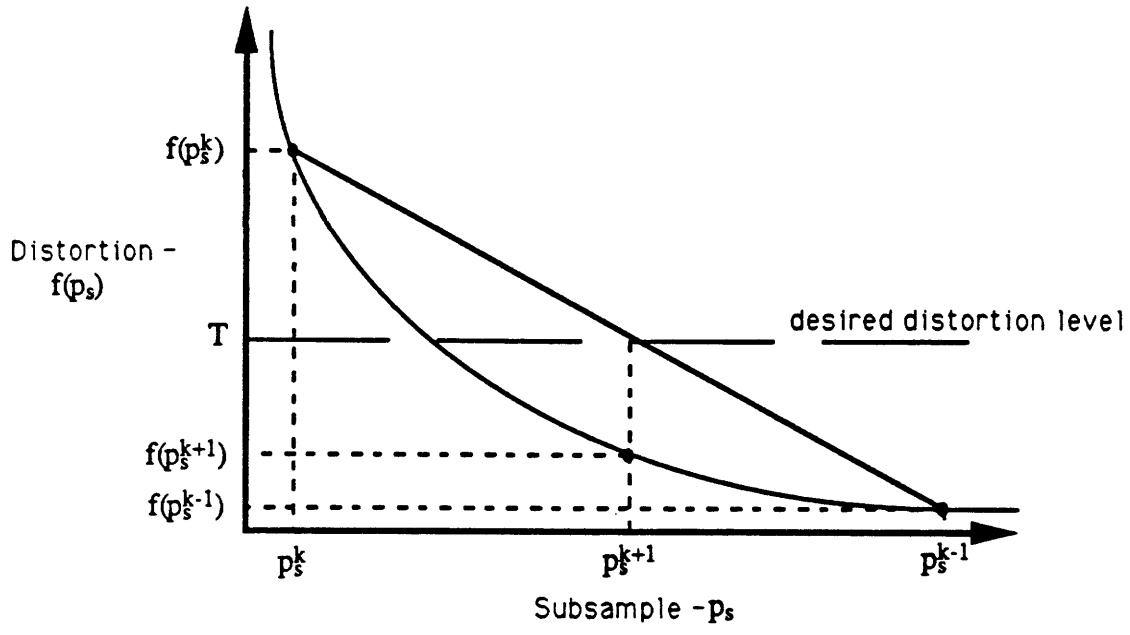


Figure 2.19: The Secant Method.

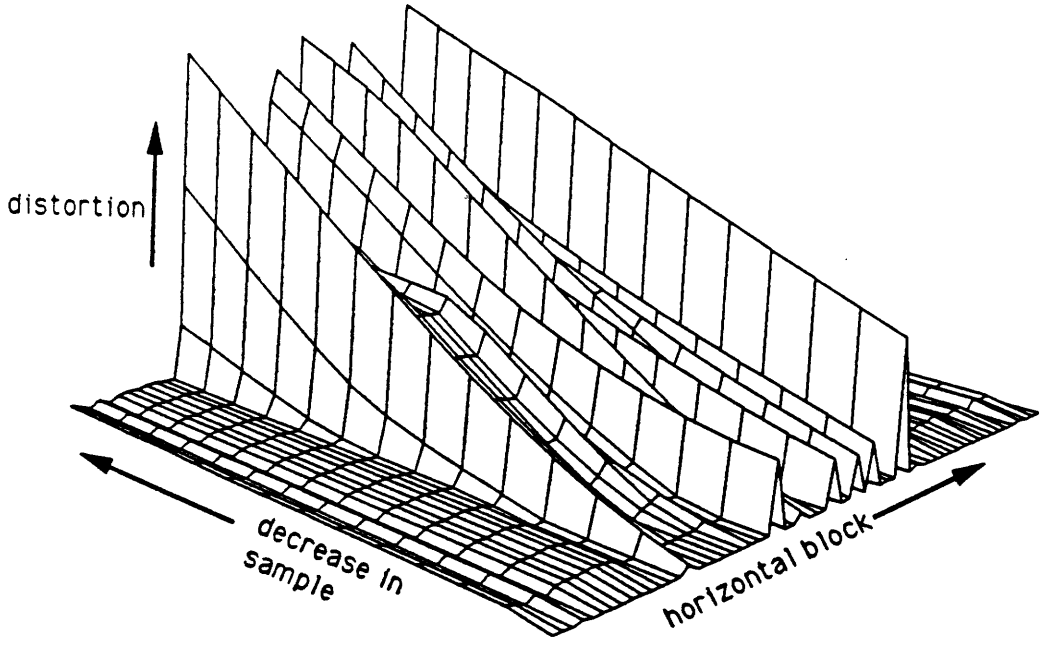


Figure 2.20: Plot of block distortion versus decrease in the value of the associated sample of that block for many blocks along a horizontal line.

function for the region of distortion of interest to us. The rate of convergence of the Secant method is somewhat slower that the quadratic rate of convergence of Newton's method, but is faster than linear convergence.

The iterative algorithm that we use is shown in Figure 2.21. Our initial guess of the $p_s[m_1, m_2, m_3]$ is obtained from finding the maximum absolute values in the blocks. The Secant search method requires an initialization step that decreases the subsamples by a large amount so that the resulting distortion is much larger that the desired level of distortion. The distortion function, f, may have places of zero or positive slope. If this happens, we force the stepsize to be a specified value (usually 2). The iterative process continues until all blocks are within $\epsilon = 10\%$ of the desired level of distortion or until the subsample is 0.

Figure 2.22 shows the demodulated image, $\hat{i}[n_1, n_2, n_3]$, the distortion in the demodulated image, $i - \hat{i}$, and the histogram of the number of blocks that are at a certain percent of the desired distortion level at five iteration numbers. The distortion measure is the root-mean-squared-error distortion measure (explained in the next section) and the distortion level is 15%. The amount of distortion is very large during initialization (iteration 0) and decreases on successive iterations. After five iterations, the distortion in the blocks has clustered around the desired distortion level. After five iterations, we assume that the algorithm has "converged."

2.5.2 The Distortion Measure and Level

We consider three distortion measures, the root-mean-squared-error

$$d_{rmse} = \sqrt{\sum_{n_1,n_2,n_3} \frac{e^2[n_1, n_2, n_3]}{N_1 N_2 N_3}}, \tag{2.18}$$

the peak-absolute-error

$$d_{peak} = \mid e[n_1, n_2, n_3] \mid_{max}, \tag{2.19}$$

and the mean-absolute-error

$$d_{mean} = \sum_{n_1,n_2,n_3} \frac{\mid e[n_1, n_2, n_3] \mid}{N_1 N_2 N_3}, \tag{2.20}$$

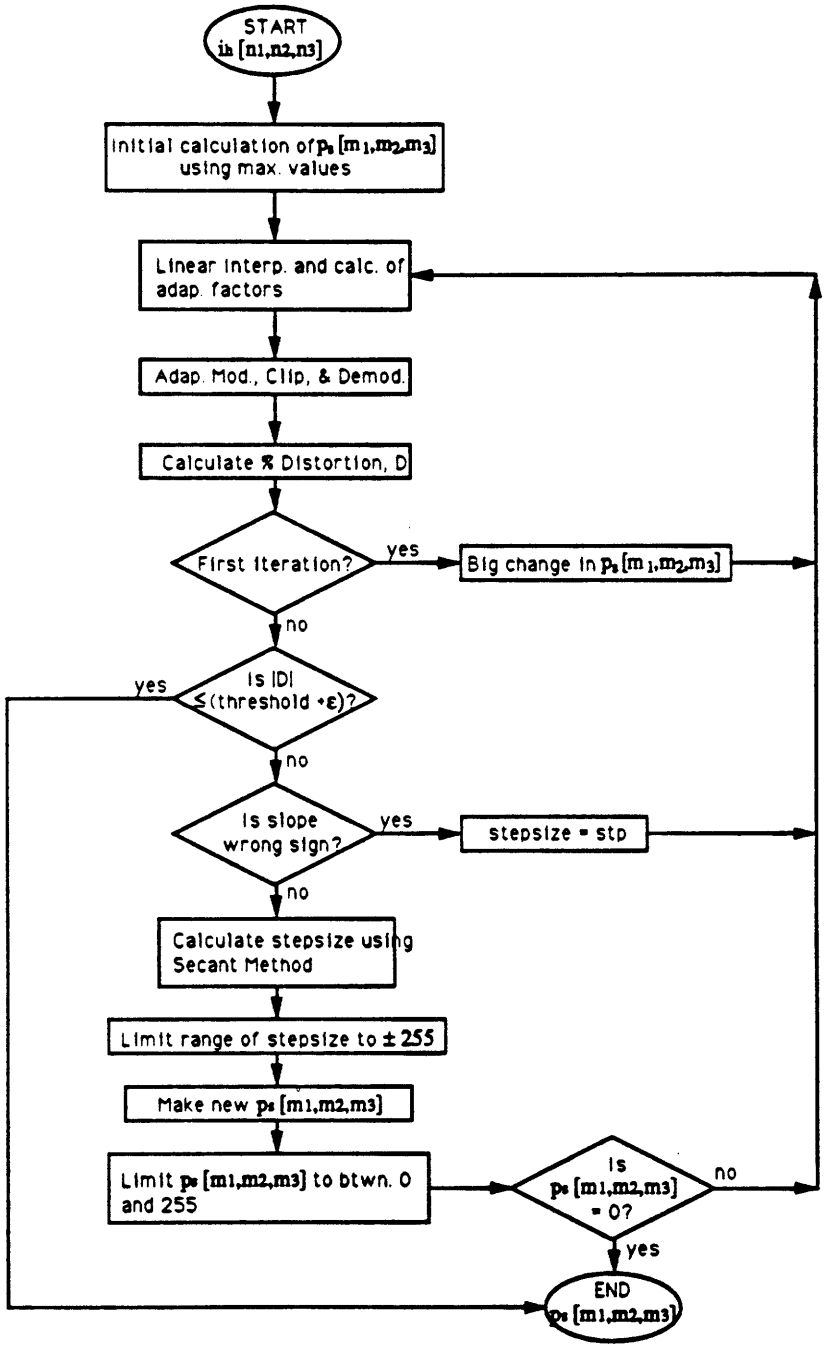


Figure 2.21: Block diagram of the iterative method of calculating the adaptation factors.

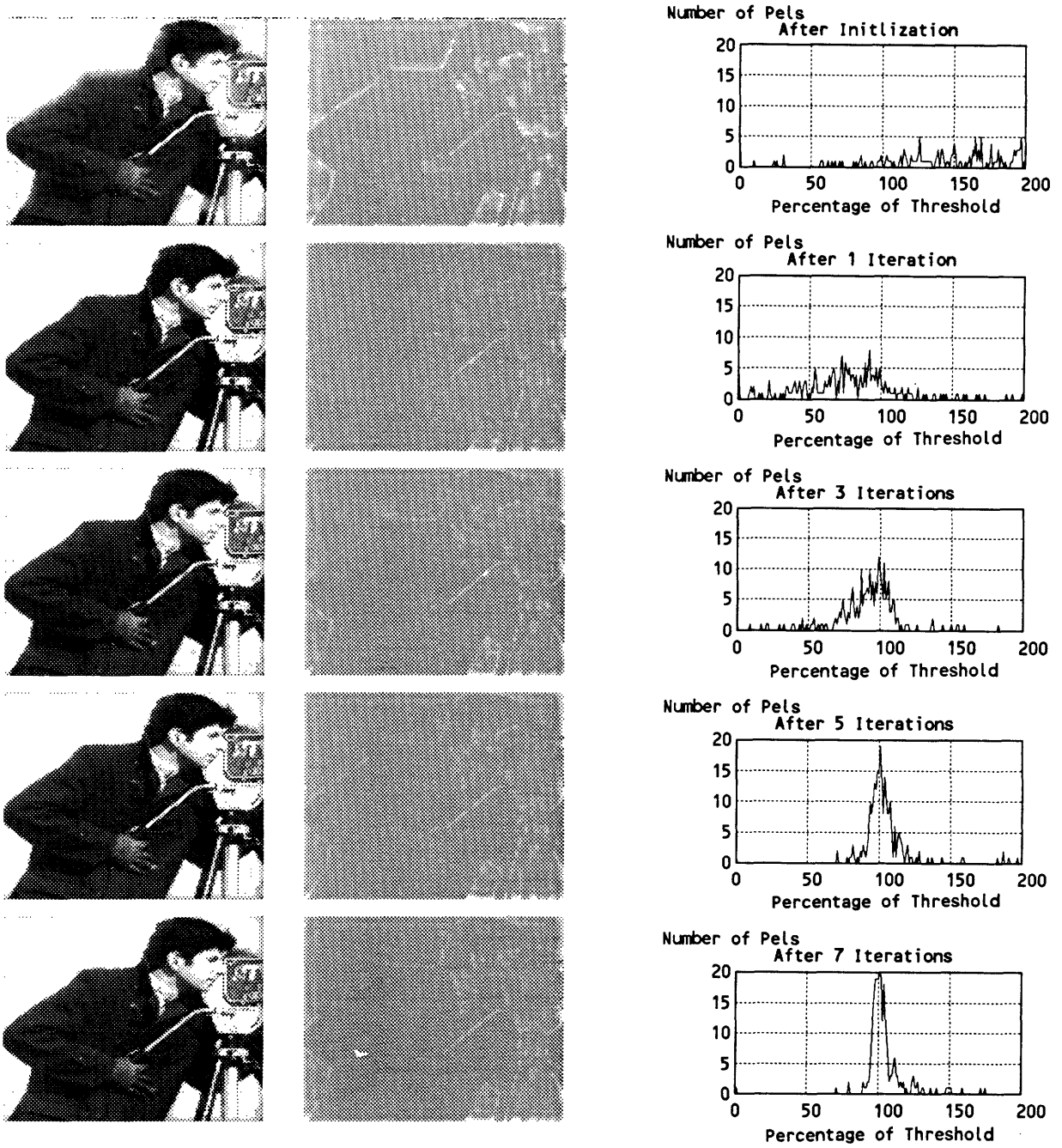


Figure 2.22: Demodulated pictures, distortion, and histograms of distortion for various iteration numbers.

where the summations and evaluations are done over the $N_1 x N_2 x N_3$ block and where

$$e[n_1, n_2, n_3] = i[n_1, n_2, n_3] - \hat{i}[n_1, n_2, n_3]. \tag{2.21}$$

The method for deciding upon which distortion measure and what distortion level to use can be stated as follows. For each distortion measure, find the resulting demodulated output for various distortion levels under noiseless channel conditions, of course. Choose the distortion level which allows the most noise reduction while, at the same time, the distortion remains imperceptible to just perceptible in the demodulated output. Do this for the different distortion measures and choose the distortion measure and level which allow the most noise reduction. In their Adaptive Frequency Modulation algorithm, Schreiber and Piot [SP88] use a similar iterative procedure to calculate their adaptation factors. They used the peak-absolute-error distortion measure and a distortion level of 3%.

2.5.3 Stills

Using each of the three distortion measures, we let our iterative algorithm run for 5 iterations at a variety of distortion levels: 5%, 10%, 15%, 20%, 25%, and 30%. Under ideal, noiseless channel conditions, the demodulated images produced when using the root-mean-squared-error and the peak-absolute-error measures at the same distortion level were of comparable picture quality and had similar appearance; whereas, the demodulated pictures produced when using the mean-absolute-error measure at the same distortion level were of lower picture quality. When a comparison is made at equal picture quality, the algorithm that uses the root-mean-squared-error has slightly greater noise suppression – about a 1.0dB SNR improvement.

Figure 2.23 shows the demodulated images that were produced when using the root-mean-squared error at three distortion levels: 5%, 15%, and 25%. The respective SNR for these distortion levels are 32.9dB, 30.3dB, and 27.6dB. Figure 2.24 shows the demodulated images that were produced when using the peak-absolute-error at three distortion levels: 10%, 20%, and 30%. The respective SNR for these distortion levels are 32.8dB, 31.4dB, and 29.2dB. The appearance of the distortion is greatest near edges. On normal

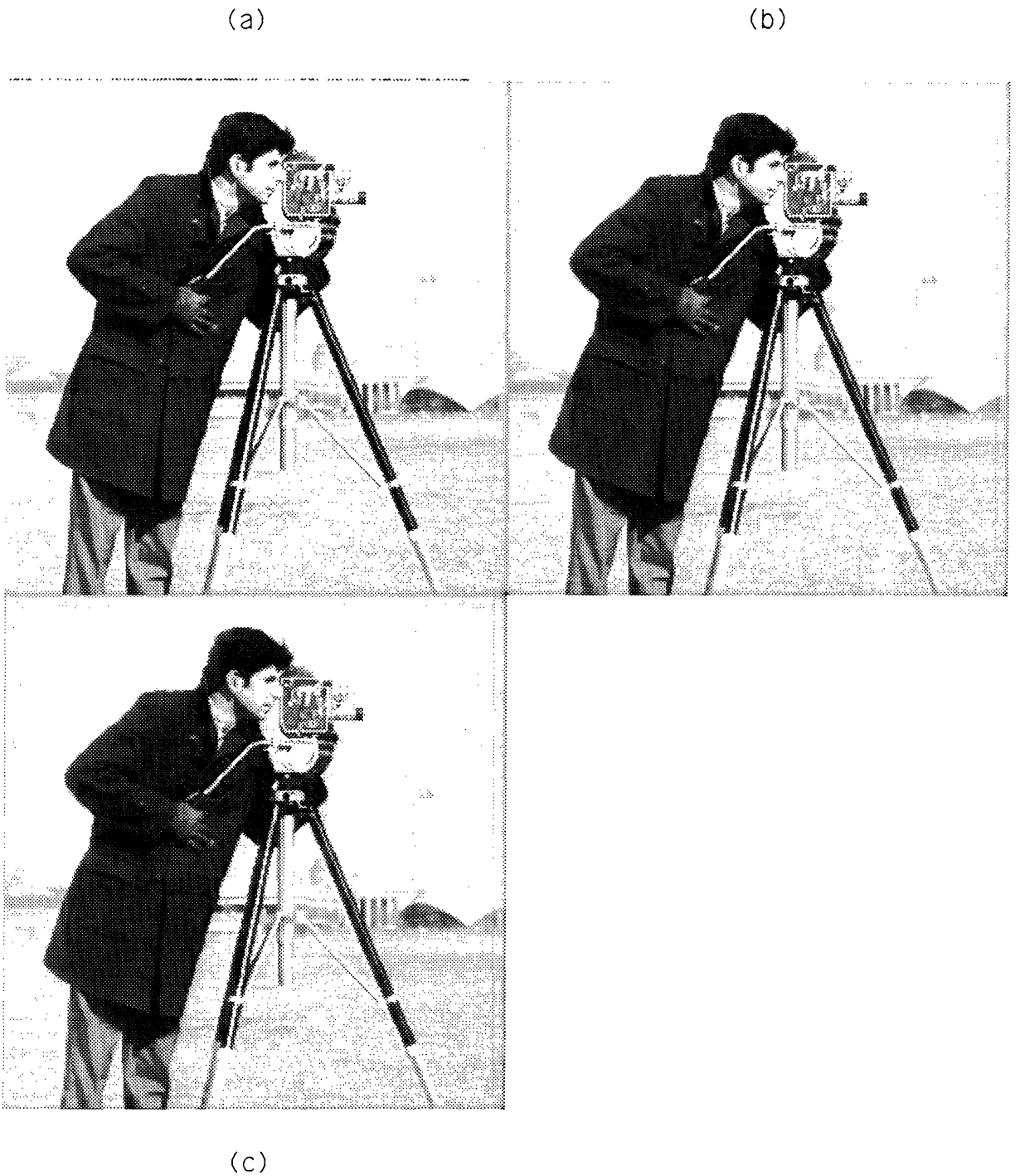


Figure 2.23: Demodulated pictures using the root-mean-squared-error distortion measure at three distortion levels: (a) 5%, (b) 15%, and (c) 25%.

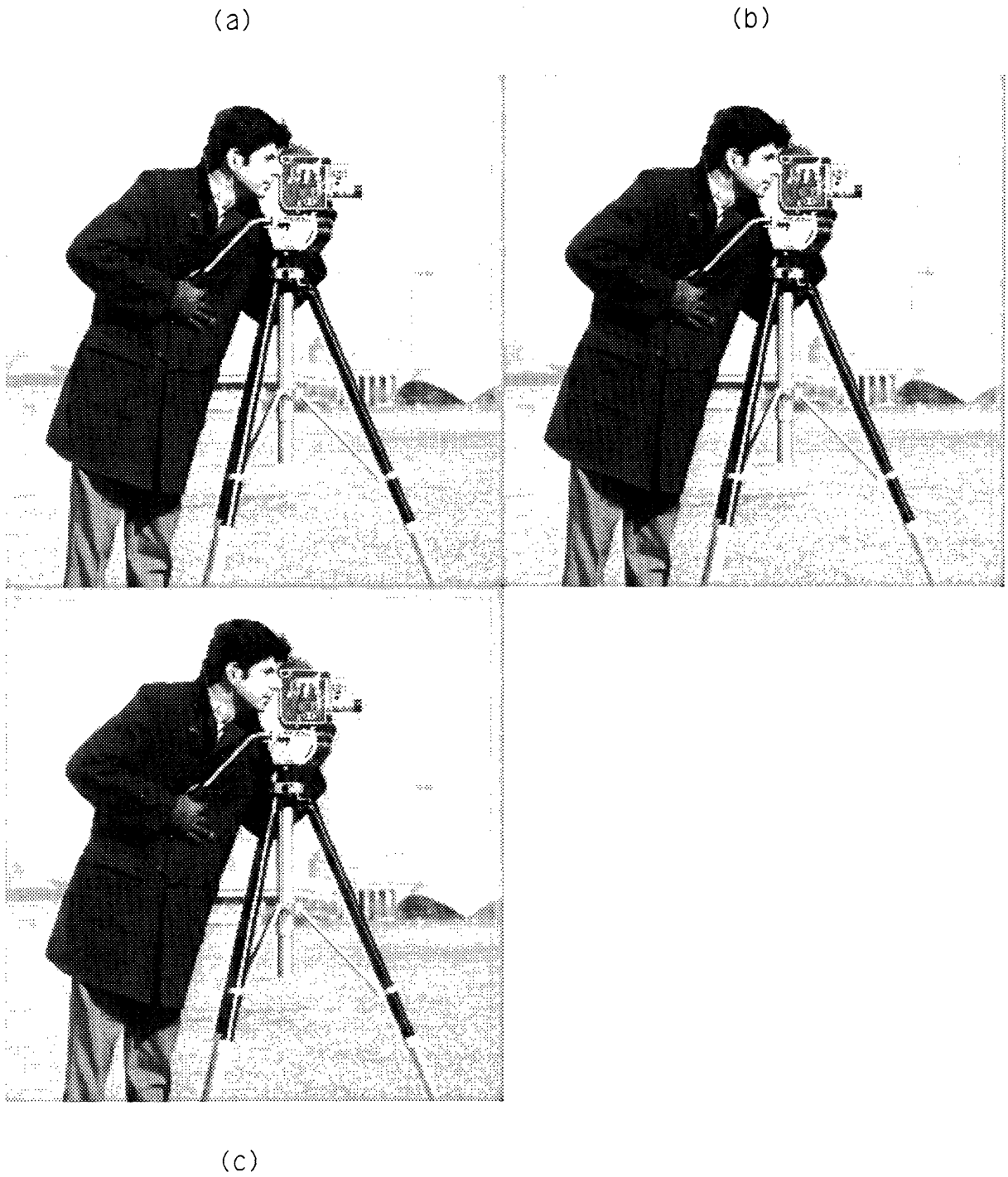


Figure 2.24: Demodulated pictures using the peak-absolute-error distortion measure at three distortion levels: (a) 10%, (b) 20%, and (c) 30%.

cathode-ray tubes at normal brightnesses, the distortion in the images when the root-mean-squared-error distortion level is 15% and when the peak-absolute-error distortion level is 20% is in the range of imperceptible to just perceptible. We conclude that when using an iterative method, which is based upon the amount of distortion in the demodulated output under ideal, noisefree channel conditions, to calculate the adaptation factors, one should use a root-mean-squared-error distortion measure and a 15% distortion level. Using this distortion measure and level will produce images with imperceptible to just perceptible distortion and will result in the greatest noise suppression of the the three distortion measures.

Now let's compare the noise suppression of the iterative method when using the above distorion measure and level to that of the method of using the maximum absolute values in the blocks and to the optimal decimation-and-interpolation method of calculating the adaptation factors. Figure 2.25 shows three demodulated images where the channel has a 20dB CNR. The upper two images use the maximum absolute values in the blocks and optimal decimation-and-interpolation with a scaling factor of 1.2 to calculate the adaptation factors. The iterative method is used in the lower image. Although the optimal decimation-and-interpolation method suppresses the noise slightly better than the iterative method (0.8 SNR gain), the iterative method produces sharper pictures. At 20dB CNR, the SNR gain is approximately 2.7 dB when using the iterative method compared to when the maximum absolute values in the blocks. This SNR gain is the same at 15dB CNR and 25dB CNR.

2.5.4 Sequences

When using iterative methods to calculate the adaptation factors for sequences, we can use three different methods. The first method uses three-dimensional blocks instead of two-dimensional blocks, the second method uses two-dimensional blocks but passes the state of the current frame to the next frame, and the third method treats each frame as a still image and uses the 2-D iterative method on each frame – no state is passed from frame to frame.

The input image, $i[n_1, n_2, n_3]$, is three-dimensionally filtered when using 3-D blocks.

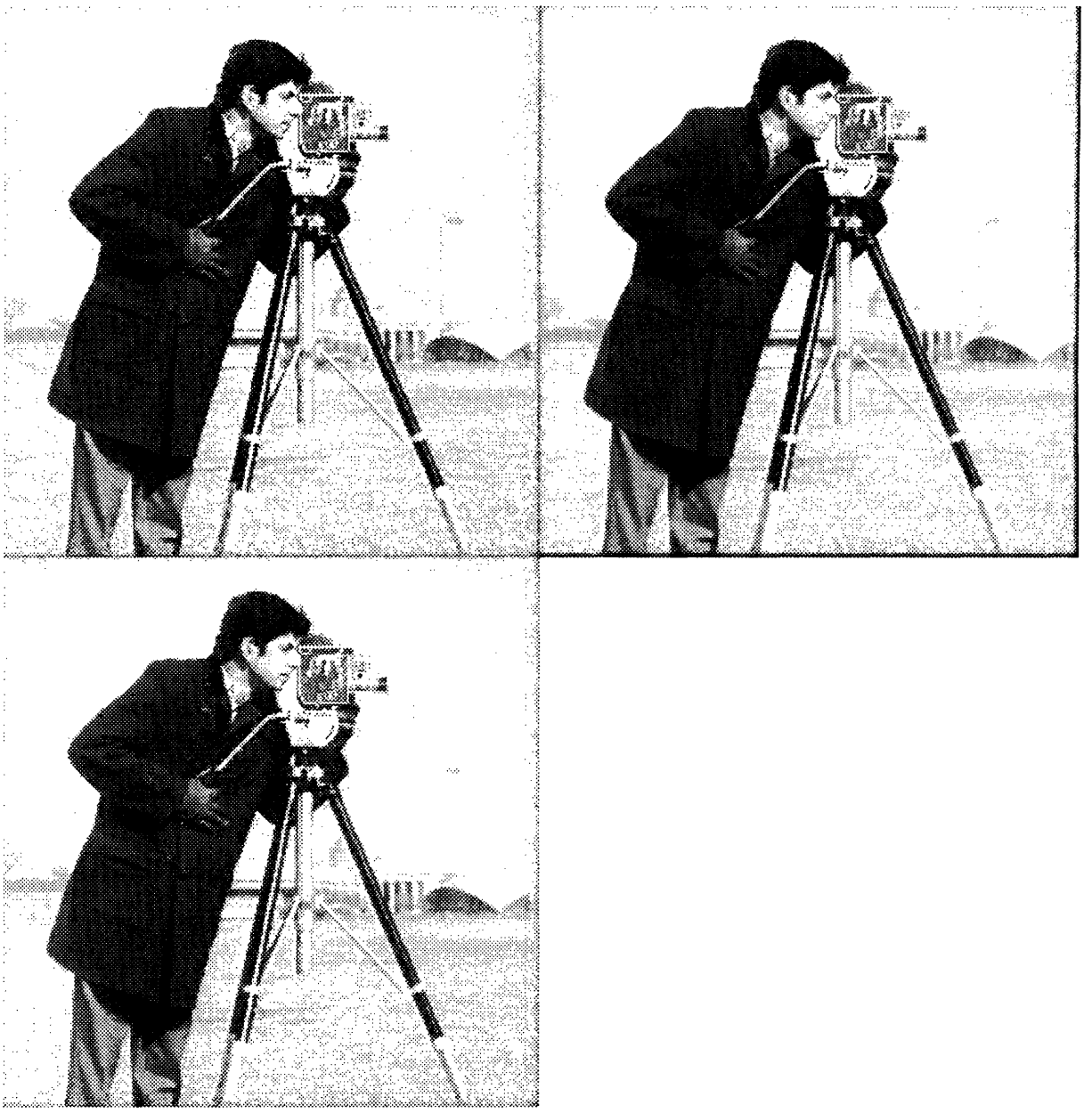


Figure 2.25: Comparison of three methods of calculating the adaptation factors. The upper left uses the linearly interpolated maximum absolute values in the blocks, the upper right uses the optimal decimation-and-interpolation method with a scaling factor of 1.2, and the lower picture uses the iterative method with the rmse distortion measure at 15%.

Figure 2.26: One original frame and one noisy frame (20dB CNR).

This method of adaptive modulation is just an extension of the ideas described in the previous section for 2-D blocks, except now one adaptation factor is calculated for each 3-D block, and the distortion is calculated over a 3-D block. Figure 2.26 shows one frame from an original sequence and one frame form a noisy sequence (20dB CNR). Figure 2.27 and Figure 2.28 show one frame for each of four demodulated sequences. The four demodulated sequences resulted from adaptively modulating two sequences of two different amounts of translational motion with an ideal and noisy (20dB CNR) channel. The maximum absolute values in 4x4x4 blocks are used to calculate the adaptation factors in Figure 2.27, and the 3-D iterative method that uses the root-mean-squared-error distortion measure at a 15% distortion level is used to calculate the adaptation factors in Figure 2.28.

When 2-D blocks and 2-D band separation are used, the effect of luminance changes on noise suppression is confined only to the spatial dimensions; however, when 3-D blocks and 3-D band separation are used, the effect of luminance changes on noise suppression

Figure 2.27: Using linear interpolation of the maximum absolute values in 4x4x4 blocks under ideal and noisy (20dB CNR) channel conditions. The upper two frames are for motion of 1 pel per frame. The lower two frames are for motion of 4 pels per frame.

Figure 2.28: Using the iterative method to calculate the adaptation factors in 4x4x4 blocks under ideal and noisy (20dB CNR) channel conditions. The upper two frames are for motion of 1 pel per frame. The lower two frames are for motion of 4 pels per frame.

is felt temporally. This means that noise will exist at the same spatial location as the temporal luminance change in frames adjacent to a temporal luminance change. When the motion of objects from frame to frame is too large (greater than 1 pel per 1/60th-of-a-second frame) and if one only looks at a single frame, then large amounts of noise can appear far from spatial luminance changes in that frame (see lower right of Figure 2.27 and Figure 2.28). Spatial masking effects will not adequately mask this noise. Temporal masking effects will mask the noise if the viewer is not tracking moving objects; however, if the viewer is tracking moving objects, then temporal masking effects cease to be significant, and spatial masking effects will fail to adequately mask the noise.

Only 2-D blocks and 2-D band separation are involved in the second method of iteratively finding the adaptation factors for sequences; however, the state of the current frame is passed to the next frame, and only one or two iterations are performed per frame, thereby saving computation time and decreasing processing delay at the transmitter. If the subscript k denotes current values in the final iteration of the current frame, then the state that is passed to the next frame consists of the past and present subsamples, $p_s^{k-1}[m_1, m_2, m_3]$ and $p_s^k[m_1, m_2, m_3]$, and the past distortion, $e^{k-1}[n_1, n_2, n_3]$. The next frame then uses these values as its past and present values for its first iteration using the Secant method. The iterative process continues on this frame, which now has become the current frame, until the specified number of iterations is reached. At this point, the current state of this frame is passed to the next frame.

Figure 2.29, where a single frame from each sequence is shown, shows the results of performing 1, 3, and 7 iterations on each frame of the sequence under ideal, noiseless channel conditions. The motion is translational in the horizontal direction with a displacement of 1 pel per frame. Also shown is the result of performing 3 iterations on a sequence where the displacement is 4 pels per frame. The block size is 4x4, and the distortion measure is the rmse distortion measure at 15%. we see that significant and highly visible amounts of distortion occur near edges even after several iterations are performed per frame.

The large distortion is due to large differences between the distortion curves of some blocks of the current frame and the new distortion curves of the same blocks of the next

(a) (b)

(c) (d)

Figure 2.29: Demodulated pictures produced when state is passed from frame to frame in the iterative method under ideal channel conditions: (a) 1, (b) 3, and (c) 7 iterations per frame when the motion is 1 pel per frame. Part (d) has 3 iterations per frame when the motion is 4 pels per frame.

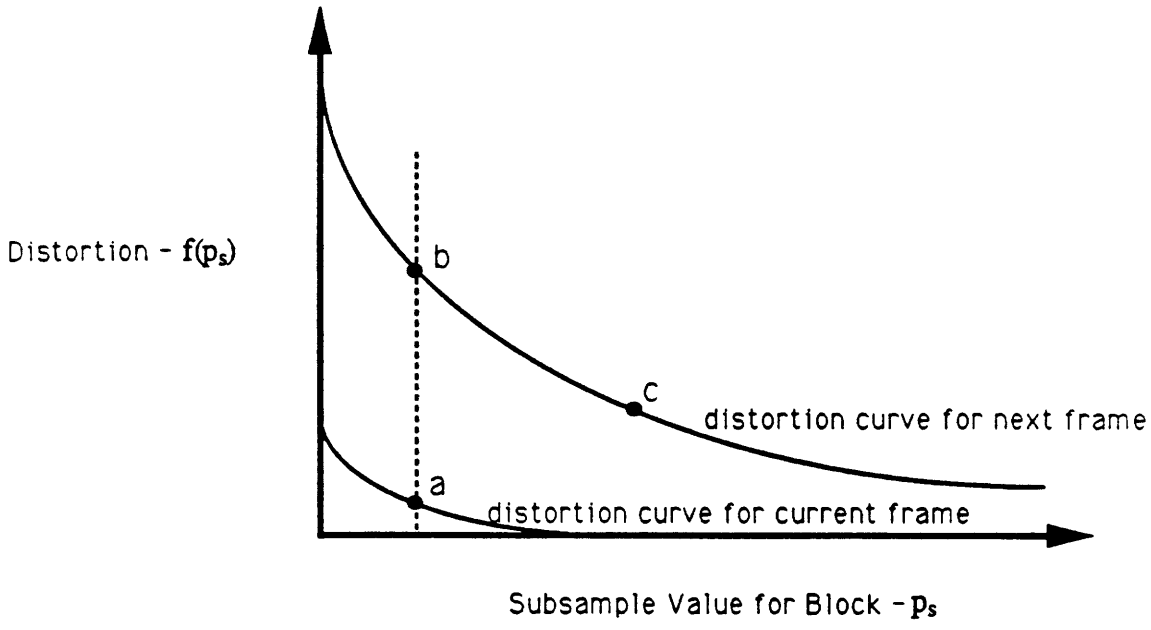


Figure 2.30: Two distortion curves. The lower curve is for a blank-area block and the upper curve is for the same block after an impulse has moved into it.

frame and to allowing an insufficient number of iterations for convergence in these blocks. Suppose that the iterative algorithm has converged to a subsample value that produces distortion of the desired level in a block of a frame. The iterative algorithm followed the distortion curve associated with this block. When the algorithm begins its search on the same block of the next frame, it must perform its search on a new distortion curve that is associated with the new frame; however, the state associated with the old frame is used. If the old and the new curves are vastly different, as is the case when an edge or impulse moves into a blank region, then the iterative algorithm requires many iterations to converge to a desirable subsample value (a subsample value that produces distortion close to the desired level) for the block of the new frame. Figure 2.30 shows two distortion curves. The lower curve is the curve that a blank-area block may have; the upper curve is the curve associated with the same block of the next frame if an edge has moved into the block. The subsample value associated with point *a* is a desirable subsample value of the current frame, and the subsample value associated with point *c* is a desirable value of the next frame. Point *b* will be the first point on the curve of the next frame when the state

Figure 2.31: Results of using the 2-D iterative method on each frame separately under ideal and noisy (20dB CNR) channel conditions. The upper two pictures use the iterative method, and, for comparison, the lower two pictures use linear interpolation of the maximum absolute values in the blocks to calculate the adaptation factors.

of the current frame is passed to the next frame. Points b and c are quite different, and many iterations will be required for convergence. The rate of convergence can be sped up by using a large stepsize on the first iteration of each frame (as mentioned before, this is done at the very beginning of the algorithm) or by using some sort of motion estimation algorithm to change the location of point b on the new distortion curve. If we use a large stepsize at the start of each frame, then we are basically ignoring the state of the past frame and may as well begin the iterative algorithm anew for each frame.

If we perform the 2-D iterative method of calculating the adaptation factors anew for each frame, we get the results shown in Figure 2.31. If there is no motion blur, then the pictures represent the results at any amounts of motion. Also shown are the demodulated pictures produced when the maximum absolute values in 2-D blocks are

used to calculate the adaptation factors. For a given number of iterations performed, starting the 2-D iterative algorithm over again for each frame produces images with less distortion than those produced by passing the state from frame to frame and continuing the algorithm from that point onward. Because 2-D blocks are used, the noise will always be masked by spatial effects, and the tracking of moving objects will not have an adverse affect on the masking of the noise. This method produces the best looking pictures of the three methods.

Chapter 3

Scrambling

Normally, a sequence of images is converted to a one-dimensional signal for transmission by scanning the sequence in a raster fashion (Figure 3.1). When this signal is subjected to interference from other similarly raster scanned signals or is degraded by intersymbol interference, the degradation produced in the received signal is highly correlated with itself and, in the case of intersymbol interference, is highly correlated with the desired signal, $i[n_1, n_2, n_3]$. The degradation can show up as ghosts or an interfering picture. In the case of quantization error, the error is subjectively most annoying when it is correlated with the desired image and least annoying when uncorrelated. Similarly, informal tests have shown that degradation produced by channel defects such as interference from other signals and intersymbol interference are less annoying when uncorrelated with the desired signal and when uncorrelated with the degradation itself. The purpose of scrambling (or pseudorandom scanning) is to decorrelate the degradation (or error between the desired signal and the demodulated signal) with itself and with the desired signal.

3.1 Picture Coding

A measure of picture quality often used in picture coding is the mean-squared-error (mse) between the input, $i[n]$ and the decoded output, $\hat{i}[n]$. The error is $e[n] = i[n] - \hat{i}[n]$. Many coding schemes can achieve a certain mse, with some coding schemes delivering

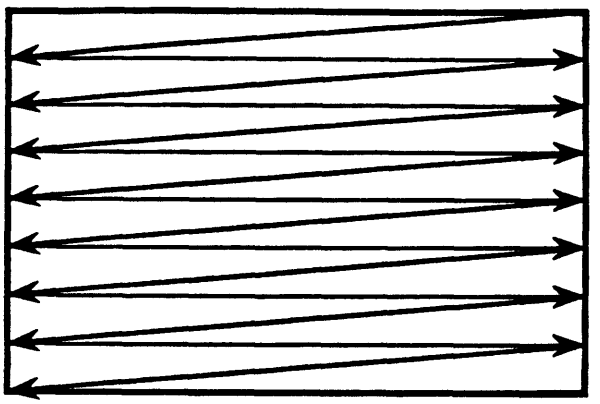


Figure 3.1: Raster scanning.

better pictures than others. In an effort to reduce the annoying effects of contours due to coarse quantization and to evaluate the relative picture quality produced by different coding schemes, L. G. Roberts divided the mse into two components [Rob62] [JN84]. We can write the mse as

$$\xi = E((i[n] - \hat{i}[n])^2) \tag{3.1}$$

where $E(.)$ denotes the expected value. In terms of the probability density function $p(i[n], \hat{i}[n])$ then

$$\begin{aligned} \xi &= \int\int (i[n] - \hat{i}[n])^2 p(i[n], \hat{i}[n]) di[n] d\hat{i}[n] && (3.2) \\ &= \int\int (i[n] - E(\hat{i}[n]|i[n]) \\ &\quad + E(\hat{i}[n]|i[n]) - \hat{i}[n])^2 p(i[n], \hat{i}[n]) di[n] d\hat{i}[n] && (3.3) \end{aligned}$$

where $E(\hat{i}[n]|i[n])$ is the expected value of $\hat{i}[n]$ given $i[n]$. Expanding and cancelling terms, we get

$$\begin{aligned} \xi &= \int (i[n] - E(\hat{i}[n]|i[n]))^2 p(i[n]) di[n] \\ &\quad + \int\int (\hat{i}[n] - E(\hat{i}[n]|i[n]))^2 p(i[n], \hat{i}[n]) di[n] d\hat{i}[n]. && (3.4) \end{aligned}$$

The mse can be divided into two terms,

$$I_1 = \int (i[n] - E(\hat{i}[n]|i[n]))^2 p(i[n]) di[n], \tag{3.5}$$

which Roberts calls the tonal error, and

$$I_2 = \int\int (\hat{i}[n] - E(\hat{i}[n]|i[n]))^2 p(i[n], \hat{i}[n]) di[n] d\hat{i}[n], \tag{3.6}$$

which Roberts calls the variance error. Tonal error shows up in images as the contours that are brought about by the coarse quantization of images, and variance error describes the random variation in the output, $\hat{i}[n]$, about the value of the input, $i[n]$. Roberts finds that human observers are more annoyed by the tonal error than by the variance error. The tonal error is eliminated when $E[\hat{i}[n]|i[n]] = i[n]$, and Roberts demonstrates that this criterion is met when the input signal is multiplied by a properly chosen pseudorandom noise sequence prior to quantization and when the received signal is divided by the same pseudorandom noise sequence. In addition, Schuchman [Sch64] shows that the tonal error, I_1, is zero if and only if the error, $e[n]$, is zero-mean and is independent of the input; that is,

$$p(e[n]|i[n]) = p(e[n]). \tag{3.7}$$

In other words, quantization noise is least visible when it is independent of the input. One additional constraint for minimum visibility of the error is that the error should resemble a white-noise process.

3.2 Channel Defects

How might one apply Robert's criteria for reducing the annoyance of coding errors to reducing the annoyance of channel defects? Suppose we let $y'[n]$ be the transmitted signal and $\hat{y}[n]$ the received signal (see Figure 2.1 and Figure 2.2), then the error due to any channel defect is

$$e[n] = \hat{y}[n] - y'[n]. \tag{3.8}$$

In the present discussion, we shall only consider channel degradations that occur at baseband. In Chapter 5, we shall consider channel degradations that occur at RF. When transmission occurs at baseband, the channel can be modelled as a linear system, so that

$$\hat{y}[n] = \sum_k h[k] y'[n-k] + b[n] \tag{3.9}$$

$$= h[0]y'[n] + \sum_{k'} h[k]y'[n-k] + b[n] \tag{3.10}$$

$$= y'[n] + \frac{1}{h[0]}(\sum_{k'} h[k]y'[n-k] + b[n]), \tag{3.11}$$

where $h[n]$ is the channel impulse response, $b[n]$ is interference from another signal source, and k' denotes all $k \neq 0$. Any frequency rolloff of the channel or any multipath that may occur in the channel is incorporated in the channel impulse response,$h[n]$. The signal $b[n]$ represents co-channel or adjacent-channel interference. The error for this system is,

$$e[n] = \frac{1}{h[0]}(\sum_{k'} h[k]y'[n-k] + b[n]). \tag{3.12}$$

Analogous to picture coding, some transmission systems will produce better looking pictures at the receiver than other systems for a given channel and corresponding SNR. Learning from Roberts, we should choose that system which makes the error resemble a white-noise process and uncorrelated with the input. To see how to do this, let's write out the equations for the cross-correlation between the input, $y'[n]$ and the error, $e[n]$. Without loss of generality, we assume that the input and output are zero-mean processes. Thus,

$$\begin{aligned} E(y'[n]e[n]) &= E(y'[n]\frac{1}{h[0]}(\sum_{k'} h[k]y'[n-k] + b[n])) \\ &= E(\frac{1}{h[0]}(\sum_{k'} h[k]y'[n]y'[n-k] + y'[n]b[n])) \\ &= \frac{1}{h[0]}(\sum_{k'} h[k]E(y'[n]y'[n-k]) + E(y'[n]b[n])), \end{aligned} \tag{3.13}$$

where $E(.)$ is the expected value operation. We can rewrite the error as,

$$e[n] = w[n] + \frac{1}{h[0]}b[n], \tag{3.14}$$

where $w[n]$ is the part of the error due to linear frequency distortion and involves contributions from adjacent pels in the channel. Normally this signal is highly self correlated. The auto-correlation of the error is,

$$\begin{aligned} E(e[n]e[n-l]) &= E((w[n] + \frac{1}{h[0]}b[n])(w[n-l] + \frac{1}{h[0]}b[n-l])) \\ &= E(w[n]w[n-l]) + \frac{1}{h[0]}E(w[n]b[n-l]) \\ &\quad + \frac{1}{h[0]}E(w[n-l]b[n]) + \frac{1}{h[0]}E(b[n]b[n-l]). \end{aligned} \tag{3.15}$$

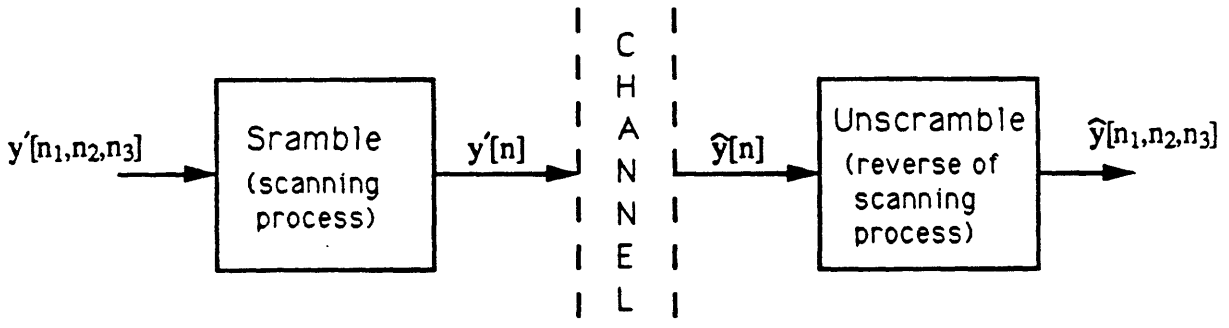


Figure 3.2: Block diagram of scrambling.

We want $E(y'[n]e[n]) = 0$ and $E(e[n]e[n-l]) = \sigma_e^2\delta[l]$. We can accomplish this by altering the scanning pattern that is used to convert the the sequence of input images to a one-dimensional signal for transmission. We first scan the sequence of input images pseudorandomly, transmit this signal through the channel, and then do the reverse of the pseudorandom scanning at the receiver (Figure 3.2). Scrambling at the transmitter makes the input signal, $y'[n]$, uncorrelated with itself (that is, $E(y'[n]y'[n-l]) = \sigma_{y'}^2\delta[l]$) and makes the input signal uncorrelated with any interfering signal so that $E(y'[n]b[n]) = 0$ for all n. Descrambling at the receiver removes the correlation from any correlated channel degradations such as $w[n]$ and $b[n]$, so that now $E(w[n]w[n-l]) = \sigma_w^2\delta[l]$ and $E(b[n]b[n-l]) = \sigma_b^2\delta[l]$. The signals $w[n]$ and $b[n]$ are uncorrelated with each other if $y'[n]$ and $b[n]$ are uncorrelated with each other. Substituting these equations into eqn. 3.13 and eqn. 3.15 we get $E(y'[n]e[n]) = 0$ and $E(e[n]e[n-l]) = \sigma_e^2\delta[l]$, where $\sigma_e^2 = \sigma_w^2 + \frac{1}{h[0]}\sigma_b^2$. This is what we desired.

By scanning the sequence in a pseudorandom fashion (scrambling), we do not change the SNR of the received signal in comparison to when raster scanning is used, but we can make the degradations caused by the channel appear as a white-noise process independent of the input. With scrambling, we make the degradations less subjectively annoying in comparison to when raster scanning is used, and the desired signal retains its full

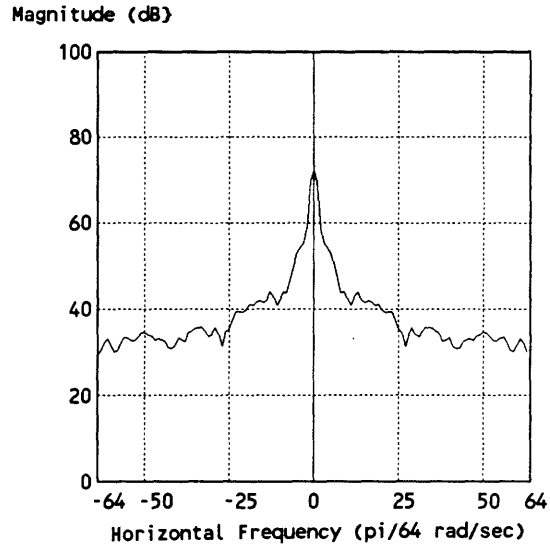
sharpness. Scrambling, also, does not change the energy of the signal, nor does it change its peak power. We have found that separable pseudorandom scanning in the horizontal and vertical directions in each frame produces an adequate noise-like appearance. We first scramble each row of a frame and, then, scramble each column, where the scrambling is independent from row to row and column to column. No scrambling is done along the temporal axis; however, a different seed for the random number generator is used for each frame.

3.3 Examples

If we scramble the entire image without decomposing it into highpass and lowpass frequency components, we get the image shown in the lower left of Figure 3.3. Also shown are the original, unscrambled image and the cross sections along the horizontal axes of the 2-D power-spectrum densities of the respective images. The mean has been subtracted from the images prior to calculating the power-spectrum densities. The energy of the original and of the scrambled images are exactly the same; scrambling does not change the energy of the signal. The power spectrum, however, is altered by the scrambling process, where scrambling tends to "whiten" the power spectrum; i.e., it tends to even out the energy over all frequencies.

Now suppose that the desired signal interferes with itself in the channel because of multipath or linear frequency rolloff. Multipath and linear frequency distortion can be treated in a similar manner because multipath can be modelled as a linear filter with an impulse at the origin and another impulse at a number of samples away that corresponds to the delay of the other path. A 40% echo with a delay of 20 pels produces the picture shown on the left of Figure 3.4. If scrambling is used, then the picture on the right of Figure 3.4 results. Scrambling decorrelates the echo with the desired signal and gives the echo the appearance of random noise.

If another signal interferes with the desired signal in the channel, the picture shown on the left of Figure 3.5 results, where the interfering picture is at 40% energy relative to the desired picture of CMAN. Scrambling decorrelates the interference and makes



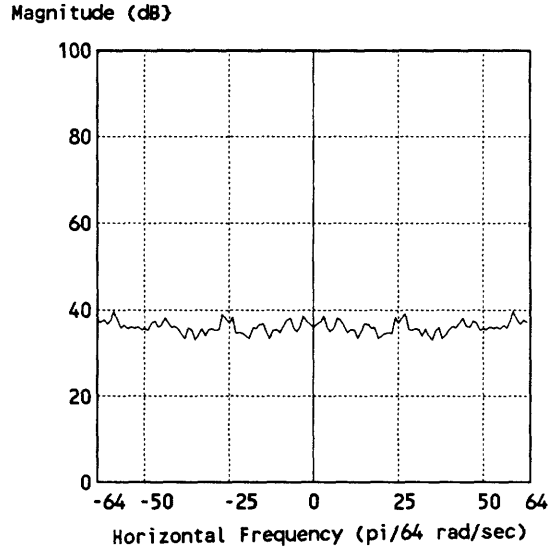


Figure 3.3: Scrambling of CMAN. The effect of scrambling upon the image and upon its power-spectrum density.

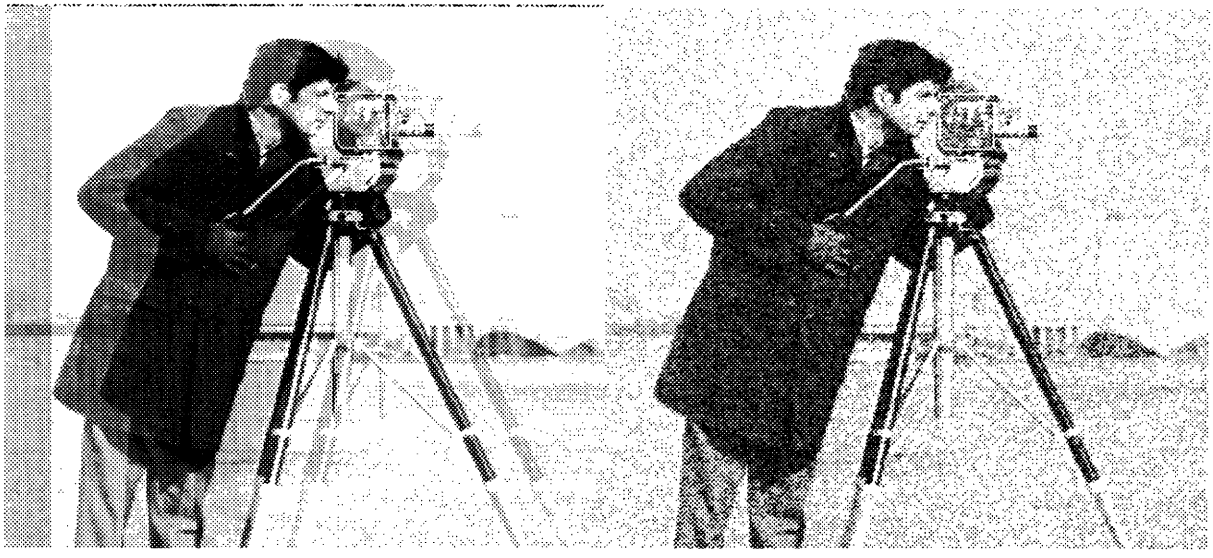


Figure 3.4: The appearance of 40% multipath without and with scrambling.

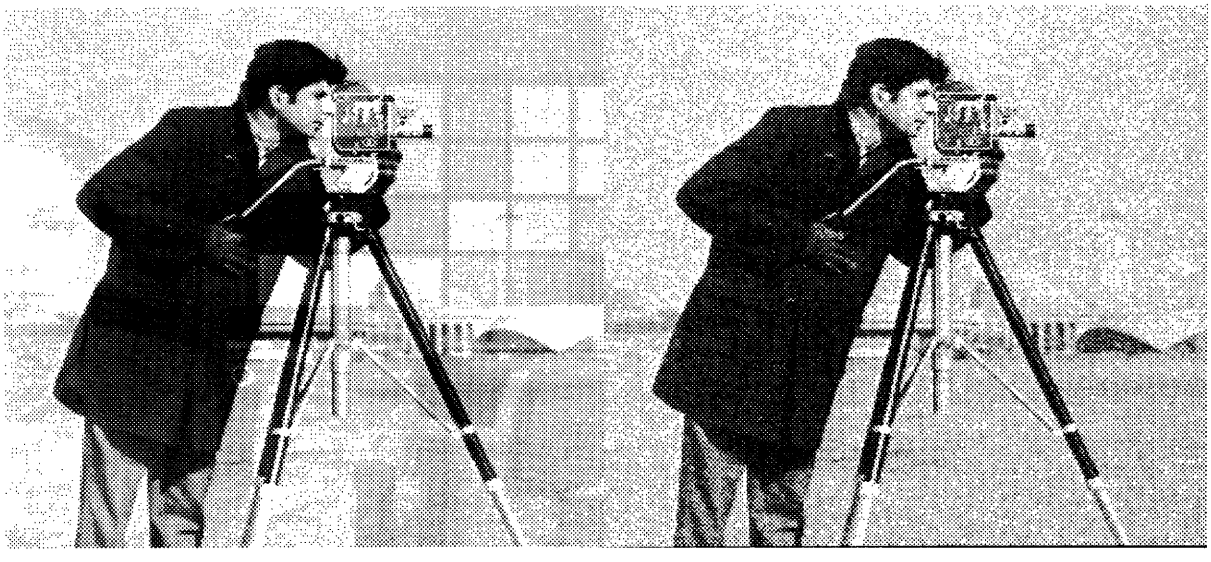


Figure 3.5: The appearance of 40% interference without and with scrambling.

the interfering signal look like random noise to the desired signal (see right picture in Figure 3.5). The interfering signal will always look like random noise to the desired signal even if the interfering signal is scrambled as long as different seeds are used in the random-number generators of the two signals.

Chapter 4

Adaptive Amplitude Modulation and Scrambling

The previous chapter investigated the use of scrambling to randomize channel defects, such as interference from other signals and linear channel defects (i.e., multipath, frequency rolloff, and other forms of intersymbol interference.) The result of descrambling at the receiver is to make channel degradations appear as additive white noise in the demodulated signal. Adaptive modulation reduces any noise, or any degradations for that matter, that may be added in the channel and can reduce the noiselike degradations that result from the application of scrambling over nonideal channels; thus, scrambling turns channel defects into pseudorandom noise, and adaptive modulation reduces the resulting pseudorandom noise. In this manner, interference is reduced and the channel is equalized. Before we investigate the use of the combination of scrambling and adaptive modulation to reduce channel degradations caused at radio frequencies (RF), let us look at the interaction of scrambling and adaptive modulation at baseband.

4.1 Why It Works

The baseband transmission system we will investigate is shown in Figure 4.1. Note that the system is only for high frequency components; the lows are assumed to be sent digitally in some noise-free manner. At first glance, it might appear that the combination

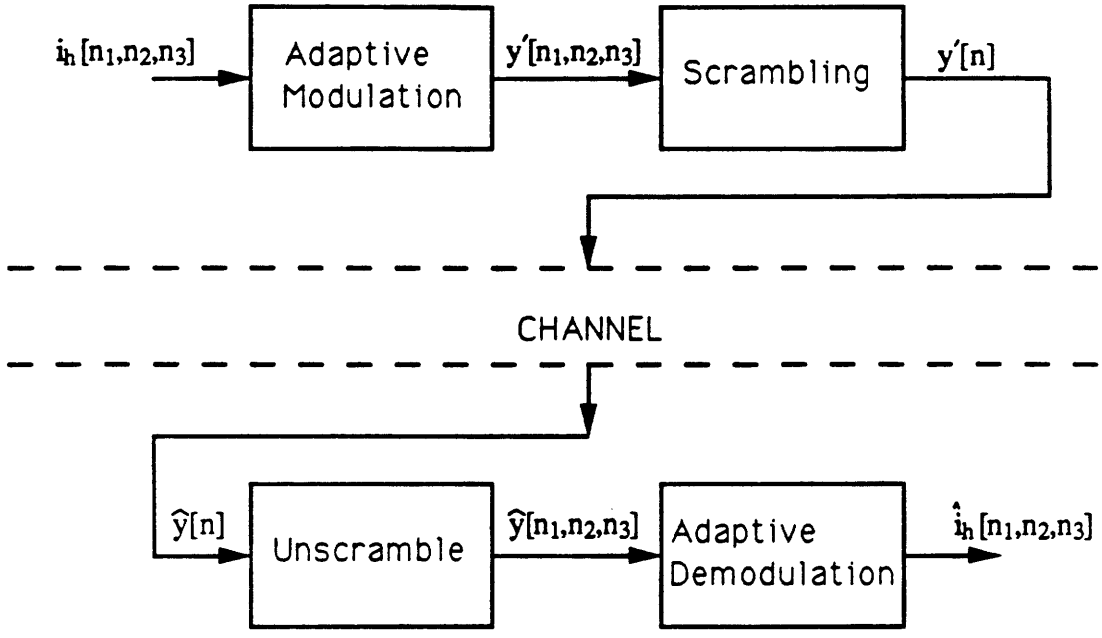


Figure 4.1: Block diagram of adaptive modulation and scrambling at baseband.

of adaptive modulation and scrambing would work against each other and not reduce the pseudorandom noise due to imperfect channels because adaptive modulation raises the magnitude of the picture elements by a multiplicative adaptation factor at the transmitter and reduces the magnitude by the same factor at the receiver, with the end result being no change in the appearance of the pseudorandom noise compared to when scrambling is used alone. The combination of adaptive modulation and scrambling does work because of the movement of degrading picture elements relative to the desired picture elements and because of the proportion of blank to busy area picture elements.

Not all picture elements after the adaptive modulation process have the same magnitude. The ones with small magnitudes tend to also have small magnitudes after adaptive modulation, have a larger adaptation factor, and constitute a larger proportion of the image. The picture elements which have large magnitudes also have large magnitudes after adaptive modulation, cause the most degradation when unscrambled, and correspond to the edges or busy areas of the picture. These picture elements have a small adaptation

factor and form a smaller proportion of the picture. The received signal is composed of a desired signal and a degrading signal. When scrambling is reversed at the receiver, four combinations of desired picture elements and degrading picture elements can occur. The first combination occurs when the desired picture element is from a blank area and when the degrading picture element is from a busy area. In this case, the adaptation factor of the desired picture element is large and the adaptation factor of the degrading picture element is small. The degrading picture element will be significantly reduced in value by this large adaptation factor.

The second combination occurs when the desired picture element is from a busy area and when the degrading picture element is from a blank area. The adaptation factor of the desired picture element is small and the adaptation factor of the degrading picture element will be large. The degrading picture element will not be greatly reduced in value; however, its degrading effect will be small since the degrading picture element will most likely have a small value. In addition, the effect of the small value of the degrading picture element will be masked. The third combination occurs when the desired picture element and the degrading picture element are both from blank areas, and the fourth combination occurs when the desired picture element and the degrading picture element are both from busy areas. I both cases, the adaptation factors of the desired picture element and of the degrading picture element will be nearly the same so that effect of the degrading picture element will be marginally reduced, if reduced at all. When both picture elements are from blank areas, the degrading effect of the degrading picture element will be small since the value of the degrading picture element will most likely be small. The worst combination occurs when both picture elements are from busy areas because the degrading picture element will have a large value and will not be significantly reduced by the small adaptation factor of the desired picture element. Luckily, this combination is least likely to occur.

Because of the proportionalities of blank-area to busy-area picture elements, most busy-area picture elements will move into blank-area picture elements; thus, the picture elements which will degrade the picture the most will be significantly reduced by large adaptation factors. It is this reason that the combination of adaptive modulation and

scrambling works so well.

4.2 Interaction of Adaptive Amplitude Modulation and Scrambling

The energy of the pseudorandom noise in the demodulated signal brought about by the scrambling operation under nonideal channel conditions depends upon the maximum value of the adaptation factors used. Because the input signal is multiplied by a set of adaptation factors, the larger the maximum value, and hence, the noise-reducing capabilities, then the higher the energy of the adaptively modulated signal. When intersymbol interference exists, the energy of the pseudorandom noise added to the received signal after descrambling is proportional to the energy of the adaptively modulated signal itself. When interference from other adaptively modulated signals exists in the channel, the energy of the pseudorandom noise is proportional to that of the interfering signal. This means that the energy of the pseudorandom noise is dependent upon the maximum value of the adaptation factors, and if the maximum value of the adaptation factors is too large, then the noise near edges or in the busy portions of the image will be too strong to be adequately masked by the detail. On the one hand, a large maximum adaptation factor means better noise reduction in the slowly varying and blank areas of the picture, but, on the other hand, the degradation near edges will be worsened. Conversely, having too small a maximum adaptation factor will not adequately reduce noise in the slowly varying or blank areas of the picture; although, a small maximum adaptation factor will reduce the strength of the pseudorandom noise that will be passed near edges. The purpose of this section is to find a value for the maximum adaptation factor that will be a subjectively pleasing compromise between suppression of noise in the blank areas and strength of noise passed in the busy areas when an adaptively modulated signal degrades itself or another adaptively modulated signal.

The plot of the root-mean-square value, which is a measure of energy, of the adaptively modulated highpass signal versus the maximum adaptation factor is shown in Figure 4.2. The energy of the adaptively modulated signal is a monotonically increasing function of

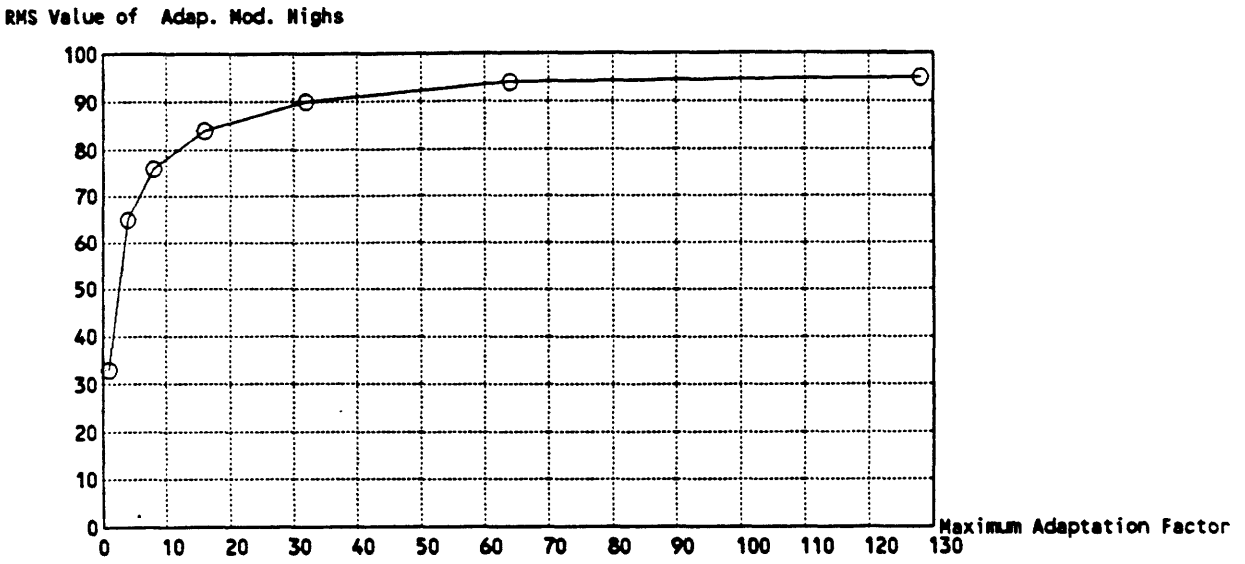


Figure 4.2: Root-mean-square value of the adaptively modulated highs versus the maximum adaptation factor.

the maximum adaptation factor. The energy increases rapidly for small values of the maximum adaptation factor, but the rate of increase decreases for large values of the maximum adaptation factor. This occurs because of the nature of adaptive modulation: not all picture elements can have the maximum adaptation factor. As we increase the value of the maximum adaptation factor, the number of pixels that can have an adaptation factor this large and, hence contribute to the increase in the overall energy of the picture, becomes smaller and smaller. Therefore, the increase in energy of the adaptively modulated highpass signal slows as the maximum adaptation factor is increased.

The behavior of this plot has important implications for the strength of the pseudo-random noise that is passed in busy areas. In busy areas, the values of the adaptation factors are generally small (approximately 1 to 4) and, hence, are not affected by the limit on the maximum value of the adaptation factors until the maximum is close to these small values. If we assume that the adaptation factors remain constant in the busy areas as we change the value of the maximum adaptation factor, then the strength of the noise in busy areas versus the maximum adaptation factor has the same shape as the plot of the rms value of the adaptively modulated signal versus the maximum adaptation factor.

The strength of the noise in the blank areas has a very different behavior. In this

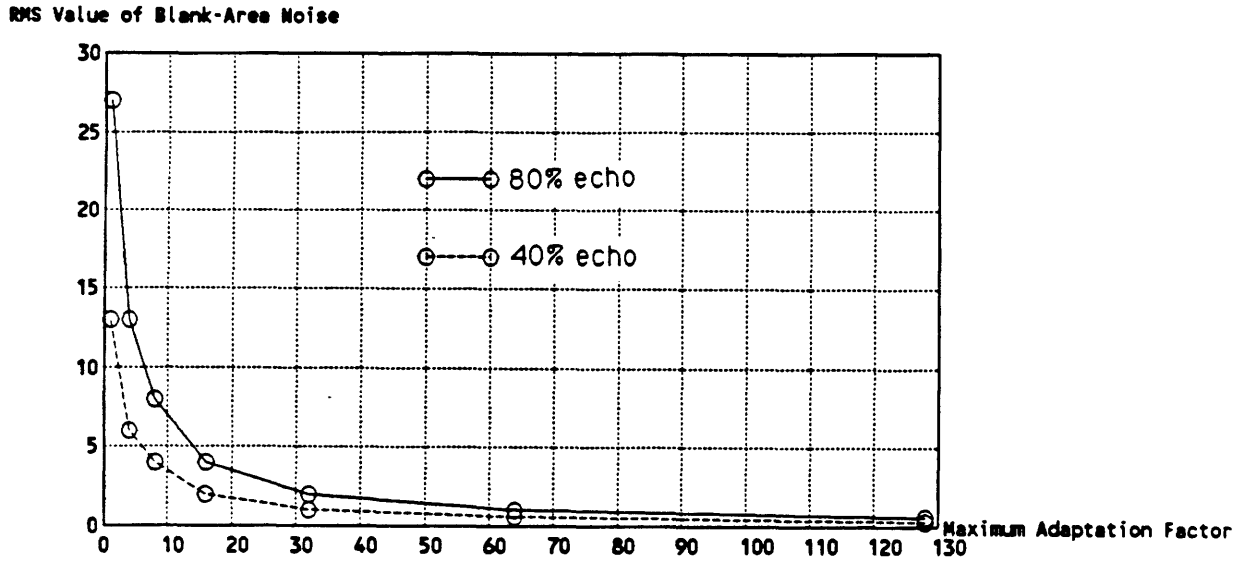


Figure 4.3: The rms value of noise in blank areas versus the maximum adaptation factor.

case, the noise suppression is directly related to the maximum value of the adaptation factors since blank-area pels will have the maximum value as the value of their adaptation factors. Figure 4.3 shows a plot of the rms value of the noise in blank areas versus the maximum adaptation factor when there is a 40% and 80% echo in the channel. The increase in the energy of the adaptively modulated signal is not as large as the increase in the maximum adaptation factor that caused the increase in energy; therefore, when an adaptively modulated signal degrades itself or another adaptively modulated signal, the blank-area noise is a decreasing function of the maximum value of the adaptation factors.

Now, let us consider the rms error over the entire image as a function of the maximum adaptation factor and try to choose a maximum value that will produce the best visual tradeoff between reduction of noise in blank areas and strength of noise in busy areas. Figure 4.4 shows the plots of the rms error versus the maximum adaptation factor for a picture that has been degraded by a 40% and 80% echo. The minimum of the graphs occur at a maximum adaptation factor of 8; however, visually, the best picture occurs when the maximum adaptation factor is approximately 16 (see Figure 4.5 and Figure 4.6 for the demodulated pictures corresponding to maximum adaptation factors of 8, 16, 32, and 64.) The discrepancy occurs because of visual masking. To explain this, let us first explain

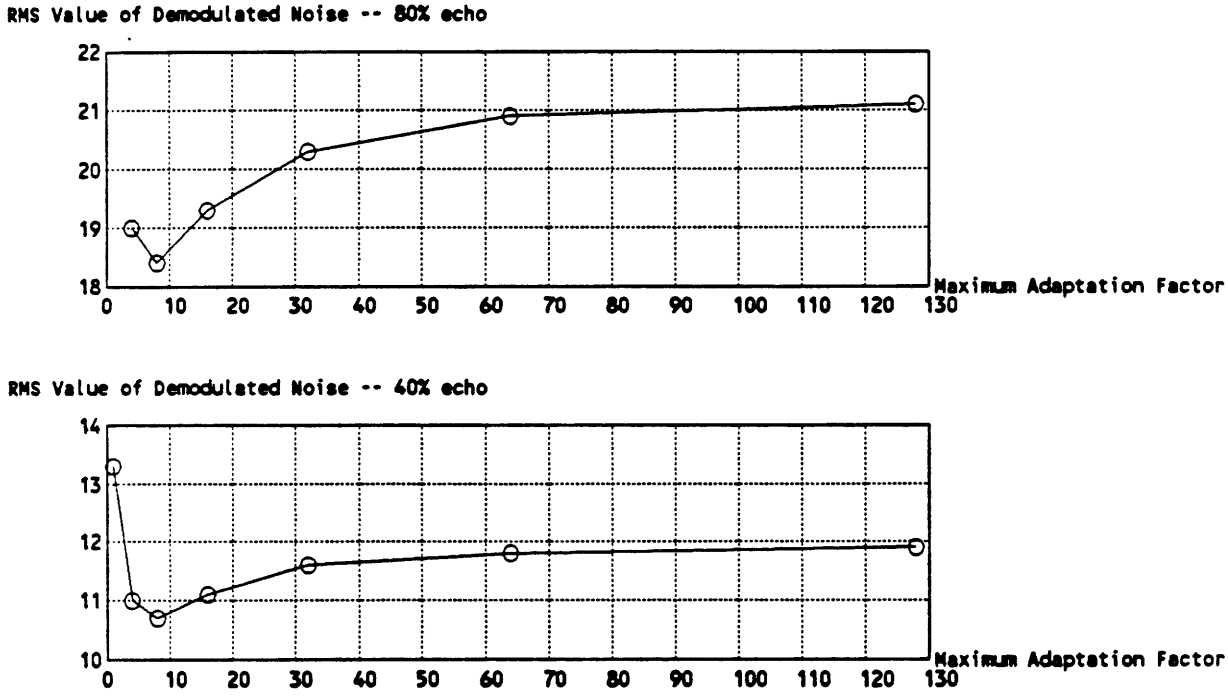


Figure 4.4: The rms error over the entire picture versus the maximum adaptation factor.

Figure 4.5: The demodulated pictures corresponding to a maximum adaptation factor of (a) 8, (b) 16, (c) 32, and (d) 64, where the picture is degraded by a 40% echo.

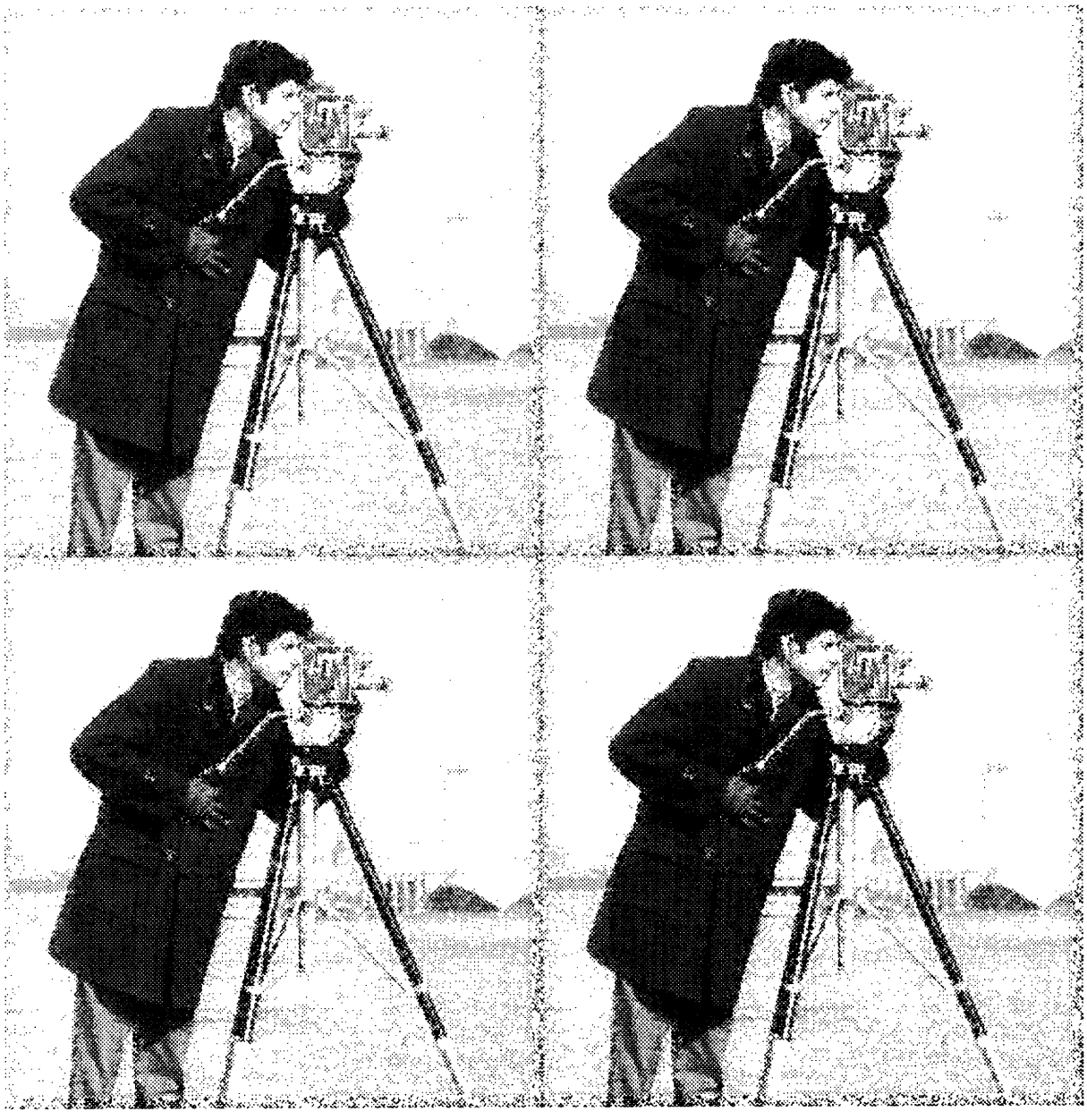


Figure 4.6: The demodulated pictures corresponding to a maximum adaptation factor of (a) 8, (b) 16, (c) 32, and (d) 64, where the picture is degraded by an 80% echo.

why we see this behavior of the graphs in Figure 4.4. For maximum adaptation factors less than 8, the noise in the blank and slowly varying areas dominate the rmse calculations and, also, the degrading effects in the image; whereas, for maximum adaptation factors greater than 8, busy areas dominate the rmse calculations and the degrading effects in the image. Visual masking reduces the degrading effects of noise in the busy areas of an image; therefore, visual masking will reduce the effective rmse contribution from busy areas. This reduction will lower the values of the error for maximum adaptation factors greater than 8 such that the minimum will occur at 16 for viewing distances of 4 times the picture height. If the viewing distance were reduced, then the minimum would move to lower values of the maximum adaptation factor.

What effect does the choice of the maximum adaptation factor have on the appearance of the demodulated signal when truly additive white Gaussian noise is added in the channel? In this case, the amount of noise that is added to the signal in the channel is independent of the maximum adaptation factor, which accounts for the behavior of the plot of the rms error of the demodulated signal versus the maximum adaptation factor (Figure 4.7.) The CNR in the channel for the above plot is 20dB and 15dB; however, the same behavior is observed for other carrier-to-noise ratios. In addition, the explanation for the behavior of Figure 4.4 pertains to the behavior of Figure 4.7. This graph would indicate that the change in picture quality is marginal when increasing the maximum value of the adaptation factor from 16 to 128. Even with visual masking effects, experiments (Figure 4.8) show little picture quality improvement when going from 16 to 128. We, therefore, choose to use a maximum value for the adaptation factors of 16.

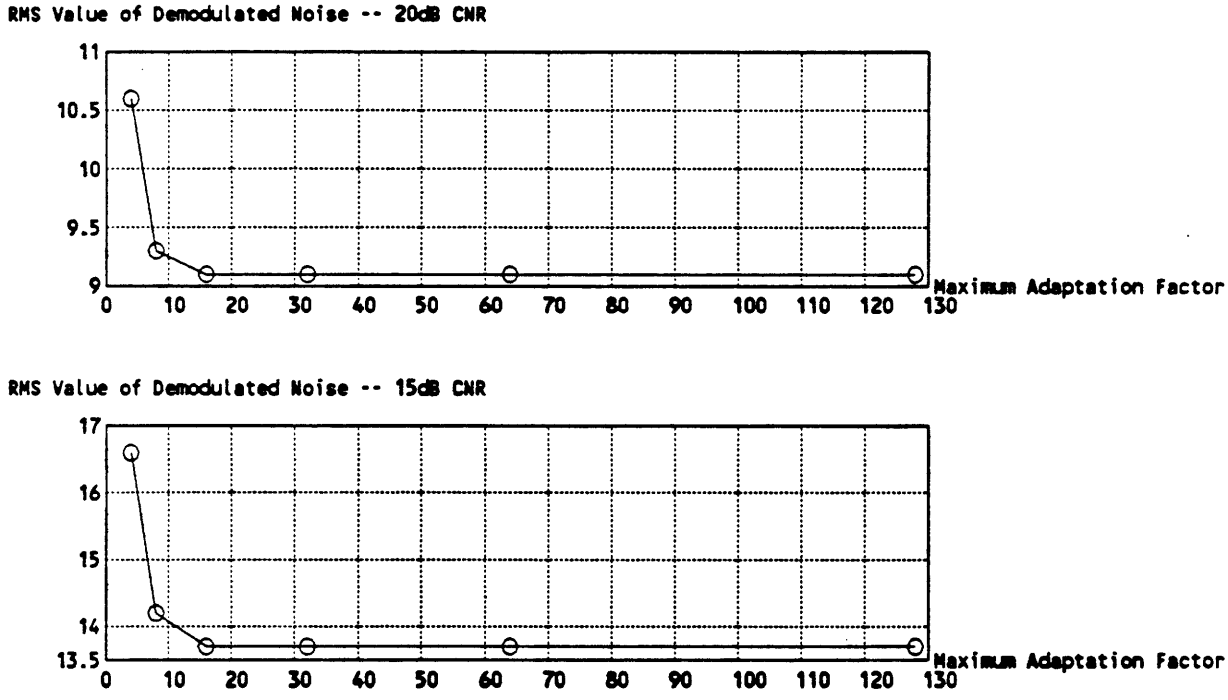


Figure 4.7: The rmse of the demodulated signal versus the maximum adaptation factor when white Gaussian noise (20dB CNR) is added in the channel.

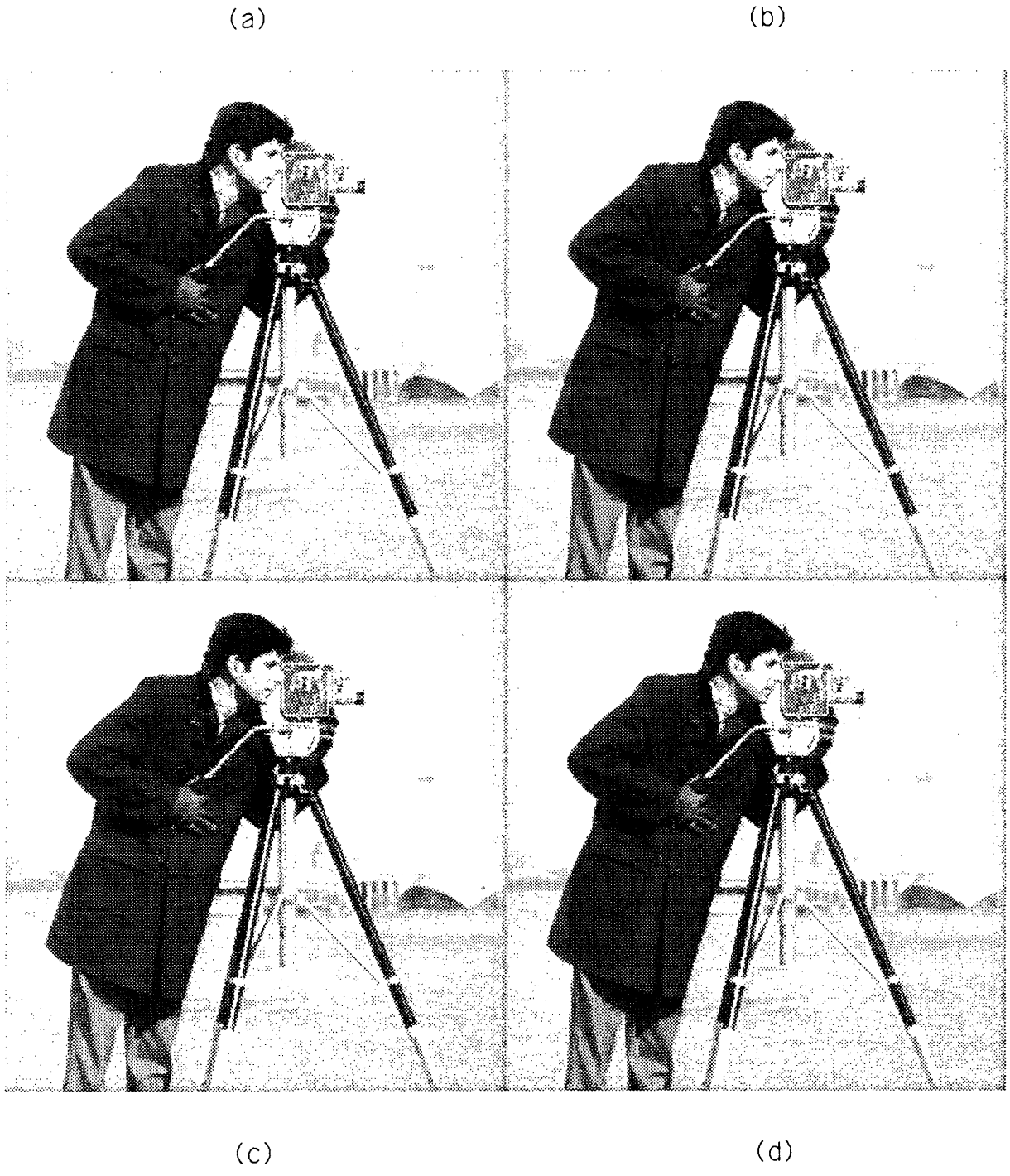


Figure 4.8: Demodulated pictures for four maximum values of the adaptation factors: (a) 8, (b) 16, (c) 32, and (d) 64. The channel is degraded by AWGN at a 20dB CNR.

Chapter 5

Interference Reduction and Channel Equalization

In this chapter, we look at the effect of and limitations of using adaptive modulation and scrambling to reduce interference and equalize a channel when the data is transmitted via pulse amplitude modulation at radio frequencies (RF). We cannot simply model the channel as a discrete-time linear system as in the previous chapter but must include the effects of modulation and demodulation at radio frequencies. Channel errors in the demodulated signal due to mistiming errors in resampling at the receiver, noise, co-channel and adjacent-channel interference, multipath, and frequency distortion at RF are discussed in this chapter. Let us first describe the transmission system that we will use and some of its parameters. Throughout this chapter, we shall use the rmse iterative method at 15% distortion to calculate the adaptation factors.

5.1 The Transmission System: AMSC-PAM

Pulse Amplitude Modulation (PAM) is used to transmit our digitized, adaptively modulated, and scrambled signal. Our modulation/transmission system is called, *Adaptively Modulated and Scrambled Pulse Amplitude Modulation* (AMSC-PAM). The transmission system is shown in Figure 5.1. We wish to transmit the discrete-time sequence, $y'[n]$. Assuming that the sequence of information samples, $y'[n]$, is wide-sense stationary

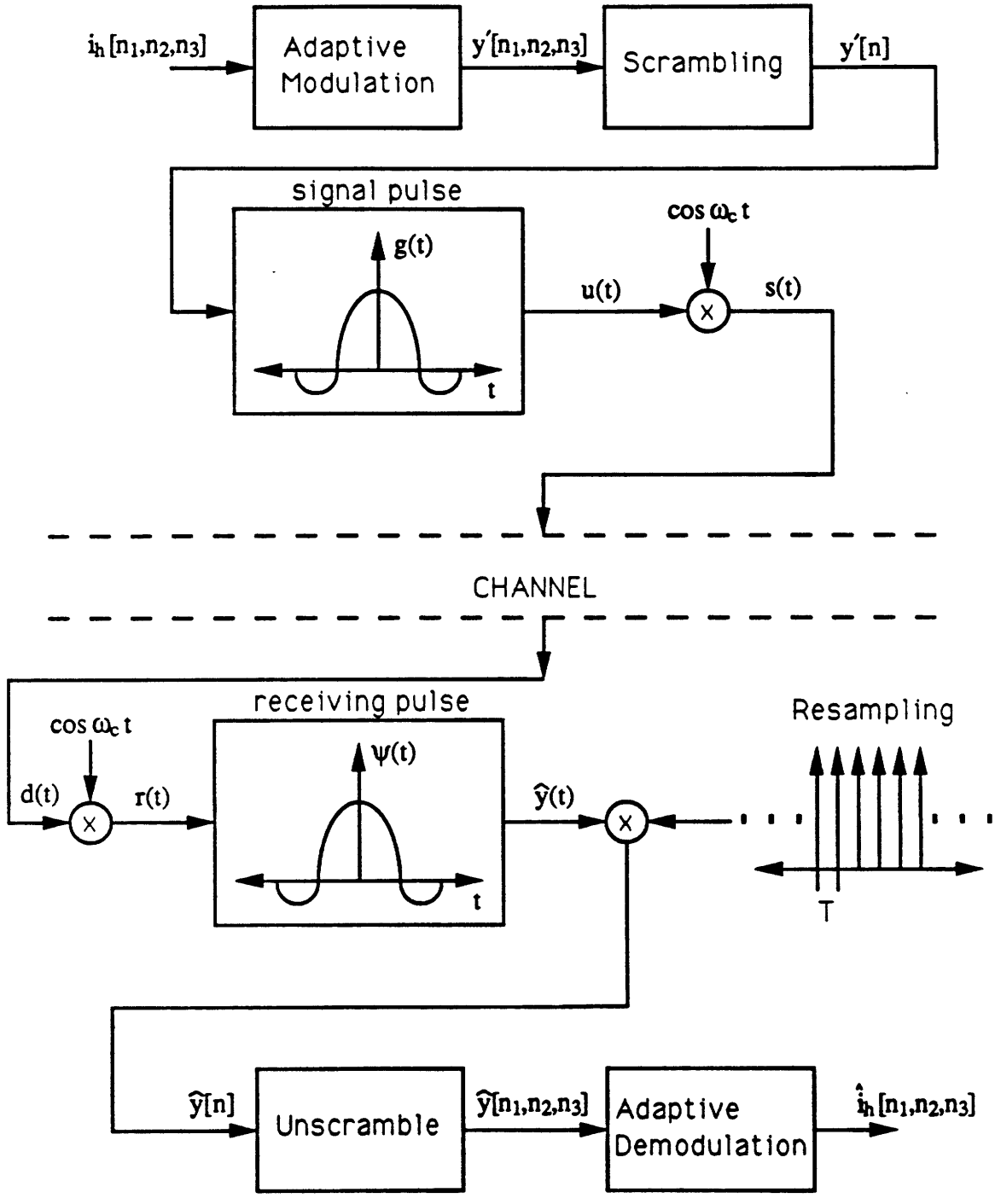


Figure 5.1: Block diagram of the AMSC-PAM transmission system.

with mean $\mu_{y'}$, then the autocorrelation function can be written as

$$\phi_{y'y'}[m] = E[y'[n]y'[n+m]]. \tag{5.1}$$

The discrete-time power spectrum of $y'[n]$ is periodic in 2π and is denoted by $\Phi_{y'y'}(\Omega)$, where Ω denotes discrete-time frequency space. The information bearing sequence, $y'[n]$, can be considered as the discrete-time representation of a wide-sense stationary continuous-time signal, y'_c, that has been sampled every T seconds to produce

$$y'_s(t) = y'_c(t)\delta_T(t), \tag{5.2}$$

where $\delta_T(t) = \sum_{m=-\infty}^{+\infty} \delta(t - mT)$ and $y'[n] = y'_s(nT)$.

Prior to modulation with a cosine, we convolve the sequence of samples with the transmission signal pulse, $g(t)$, which, for the purposes of this discussion, is assumed to have a bandlimited frequency response $G(\omega) = 0$ for $\mid \omega \mid > W$, where ω denotes continuous-time frequency space. The resulting signal is,

$$u(t) = \sum_{n=-\infty}^{+\infty} y'[n]g(t - nT). \tag{5.3}$$

The signal, $u(t)$, can also be written as

$$u(t) = \int_{-\infty}^{+\infty} g(\gamma)y'_s(t - \gamma)d\gamma. \tag{5.4}$$

This convolution transforms the discrete-time sequence into a continuous-time waveform and determines the spectral characteristics of the transmitted signal. To find out what the spectrum of this signal looks like, we need to calculate its autocorrelation function.

The autocorrelation function of $u(t)$ is

$$\begin{aligned}
\phi_{uu}(t+\tau, t) &= E[u^*(t)u(t+\tau)] \\
&= E[\int_{-\infty}^{+\infty} g^*(\gamma)y'^*_s(t-\gamma)d\gamma \int_{-\infty}^{+\infty} g(\nu)y'_s(t+\tau-\nu)d\nu] \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g^*(\gamma)g(\nu)E[y'^*_s(t-\gamma)y'_s(t+\tau-\nu)]d\gamma d\nu \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g^*(\gamma)g(\nu)E[y'^*_c(t-\gamma)y'_c(t+\tau-\nu)]\delta_T(t-\gamma)\delta_T(t+\tau-\nu)d\gamma d\nu \\
&= \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g^*(\gamma)g(\nu)\phi_{y'_cy'_c}(\tau-\nu+\gamma)\delta_T(t-\gamma)\delta_T(t+\tau-\nu)d\gamma d\nu. \qquad (5.5)
\end{aligned}$$

The autocorrelation function, $\phi_{uu}(t+\tau,t)$, is periodic in t with period T and is a cyclostationary random process [Pro83] [GF75]. The dependence of the autocorrelation function on t must be eliminated before we can compute the power spectrum of $u(t)$. We can do this by averaging $\phi_{uu}(t+\tau,t)$ over one period, so that

$$\begin{aligned}\phi_{uu}(\tau) &= \frac{1}{T}\int_{-T/2}^{T/2}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g^*(\gamma)g(\nu)\phi_{y'_c y'_c}(\tau-\nu+\gamma)\delta_T(t-\gamma)\delta_T(t+\tau-\nu)d\gamma d\nu dt \\ &= \frac{1}{T}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} g^*(\gamma)g(\nu)\phi_{y'_c y'_c}(\tau-\nu+\gamma)\delta_T(\tau-\nu+\gamma)d\gamma d\nu \\ &= \frac{1}{T}\phi_{y'_c y'_c}(\tau)\delta_T(\tau)\star g(\tau)\star g^*(-\tau), \end{aligned} \tag{5.6}$$

where $\star$ denotes convolution.

Taking the Fourier transform of the above equation yields,

$$\Phi_{uu}(\omega) = \frac{1}{T}\mid G(\omega)\mid^2 \sum_{n=-\infty}^{+\infty}\Phi_{y'_c y'_c}(\omega+n\omega_s), \tag{5.7}$$

where $\Phi_{y'_c y'_c}(\omega)$ is the power spectrum of $y'_c(t)$, and $\omega_s = 2\pi/T$ is the sampling frequency. Since $\phi_{y'y'}[m] = \phi_{y'_c y'_c}(mT)$ [Pro83], then

$$\sum_{n=-\infty}^{+\infty}\Phi_{y'_c y'_c}(\omega+n\omega_s) = \Phi_{y'y'}(T\omega), \tag{5.8}$$

where $\Phi_{y'y'}(\Omega)$ is the discrete-time power spectrum of $y'[n]$. Finally,

$$\Phi_{uu}(\omega) = \frac{1}{T}\mid G(\omega)\mid^2 \Phi_{y'y'}(T\omega), \tag{5.9}$$

which reveals the direct relationship between the spectral characteristics of the continuous-time signal, $u(t)$, and the discrete-time signal,$y'[n]$. The power spectrum, $\Phi_{y'y'}(T\omega)$, is periodic in $2\pi/T$; however, the signal pulse, $g(t)$ bandlimits the signal to produce the bandlimited signal $u(t)$. The spectral characteristics of $y'[n]$ determines the spectral characteristics of $u(t)$, and if $y'[n]$ is lowpass in nature, then $u(t)$ will also be lowpass. Likewise, if $y'[n]$ has a flat spectrum, then $u(t)$ will also have a flat spectrum (within the limits of $g(t)$.)

As an example, let's consider the characteristics of $y'[n]$ after scrambling. The sequence resembles a wide-sense stationary, zero-mean, white-noise process. Then,

$$\phi_{y'y'}[m] = \sigma^2_{y'}\delta[m], \tag{5.10}$$

and

$$\Phi_{y'y'}(\Omega) = \sigma_{y'}^2; \tag{5.11}$$

therefore,

$$\Phi_{uu}(\omega) = \frac{1}{T} \mid G(\omega) \mid^2 \sigma_{y'}^2. \tag{5.12}$$

In the chapter on scrambling, we observed that scrambling alters the spectral characteristics of a signal, but leaves the energy and peak value unchanged. Because of the direct relationship between the power spectrum of $u(t)$ and of $y'[n]$, we see that scrambling also alters the spectral characteristics of the continuous-time signal, $u(t)$, and leaves the energy and peak value unchanged. This is an important property because it means that we can make our transmitted signal less annoying to other signals by making it look like random noise without changing the energy of the signal. Another way of looking at it is that scrambling evens out the energy over all frequencies in the band. When adaptive modulation is used, the energy of the transmitted signal increases accordingly.

Modulating $u(t)$ by a cosine gives

$$\begin{aligned} s(t) &= u(t)\cos\omega_c t \\ &= \Big(\sum_{n=-\infty}^{+\infty} y'[n]g(t-nT)\Big)\cos\omega_c t, \end{aligned} \tag{5.13}$$

where ω_c is the carrier frequency. Suppose, for now, that the channel has an ideal frequency response characteristic, such that $C(\omega) = 1$ for all ω, and only additive white Gaussian noise degrades the signal. The signal, therefore, appearing at the receiver is

$$d(t) = (u(t) + n(t))\cos\omega_c t, \tag{5.14}$$

where $n(t)$ represents the equivalent lowpass noise, which is assumed to be a zero-mean, white-Gaussian-noise process. Synchronous demodulation at the receiver produces

$$\begin{aligned} r(t) &= (u(t) + n(t))\cos^2\omega_c t \\ &= \frac{1}{2}(u(t) + n(t)) + \frac{1}{2}(u(t) + n(t))\cos 2\omega_c t. \end{aligned} \tag{5.15}$$

The optimal receiving filter is a filter matched to the received pulse

$$\psi(t) = \int_{-\infty}^{+\infty} g(\tau)c(t-\tau)d\tau, \tag{5.16}$$

where $c(t)$ is the equivalent baseband impulse response of the channel. That is, the frequency response of the receiving filter is $\Psi^*(\omega)$. In the present case, $c(t)$ is the unit impulse, since we are assuming an ideal channel, and the receiving filter has a frequency response of $G^*(\omega)$.

Filtering $r(t)$ with the receiving filter eliminates the components at $2\omega_c$ leaving

$$\begin{aligned} \hat{y}(t) &= (u(t)+n(t)) \star g^*(-t) \\ &= \sum_{n=-\infty}^{+\infty} y'[n]x(t-nT) + v(t), \end{aligned} \tag{5.17}$$

where the 1/2 has been discarded, and $x(t)$ is the response of the receiving filter to the input pulse, $\psi(t)$. In this discussion, $x(t)$ has the spectral characteristic $X(\omega) =\mid G(\omega) \mid^2$. The signal $v(t)$ represents the response of the channel noise to the receiving filter.

Sampling the signal $\hat{y}(t)$ at times $t = kT$ for $k = \ldots -1, 0, 1 \ldots$ and, ignoring propagation delay, yields

$$\begin{aligned} \hat{y}[k] = \hat{y}(kT) &= \sum_{n=-\infty}^{+\infty} y'[n]x(kT-nT) + v(kT) \\ &= \sum_{n=-\infty}^{+\infty} y'[n]x[k-n] + v[k] \\ &= y'[k] + \sum_{n=-\infty \,\&\, n\neq k}^{+\infty} y'[n]x(kT-nT) + v(kT), \end{aligned} \tag{5.18}$$

where $x[0]$ is an arbitrary scaling factor that we set to unity. The term $y'[k]$ represents the desired sample at time kT and the term

$$\sum_{n=-\infty \,\&\, n\neq k}^{+\infty} y'[n]x(kT-nT) \tag{5.19}$$

represents the intersymbol interference resulting from the filtering process. In order for there to be no intersymbol interference, we want

$$x(t=kT) = x[k] = \begin{cases} 1 & \text{for } k = 0 \\ 0 & \text{for } k \neq 0 \end{cases} \tag{5.20}$$

For a bandlimited $x(t)$, such that $X(\omega) = 0$ for $\mid \omega \mid > W$, and when the sample rate of $y'[n]$ is at the Nyquist rate ($\frac{1}{T} = \frac{2W}{2\pi}$), then the pulse, x(t), which produces no intersymbol interference is

$$x(t) = \frac{\sin(\pi t/T)}{(\pi t/T)}. \tag{5.21}$$

Two problems exist with this pulse shape. The first is that this pulse is unrealizable. The second has to do with mistiming errors when resampling at the receiver. The tails of $x(t)$ decay as $1/t$; hence, a mistiming error in sampling results in an infinite series of intersymbol interference that is not absolutely summable and, therefore, does not converge. These two problems can be avoided if we restrict the sample rate to $\frac{1}{T} < \frac{2W}{2\pi}$ samples per second, where W is still the bandwidth of the signal pulse, $g(t)$, and the spectrum of $y'_s(t)$ is periodic with period $\frac{2\pi}{T}$.

With the sample rate less than the Nyquist rate, it is possible to not only have the tails of $x(t)$ decay at a faster rate but to not have intersymbol interference. One pulse widely used in the telephony business [LM88] is the raised cosine spectral characteristic,

$$X(\omega) = \begin{cases} T, & 0 \leq \mid \omega \mid \leq \frac{\pi(1-\beta)}{T} \\ \frac{T}{2}(1 - \sin T\frac{(\frac{\omega}{2} - \frac{\pi}{2T})}{\beta}), & \frac{\pi(1-\beta)}{T} \leq \mid \omega \mid \leq \frac{\pi(1+\beta)}{T}, \end{cases} \tag{5.22}$$

where β is called the *rolloff parameter.* The pulse associated with this spectrum is

$$x(t) = \frac{\sin(\pi t/T)}{(\pi t/T)} \frac{\cos(\beta \pi t/T)}{(1 - 4\beta^2 t^2/T^2)}. \tag{5.23}$$

When using this pulse shape, the receiving pulse is just $g(t)$. For demonstrative purposes and without losing any generality, we choose $\beta = 0.25$ throughout the rest of this chapter. The actual parameter chosen depends upon the nature of the transmission medium and is a parameter to be determined by the system designer. A larger β means that less intersymbol interference will result for a given mistiming error; however, a larger β also means that fewer samples per second can be transmitted in a given bandwidth W. With $\beta = 0.25$, we can transmit 9.6M samples per second when $W = 2\pi \cdot 6\text{MHz}$.

Figure 5.2 shows how this pulse shape allows recovery of the sequence $y'[n]$ at the receiver without intersymbol interference. After resampling at the receiver, $\hat{Y}(\omega)$ is replicated every $\frac{2\pi}{T}$, and because $X(\omega)$ has a raised cosine characteristic, the aliasing cancels, and we regain the original spectrum of $y'_s(t)$. Now that we have explained the evolution of the signals in the modulation and demodulation process, and have chosen a pulse shape and rolloff factor, let us look at what happens to the signal, $y'[n]$, when channel defects exist.

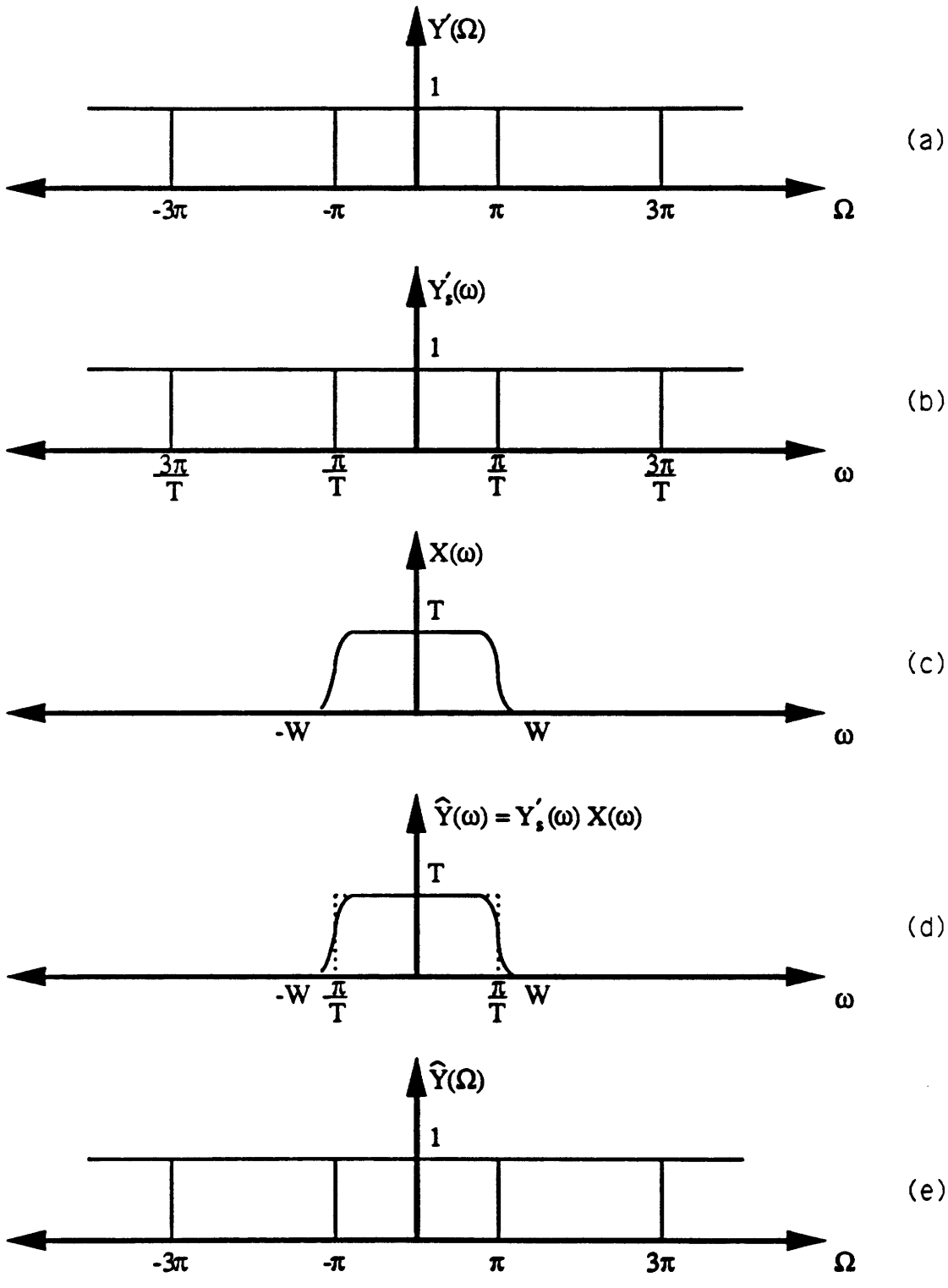


Figure 5.2: Filtering action of the pulse $x(t)$ on the sequence $y'[n]$.

5.2 Mistiming Errors

Resampling at the receiver needs to occur in phase with the nulls of $x(t)$ in order to recover the input samples, $y'[n]$, exactly under ideal channel conditions. If the transmission delay has been compensated for, then resampling needs to occur at kT for $k = \ldots -1, 0, 1 \ldots$. Any deviation from these time instances will result in mistiming errors that show up as intersymbol interference. Suppose the demodulated signal, before resampling, is

$$\hat{y}(t) = \sum_{n=-\infty}^{+\infty} y'[n]x(t - nT), \tag{5.24}$$

where no degradations have occurred in the channel. If resampling occurs at time $t = kT + \tau_o$ for $0 \leq \tau_o < T$ then the received samples are

$$\hat{y}[k] = \hat{y}(kT + \tau_o) = \sum_{n=-\infty}^{+\infty} y'[n]x(kT - nT + \tau_o). \tag{5.25}$$

Mistiming errors will occur for values of $\tau_o \neq 0$. writing the above equation in discrete notation gives

$$\hat{y}[k] = \sum_{n=-\infty}^{+\infty} y'[n]\hat{x}[k - n], \tag{5.26}$$

where

$$\hat{x}[k] = x(kT + \tau_o) = \frac{\sin[\pi(kT + \tau_o)/T)]}{\pi(kT + \tau_o)/T} \frac{\cos[\beta\pi(kT + \tau_o)/T]}{1 - [4\beta^2(kT + \tau_o)^2/T^2]}. \tag{5.27}$$

If we rewrite eqn. 5.26 as

$$\hat{y}[k] = \hat{x}[0]y'[k] + \sum_{n=-\infty \,\&\, n\neq k}^{+\infty} y'[n]\hat{x}[k - n], \tag{5.28}$$

then $y'[k]$ is the desired sample and

$$\sum_{n=-\infty \,\&\, n\neq k}^{+\infty} y'[n]\hat{x}[k - n] \tag{5.29}$$

represents the intersymbol interference due to the error, τ_o, in timing at the receiver when resampling.

We performed experiments to determine the amount of mistiming error our modulation system can tolerate by producing demodulated pictures for various values of the fraction, τ_o/T. Figure 5.3 shows the demodulated pictures for four values of τ_o/T: 0.05,

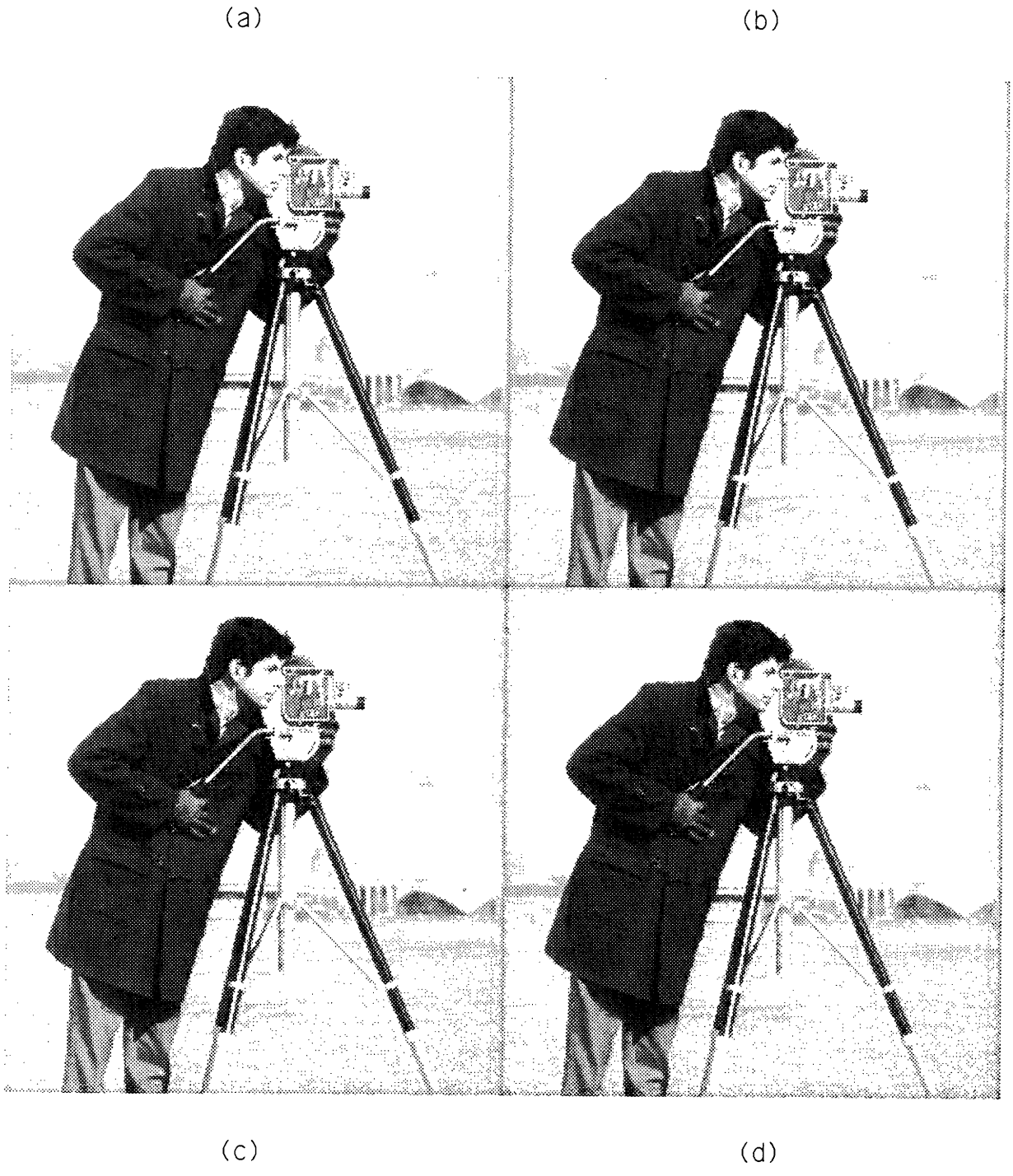


Figure 5.3: The demodulated pictures for four values of τ_o/T: (a) 0.05, (b) 0.1, (c) 0.15, and (d) 0.2.

0.1, 0.15, and 0.2. The largest fraction which still produces imperceptible distortion is 0.12. This picture along with the demodulated pictures when neither adaptive modulation nor scrambling is used, when only adaptive modulation is used, and when only scrambling is used are shown in Figure 5.4. The mistiming fraction is 0.12 in all cases. Because of the high degree of correlation between samples when neither adaptive modulation nor scrambling is used and when only adaptive modulation is used, mistiming errors produce little blur.

5.3 Noise

If we assume that we have chosen the proper signal pulse, $x(t)$, no timing errors exist, and only additive white Gaussian noise degrades the channel, then

$$\hat{y}[k] = y'[k] + v[k], \tag{5.30}$$

where

$$v[k] = \int_{-\infty}^{+\infty} n(t)g(kT - t)dt \tag{5.31}$$

and $n(t)$ is a white-Gaussian-noise process. Figure 5.5 shows the demodulated pictures at 4 levels of noise: 15dB, 20dB, 25dB, and 30dB CNR. At a CNR of 25dB, the degradation in the demodulated picture is imperceptible (see Figure 5.6.) Also shown is the demodulated picture that results when no adaptive modulation is used and when the same CNR exists in the channel.

5.4 Co-Channel Interference

Utilizing the spectrum efficiently requires that at least two sources transmit their signals using the same carrier frequency, but the sources should be sufficiently far apart geographically from each other so that the interference from the undesired signal does not severly degrade the picture quality of the desired signal. The relative strengths required of the two signals at a receiver to reproduce the desired signal at a certain picture quality will determine the distance that the two sources must be apart in order to provide service

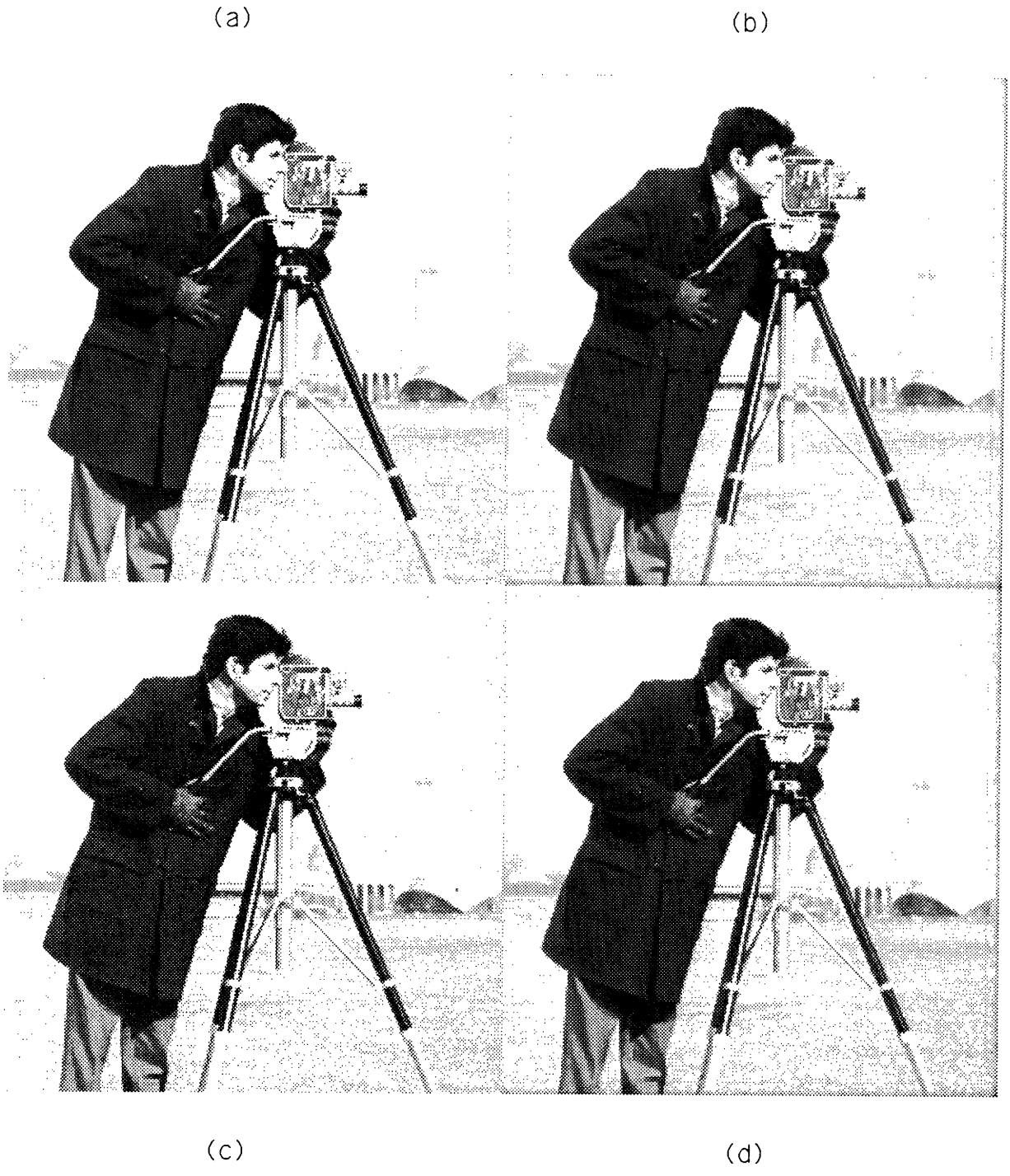


Figure 5.4: A comparison of four demodulated pictures that suffer from mistiming errors: (a) when neither adaptive modulation nor scrambling is used, (b) when only adaptive modulation is used, (c) when only scrambling is used, and (d) when both adaptive modulation and scrambling are used. The mistiming error fraction is 0.12 in all cases.

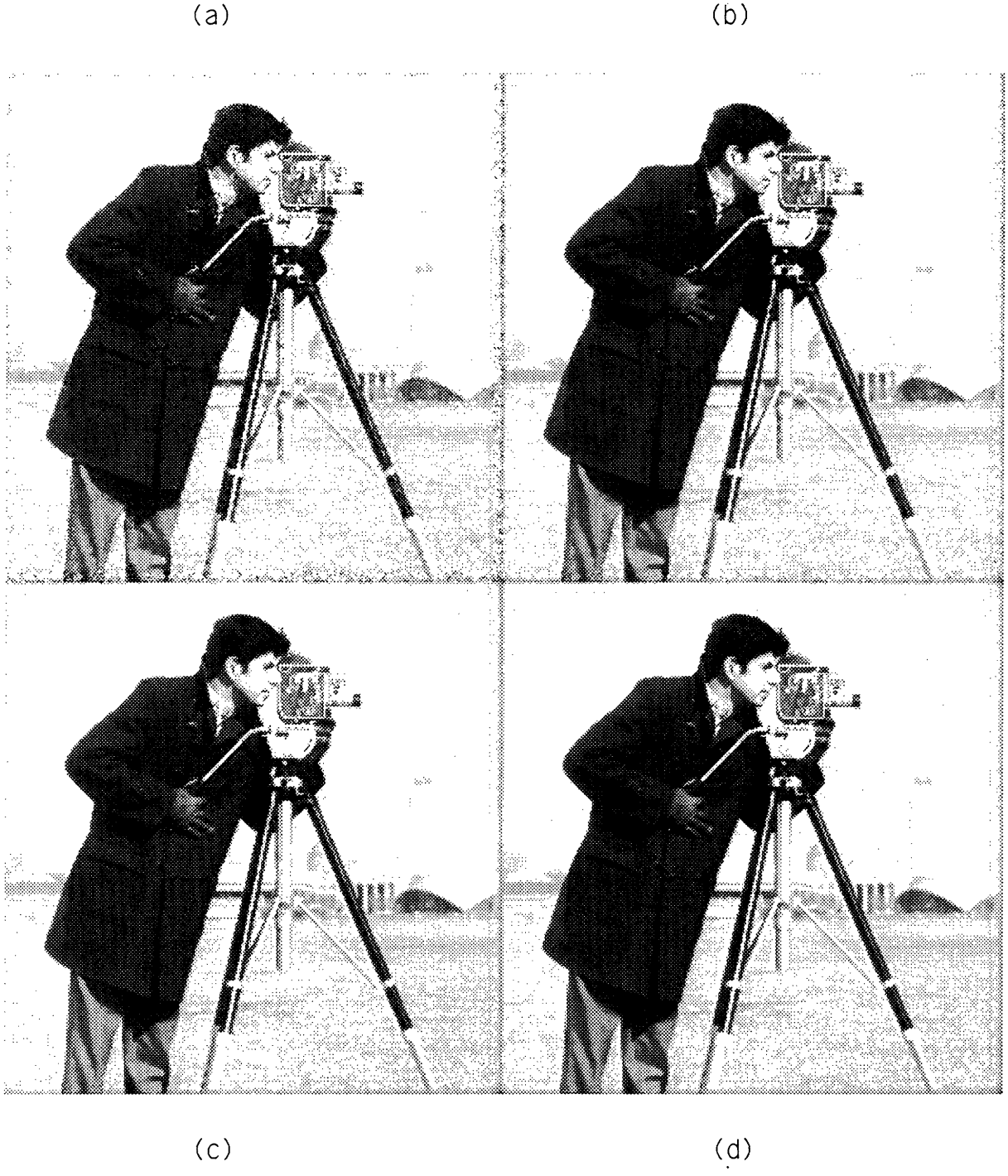


Figure 5.5: Demodulated pictures at 4 levels of noise: (a) 15dB, (b) 20dB, (c) 25dB, and (d) 30dB CNR.

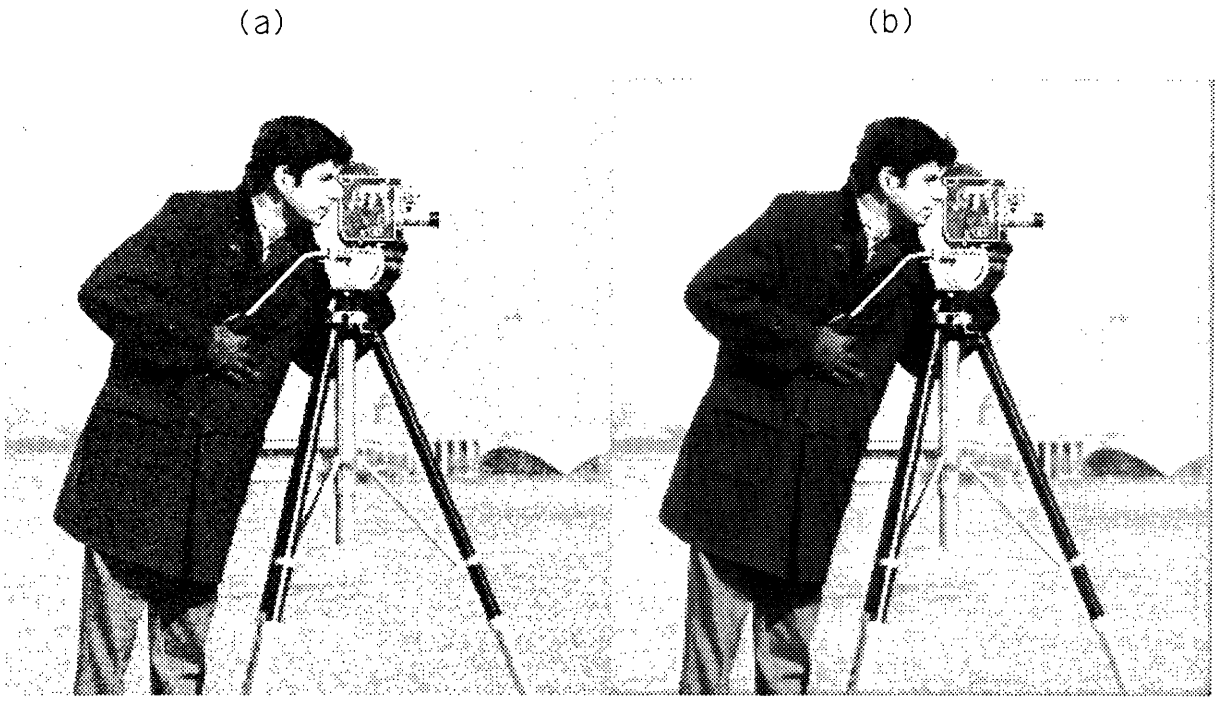


Figure 5.6: Demodulated picture at a CNR of 25dB. Picture (a) results when no adaptive modulation is used, and picture (b) results when both adaptive modulation and scrambling are used.

of that quality. This ratio of the desired signal strength to the undesired signal strength at a receiver is termed the D/U ratio. The smaller the required D/U ratio, then the closer the sources can be to one another, which means that more sources can transmit in a given geographical area on the same carrier. Said another way, smaller D/U ratios mean that, for the existing source locations, the area of service can be increased.

In the early years of color television, studies were performed on co-channel and adjacent-channel interference both in the United States and in Europe [FB60] [Dea60] [AML63]. Each study found that having a small frequency offset in the carrier frequencies between stations that transmit with the same nominal carrier frequency reduces the visibility of the interference of the undesired signal in the desired signal. For example, station A may transmit at ω_c and station B may transmit at $\omega_c + \omega_\Delta$, where ω_Δ is the frequency offset. The studys performed by Dean [Dea60] in 1960 for the Television Allocations Study Organization (TASO) explored 6 values for the offset frequency: 360, 604, 9985, 10010, 19995, and 20020 Hz. Offset frequencies of 360 and 604 Hz produced the most objectionable interference and require D/U ratios of 22dB and 44dB, respectively,

to produce a picture of passable picture quality or better. Offset frequencies of 10010 and 200020 Hz produced the least objectionable interference, and each required a D/U ratio of 17dB to produce a picture of passable picture quality or better and 31dB to produce a picture of excellent picture quality. The appearance of the interfering signal for these four offset frequencies are shown in Figure 5.7. Figure 5.8 shows a picture that has been interfered with at a D/U ratio of 17dB by another source that is offset by 10010 Hz.

In our transmission system, the signal in the channel is,

$$\begin{aligned} s(t) &= u(t)\cos\omega_c t \\ &= [\sum_{n=-\infty}^{+\infty} y'[n]g(t-nT)]\cos\omega t. \end{aligned} \tag{5.32}$$

If another signal $u_1(t)$, which is modulated at the same nominal carrier frequency, interferes with our desired signal, $s(t)$, then the signal appearing at the receiver is

$$d(t) = u(t)\cos\omega_c t + \alpha u_1(t)\cos(\omega_c + \omega_\Delta)t, \tag{5.33}$$

where α is a scaling factor. Synchronous demodulation yields,

$$\begin{aligned} r(t) &= d(t)\cos\omega_c t \\ &= u(t)\cos^2\omega_c t + \alpha u_1(t)\cos(\omega_c+\omega_\Delta)t\cos\omega_c t \\ &= \frac{1}{2}u(t) + \frac{1}{2}u(t)\cos 2\omega_c t + \frac{1}{2}\alpha u_1(t)\cos(2\omega_c+\omega_\Delta)t + \frac{1}{2}\alpha u_1(t)\cos\omega_\Delta t. \end{aligned} \tag{5.34}$$

Filtering at the receiver with $g(t)$ and ignoring the $\frac{1}{2}$ gives

$$\begin{aligned} \hat{y}(t) &= u(t)\star g(t) + (\alpha u_1(t)\cos\omega_\Delta t)\star g(t) \\ &= y'_s(t)\star g(t)\star g(t) + (\alpha y'_{s_1}(t)\star g(t))\cos\omega_\Delta t\star g(t). \end{aligned} \tag{5.35}$$

The evolution of the spectrum of $(u_1(t)\cos\omega_\Delta t)\star g(t)$ is shown in Figure 5.9.

Using an offset frequency reduces the energy of the interfering signal by half; therefore, we shall use offset frequencies in our transmission system. We shall use the same offset frequency of 10010 Hz that is used in NTSC, so that the degrading effects of interference from our signal into NTSC is minimized.

With $\omega_\Delta = 2\pi\cdot 10010$ Hz, we must find α such that the degradations in the resampled signal $\hat{y}[n]$, due to the interference from $u_1(t)$ are imperceptible. Figure 5.10 shows the

Figure 5.7: Interfering pictures when the offset frequency is: (a) 360 Hz, (b) 604 Hz, (c) 10010 Hz, and (d) 20020 Hz.

Figure 5.8: Interference when an undesired signal has been offset by 10010 Hz and is at a D/U ratio of 17dB.

demodulated desired signal at four values of α: 1.0, 0.5, 0.25, and 0.125. The largest value of alpha that still produces imperceptible degradation is 0.25. This corresponds to a D/U ratio of 12dB. With this D/U ratio, Figure 5.11 shows the demodulated desired pictures that result when neither adaptive modulation nor scrambling is used, when only adaptive modulation is used, and when only scrambling is used. The case when adaptive modulation and scrambling are used is also shown. When NTSC is used for transmission, a D/U ratio of 31dB is required to produce excellent pictures. When AMSC-PAM is used, a D/U ratio of only 12dB is required to produce excellent pictures.

5.5 Adjacent-Channel Interference

Adjacent-channel interference is essentially co-channel interference when $\omega_\Delta = 2W$ in eqn. 5.35. In his studies on NTSC, Dean [Dea60] found that a D/U ratio of -27dB produces pictures of passable quality or better and -10dB produces pictures of excellent quality. For our system, we must find α such that the degradations in the resampled signal $\hat{y}[n]$, due to adjacent-channel interference from $u_1(t)$ are imperceptible. Figure 5.12

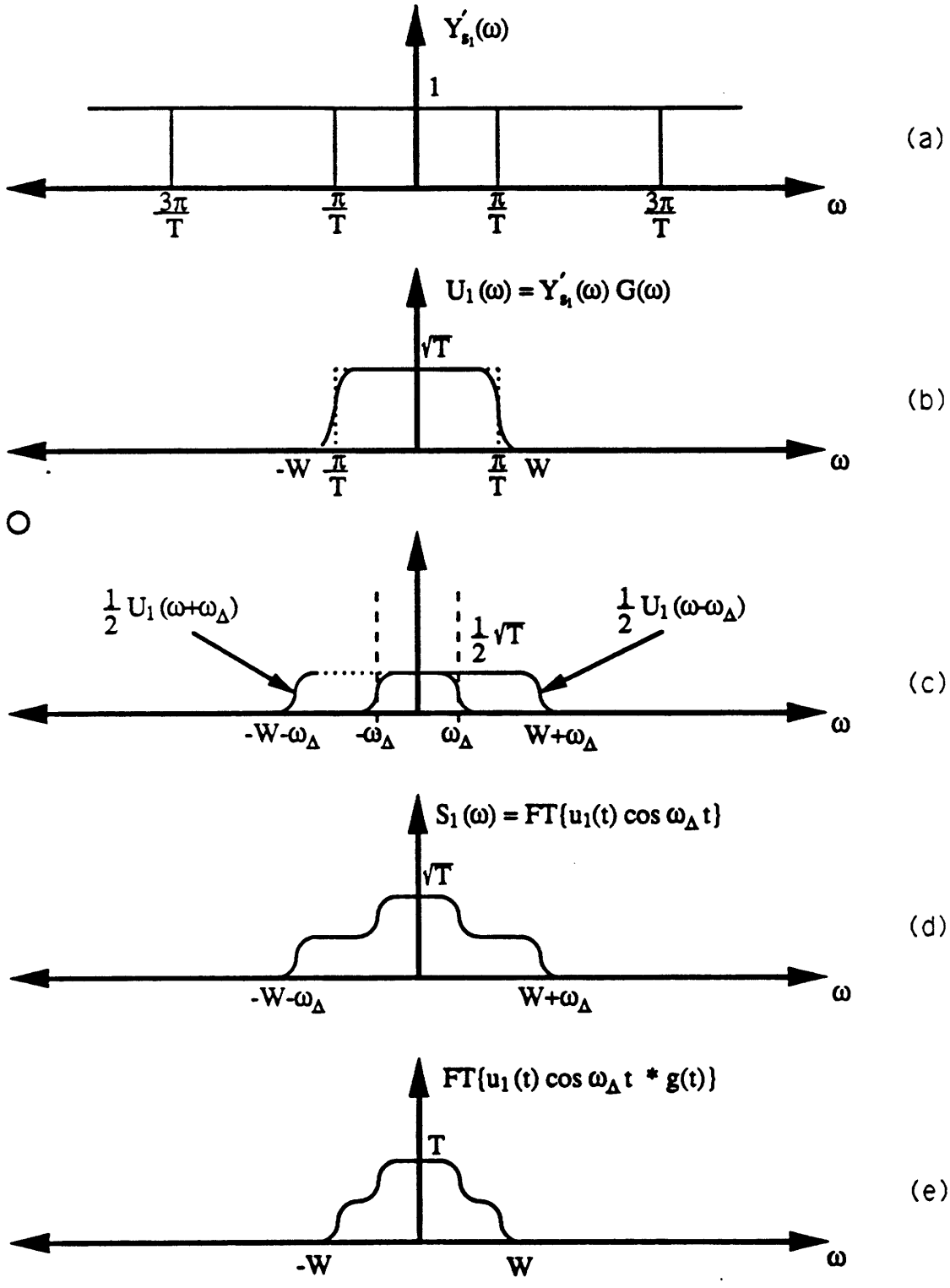


Figure 5.9: The evolution of the specturm of $(u_1(t)\cos\omega_\Delta t) \star g(t)$. Part (c) shows the shifting action caused by multiplication with $\cos\omega_\Delta t$.

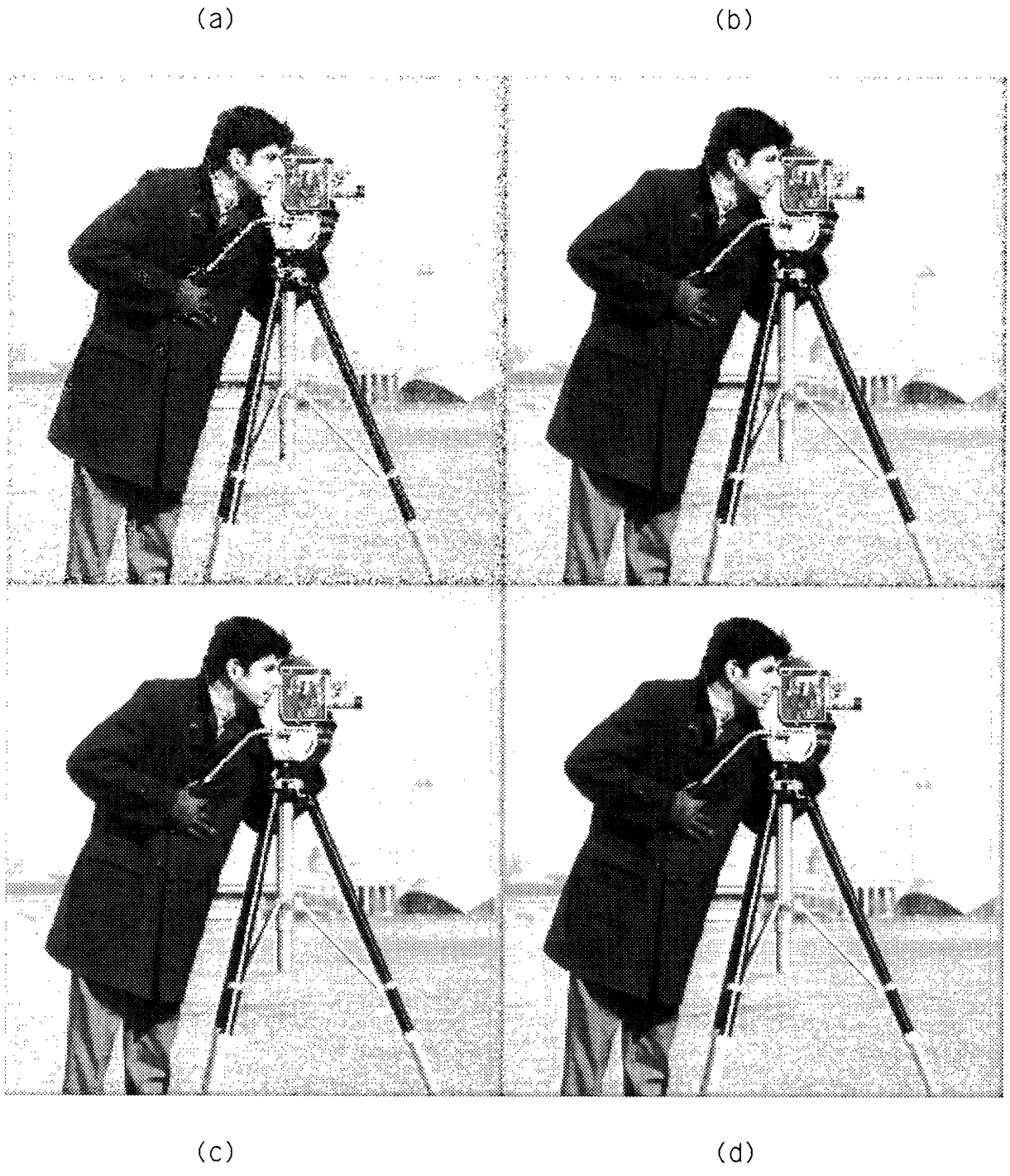


Figure 5.10: The demodulated signal when co-channel interference exists in the channel using AMSC-PAM for four values of α: (a) 1.0, (b) 0.5, (c) 0.25, and (d) 0.125.

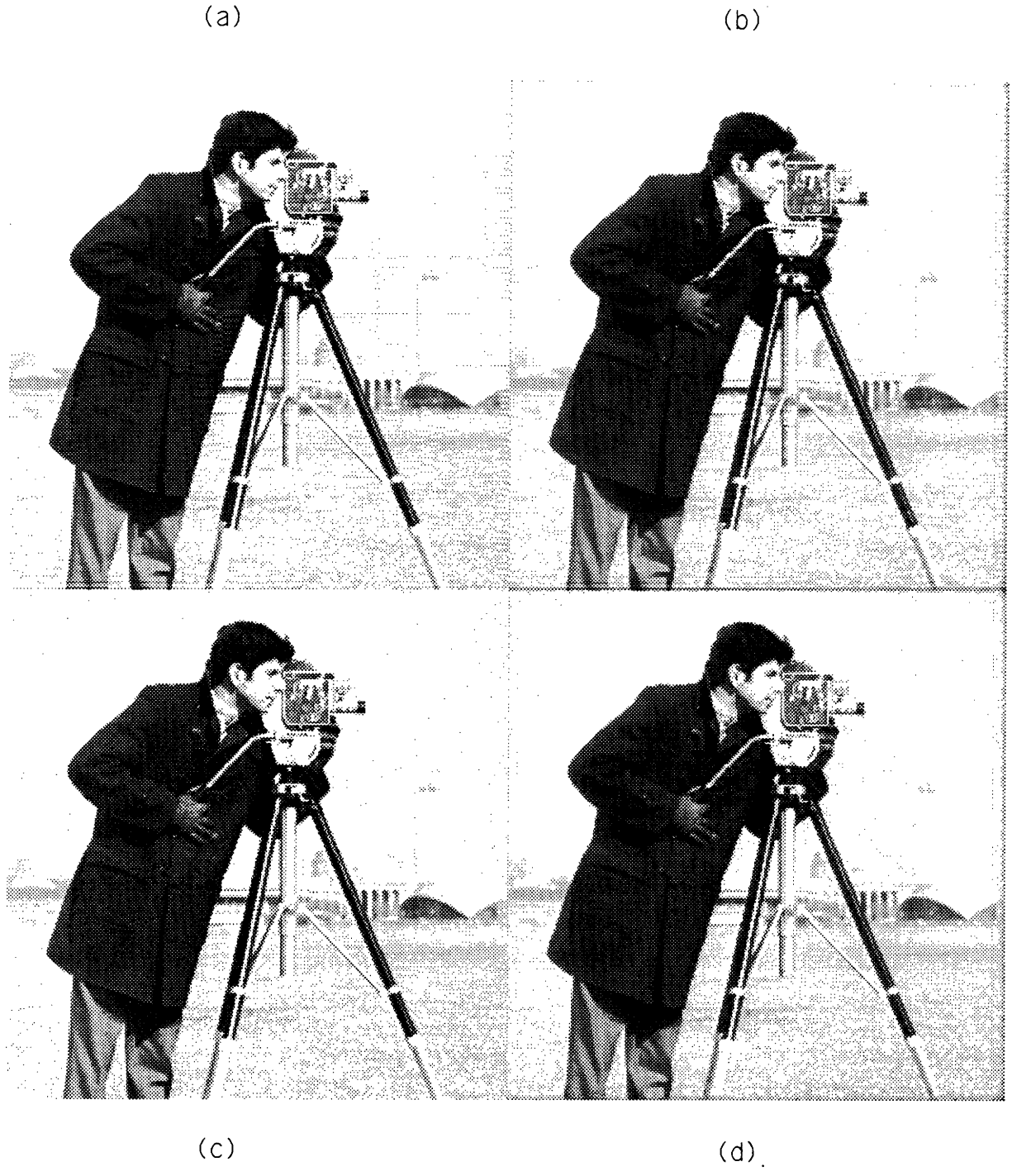


Figure 5.11: The appearance of co-channel interference at a 12dB D/U ratio using PAM when (a) neither adaptive modulation nor scrambling is used, (b) only adaptive modulation is used, (c) only scrambling is used, and (d) both are used.

shows the demodulated desired signal at four values of α: 100, 50, 35, and 10. The largest value of alpha that still produces imperceptible degradation is 35. This corresponds to a D/U ratio of -31dB. With this D/U ratio, Figure 5.13 shows the demodulated desired pictures that result when neither adaptive modulation nor scrambling is used, when only adaptive modulation is used, and when only scrambling is used. The case when adaptive modulation and scrambling are used is also shown. When NTSC is used for transmission, a D/U ratio of -10dB is required to produce excellent pictures. When AMSC-PAM is used, a D/U ratio of only -31dB is required to produce excellent pictures.

5.6 Multipath and Frequency Distortion

The signal in the channel is

$$s(t) = u(t)\cos\omega_c t. \tag{5.36}$$

If one source of multipath exists in the channel, then the signal appearing at the receiver is

$$d(t) = u(t)\cos\omega_c t + u(t-t_o)\cos\omega_c(t-t_0). \tag{5.37}$$

Synchronous demodulation gives

$$\begin{aligned} r(t) &= d(t)\cos\omega_c t \\ &= u(t)\cos^2\omega_c t + u(t-t_o)\cos\omega_c(t-t_o)\cos\omega_c t \\ &= \frac{1}{2}[1+\cos 2\omega_c t]u(t) + \frac{1}{2}[1+\cos 2\omega_c t]\cos\omega_c t_o u(t-t_o) \\ &\quad +\frac{1}{2}\sin 2\omega_c t \sin\omega_c t_o u(t-t_o). \end{aligned} \tag{5.38}$$

The filtering operation at the receiver with $g(t)$ will eliminate the $\cos 2\omega_c t$ and $\sin 2\omega_c t$ components, leaving (ignoring the $\frac{1}{2}$)

$$\begin{aligned} \hat{y}(t) &= u(t)\star g(t) + \cos\omega_c t_o u(t-t_o)\star g(t) \\ &= \sum_{n=-\infty}^{+\infty} y'[n]x(t-nT) + \cos\omega_c t_o \sum_{n=-\infty}^{+\infty} y'[n]x(t-t_o-nT). \end{aligned} \tag{5.39}$$

Sampling at kT, we get

$$\hat{y}[k] = y'[k] + \cos\omega_c t_o \sum_{n=-\infty}^{+\infty} y'[n]x(kT-nT-t_o). \tag{5.40}$$

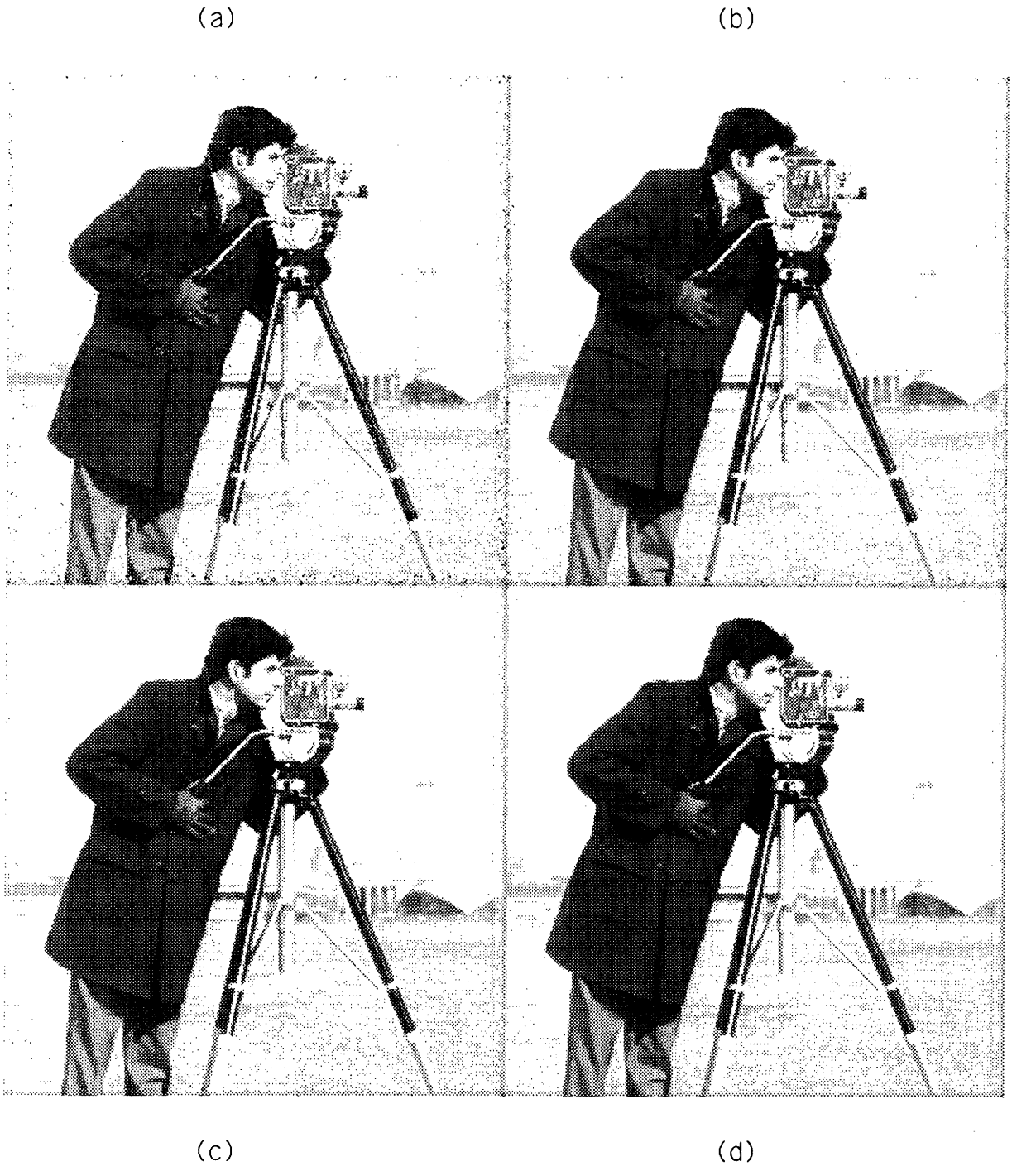


Figure 5.12: The demodulated signal when adjacent-channel interference exists in the channel using AMSC-PAM for four values of α: (a) 100, (b) 50, (c) 35, and (d) 10.

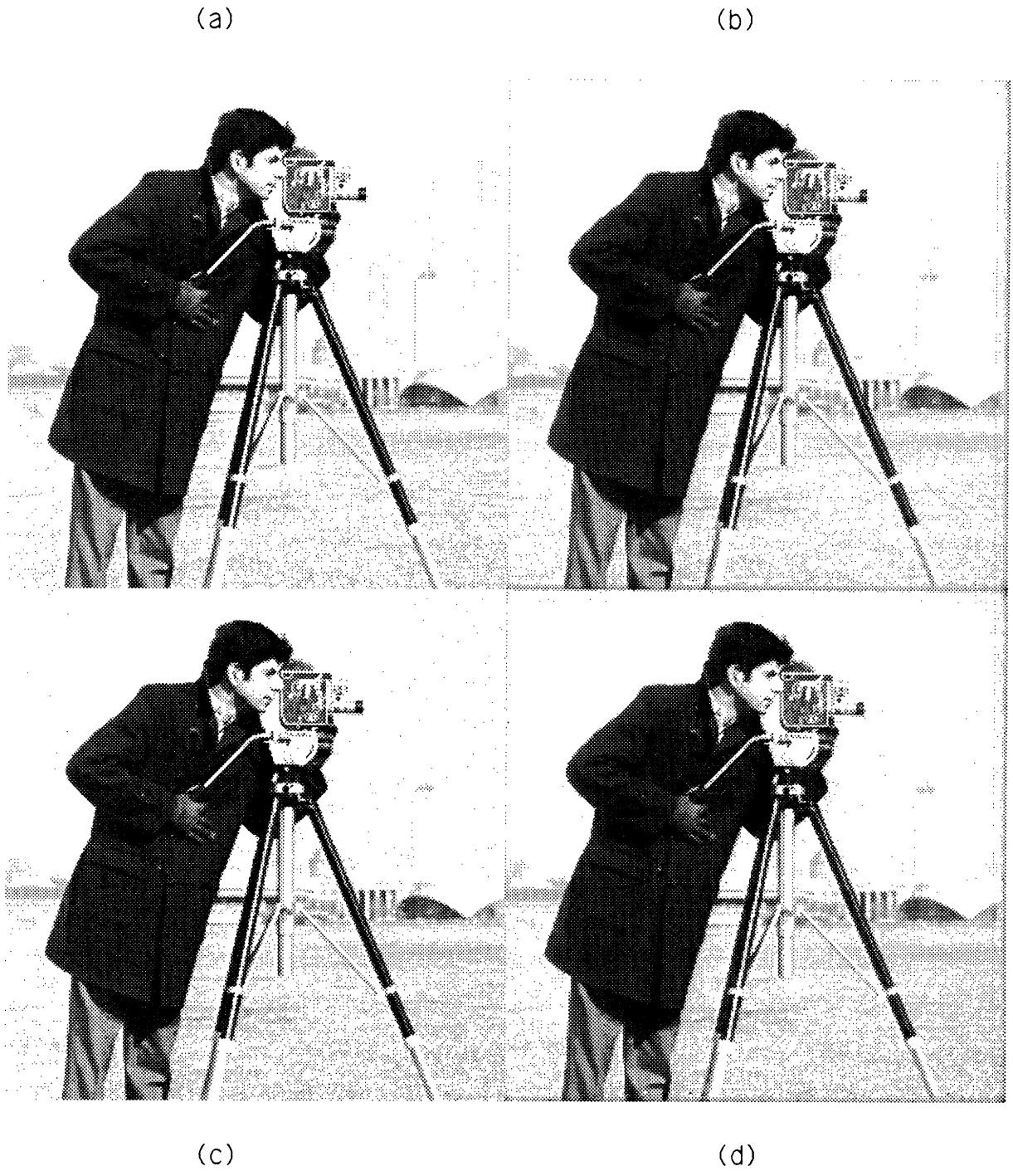


Figure 5.13: The appearance of adjacent-channel interference at a -31dB D/U ratio using PAM when (a) neither adaptive modulation nor scrambling is used, (b) only adaptive modulaton is used, (c) only scrambling is used, and (d) both are used.

The second term represents the attenuated intersymbol interference that is added to the desired sample, $y'[k]$.

If the channel is not ideal ($C(\omega) \neq 1$ for all ω) and frequency distortion exists, then the signal pulse seen at the receiver (eqn. 5.16) is

$$\psi(t) = \int_{-\infty}^{+\infty} g(\tau)c(t-\tau)d\tau. \tag{5.41}$$

Since the receiving pulse is matched to $g(t)$, then the sample value at the receiver will be

$$\hat{y}[k] = \sum_{n=-\infty}^{+\infty} y'[n]x(kT - nT), \tag{5.42}$$

where $x(t) = \psi(t) \star g(t)$ and does not satisfy eqn. 5.20. By expressing $\hat{y}[n]$ in the form

$$\hat{y}[k] = x(0)y'[k] + \sum_{n=-\infty \,\&\, n\neq k}^{+\infty} y'[n]x(kT - nT), \tag{5.43}$$

we see that $y'[k]$ is the desired sample and the second term represents the intersymbol interference.

Both multipath and frequency distortion result in intersymbol interference, and the variance of the intersymbol interference can be calculated using properties of the sums of independent identically distributed random variables. This means that the variance of the intersymbol interference is some multiple of $\sigma_{y'}^2$, and all that we need to know in order to determine the greatest amount of intersymbol interference that is still imperceptible is the amount of $y'[n]$ that can be added to itself in an uncorrelated manner and still remain imperceptible. This experiment entailed adding $\alpha y'[n-l]$, for $l = 20$, to $y'[n]$ and finding the largest α that still produces imperceptible distortion. Figure 5.14 shows the demodulated signal for four values of α: 0.4, 0.3, 0.2, and 0.15. The largest value α that produces no perceptual distortion is 0.2. This corresponds to a variance of $\alpha^2\sigma_{y'}^2 = 0.04\sigma_{y'}^2$. With $\alpha = 0.2$, Figure 5.15 shows the demodulated pictures that result when neither adaptive modulation nor scrambling is used, when only adaptive modulation is used, when only scrambling is used, and when both are used.

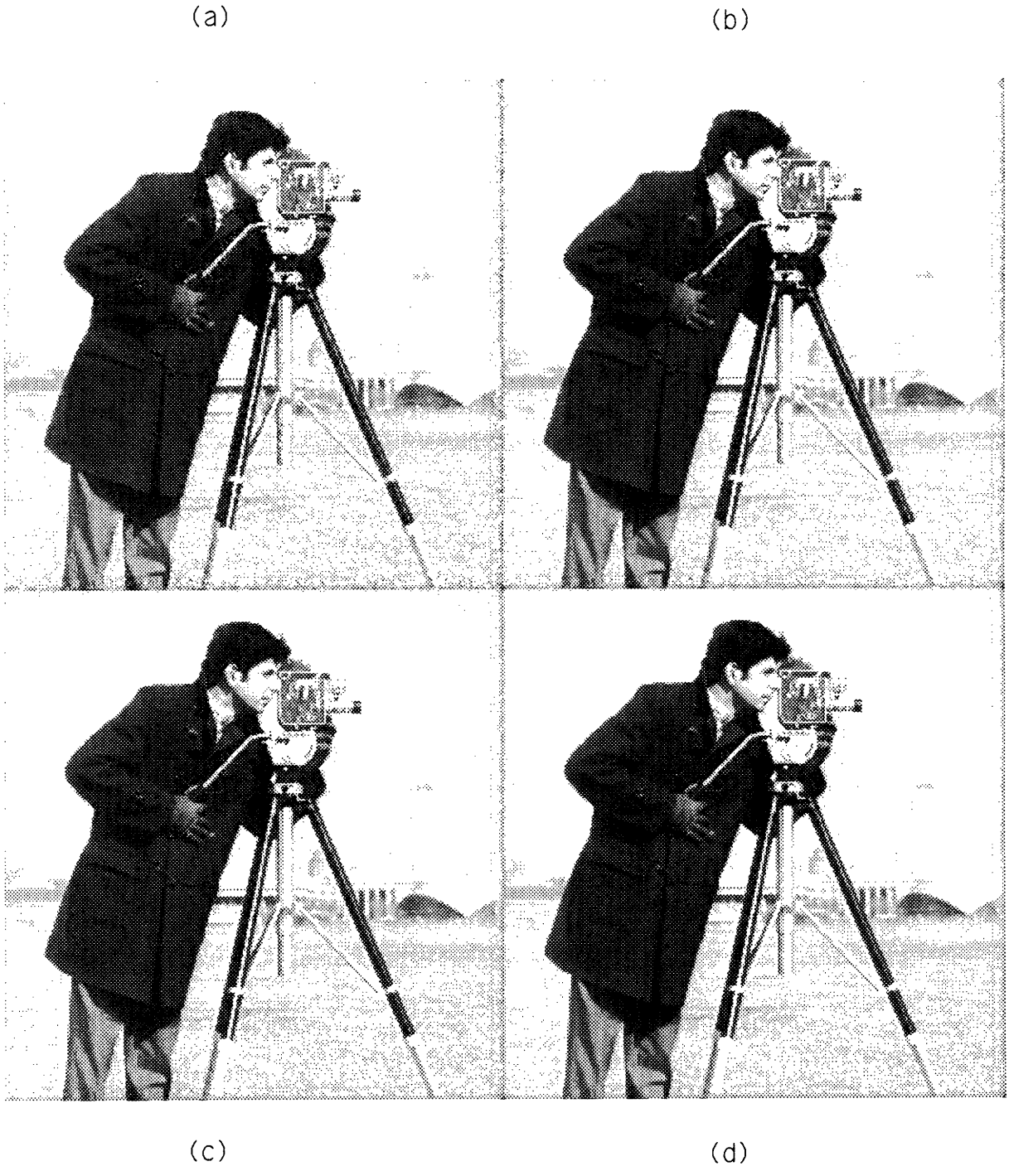


Figure 5.14: The demodulated signal when multipath exists in the channel using AMSC-PAM for four values of α: (a) 0.4, (b) 0.3, (c) 0.2, and (d) 0.15.

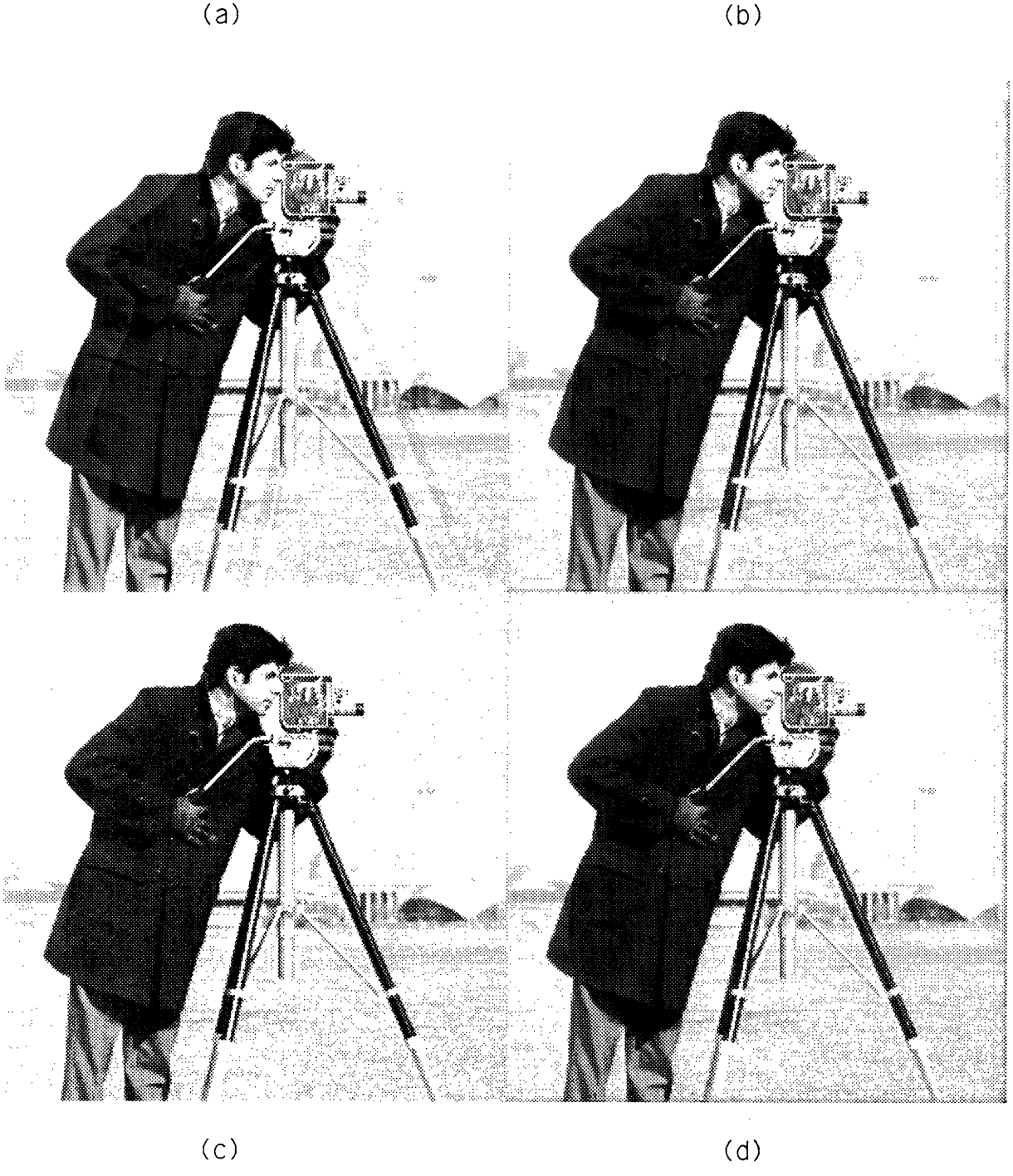


Figure 5.15: The appearance of 20% multipath using PAM when (a) neither adaptive modulation nor scrambling is used, (b) only adaptive modulation is used, (c) only scrambling is used, and (d) both are used.

Chapter 6

Conclusion

In the preceding chapters, we have presented a new method of interference reduction and channel equalization that is composed of adaptive amplitude modulation in conjunction with scrambling, and we have separately described each component in detail. Scrambling causes channel defects (such as co-channel and adjacent-channel interference, multipath, frequency distortion, and other forms of intersymbol interference) to manifest themselves in the demodulated image as degradations that have the appearance of additive random noise. Adaptive amplitude modulation reduces the appearance of these degradations, of additive channel noise, and of any other additive degradations, thus, equalizing the channel and reducing interfence. the combination of adaptive amplitude modulation and scrambling produces a significantly improved image at the receiver.

The first topic in our discussion focused on finding pairs of filters and cutoff frequencies to separate the input signal into a 3-D highpass and a 3-D lowpass frequency component. The goal was to minimize the amount of lowpass data that needs to be transmitted digitally while producing a highpass unit-step response whose rate of decay matches that of the visual masking effect. The cutoff frequency in each dimension determines the decimation factors of the lowpass component and, thus determines the amount of digital information that must be sent to represent the lowpass frequency component. We restricted ourselves to using Gaussian prefilters and sharpened-Gaussian postfilters. In selecting the filter pairs, attention was paid to the amplitude of the sampling structure and the magnitude of the overshoot in the highpass unit-step rewponse. A large sampling

structure produces annoying fluctuations in the noise that is passed in the blank areas of the demodulated image, and a large overshoot produces a "halo" of noise near edges. The amplitude of the sampling structure was limited to ± 1, and the magnitude of the overshoot was limited to 5.

We chose a Gaussian prefilter with a standard deviation of 5.0, a sharpened-Gaussian postfilter with a standard deviation of 4.5, and a decimation factor of 8 to perform the spatial band separation of a 512x512 image. We also chose an averaging prefilter, a linear postfilter, and a decimation factor of 8 to perform the temporal band separation of a 60 frame-per-second sequence. If we only perform a spatial frequency decomposition with a decimation factor of 8, then 245,760 samples of lowpass data per second need to be sent. The band-separation filters and the highpass and lowpass images produced from these filters are shown in Section 2.1.

The amount of side information required to implement adaptive amplitude modulation is small because only one digital sample of information is transmitted per $N_1 \text{x} N_2 \text{x} N_3$ block of highpass data. If 4x4 spatial blocks are used for a 512x512, 60 fps sequence, then 983,040 samples of side information per second need to be sent digitally. The existing adaptive amplitude modulation algorithm put forth by Schreiber and Buckley [SB81] uses the maximum absolute value in an $N_1 \text{x} N_2 \text{x} N_3$ block of highpass data as the transmitted sample. These samples are linearly interpolated, and then this interpolated signal is inverted and scaled by a constant to produce the adaptation factors. The adaptation factors are multiplied by the highpass component to produce the adaptively modulated signal. Under noisy channel conditions ($<$ 25dB CNR) this algorithm does not sufficiently suppress noise around edges.

We improved the performance of adaptive modulation by prefiltering the absolute value of the highpass component with a Gaussian filter, subsampling by N_1, N_2, and N_3 in the respective dimensions, and interpolating with a sharpened-Gaussian filter. This interpolated signal is used to calculate the adaptation factors, and the subsample associated with each block of highpass data is transmitted to the receiver. Prefiltering with a Gaussian filter and interpolating with a sharpened-Gaussian filter increases the rates of decay of the interpolated signal near edges and improves noise reduction near

edges. Our experiments focused on 4x4 spatial blocks and found that a Gaussian prefilter of standard deviation 1.5 and a sharpened-Gaussian postfilter of standard deviation 2.0 produced the best looking pictures.

We can increase noise reduction near edges by raising the values of the adaptation factors or, equivalently, by lowering the value of the interpolated signal that is used to calculate the adaptation factors near edges; however, this will introduce distortion near edges in the demodulated image under ideal channel conditions. The distortion will appear as a softening of the edges. We found the optimal tradeoff between noise reduction and distortion for a given anticipated CNR in the channel by scaling the interpolated signal with a constant factor and performing subjective tests. This scaling factor controls the amount of distortion and noise reduction in the demodulated signal. The results of this investigation are in Section 2.4. For example, the interpolated signal, which results from using the above Gaussian filters, should be scaled by 1.2 for a 20dB-CNR channel. In addition, for any level of distortion, using the Gaussian filters with a scaling factor provides a 1.0 dB SNR improvement over using Schreiber and Buckley's method with a scaling factor.

Another way to increase noise reduction near edges is to more strictly control the amount of distortion in the demodulated image by allowing only a specified amount of distortion to exist in each block of highpass data. The sample which is to be transmitted for a block of highpass data and that yields this level of distortion is found by performing an iterative Secant search over the distortion in the block. For a given distortion measure, the level of distortion is based on a percentage of a corresponding metric of the highpass data. We considered three distortion measures: root-mean-squared error, peak-absolute error, and mean-absolute error. The corresponding three metrics of the highpass data are the root-means-squared value, the peak-absolute value, and the mean-absolute value. the root-mean-squared-error distortion measure allowed the most noise reduction for a given level of picture quality and, hence, was chosen as our distortion measure. With this distortion measure, the maximum level of distortion that is still imperceptible to just perceptible is 15%; therefore, with a rmse distortion measure and a 15% distortion level, the demodulated image under ideal channel conditions has insignificant perceptual

differences from the original image.

If we compare Schreiber and Buckley's method (without scaling) of calculating the adaptation factors to the above iterative method and to the method of using Gaussian filters with a scaling factor for still images, we find that when using 2-D blocks, the iterative method provides a 2.7dB SNR improvement and the Gaussian method provides a 3.5dB SNR improvement over Schreiber and Buckley's method. The demodulated images that result from the iterative method, however, retain their sharpness; hence, the iterative method produces better looking images than the Gaussian method.

Three options exist when applying the iterative method to sequences. One can simply treat each frame as a still picture and perform the iterative search using 2-D blocks on each frame separately, or one can perform the iterative search using 2-D blocks on each frame but pass the state of the search from frame to frame. The third method uses 3-D blocks in the iterative search. when 3-D blocks are sued, the amount of motion becomes an important factor. Noise may exist far from spatial luminance changes within a frame when the motion is great (> 1 pel per frame); therefore, visual masking effects will not adequately mask this noise when the human observer is tracking moving objects. when 2-D blocks are used, the amount of motion is not a factor. Unfortunately, the 2-D iterative search algorithm cannot compensate quickly enough for changes in the distortion functions of the blocks when the state of the blocks is passed from frame to frame and, thus, requires too many iterations to produce images with the specified level of distortion. consequently, we found that performing the 2-D iterative search on each frame as if it were a separate, still image gave the best looking images.

Because the maximum adaptation factor determines the strength of the pseudorandom noise that gets added to the demodulated picture when intersymbol interference or interference from other adaptively modulated sources exists, we had to find the value of the maximum allowable adaptation factor that provides a balance between the visibility of noise in the busy areas and the amount of noise reduction in the blank areas. A maximum value of 16 produces the best balance. In addition, there appears to be no perceptual advantage to using larger maximum adaptation factors when additive white Gaussian noise degrades the channel.

Defect	CNR
Mistiming Errors	24.4
Additive White Gaussian Noise	25.0
Co-Channel Interference	26.5
Adjacent-Channel Interference	25.7
Multipath and Frequency Distortion	23.4

Table 6.1: Table of defect and its maximum equivalent CNR for imperceptible to just perceptible degradation in the demodulated images.

The next and most important step was to apply adaptive amplitude modulation and scrambling to reducing interference in and equalizing an RF channel. In these experiments, we looked at mistiming errors, additive white Gaussian noise, co-channel and adjacent-channel interference, multipath, and frequency distortion. We found the maximum level of each defect that still produces imperceptible to just perceptible degradation in the demodulated images. In Table 6.1 we list each type of defect and its maximum equivalent CNR for imperceptible to just perceptible degradation in the demodulated images. From this table we see that there is a fairly consistent level of degradation that our equalization and interference reduction method can make imperceptible to just perceptible. the maximum level is approximately at a 25dB CNR. This level is independent of the type of degradation.

We have presented in this thesis a new channel equalization and interference reduction method and a robust transmission system that codes a signal so that it can better withstand channel defects. AMSC-PAM transmission provides greatly improved images at the receiver, in comparison to ordinary pulse amplitude modulation systems, without having

to raise the peak transmission power. AMSC-PAM can also be used in conjunction with conventional channel equalization methods. Unlike conventional equalization methods which employ adaptive filters, our equalization and interference reduction scheme does not require long convergence times to find the filter coefficients, does not require complex calculations at the receiver, is not limited to a maximum length of the channel impulse response that it can equalize – our method can equalize a channel with an impulse response of any duration, – and does not require the transmission of training sequences. Our method, however, is limited by the power of the channel defects and is more sensitive to mistiming errors in resampling at the receiver than transmission methods that do not employ scrambling. In addition, digital side information, which is composed of lowpass frequency information and adaptation information, must be sent to the receiver.

When a video signal is decomposed into spatio-temporal frequency subbands via block transforms or quadrature mirror filtering and is transmitted in analog form, it becomes very important to guard against channel noise. Methods exist [Cho90] which apply adaptive modulation to the subbands; however, the tradeoff between distortion and noise reduction has not been investigated when modulating subband data. In addition, further work needs to be performed on the interaction and application of adaptive amplitude modulation and scrambling to subband data.

As one last example of AMSC-PAM, we degraded the channel with co-channel interference at a 14dB D/U ratio, a 15% echo at 20 pels, and additive white Gaussian noise at a 28dB CNR. If conventional PAM is used, we would receive the picture in Figure 6.1, and if AMSC-PAM is used, we would receive the picture in Figure 6.2.

Figure 6.1: The demodulated image when conventional PAM is used for transmission. The channel is degraded by co-channel interference (14dB D/U), multipath (15%), and additive white Gaussian noise (28dB CNR).

Figure 6.2: The demodulated image when AMSC-PAM is used for transmission. The channel is degraded by co-channel interference (14dB D/U), multipath (15%), and additive white Gaussian noise (28dB CNR).

Bibliography

[AML63] J. W. Allnatt, D. E. Mills, and E. R. Loveless. The subjective effect of co-channel and adjacent-channel interference in television reception. *Proceeding of the Institute of Electrical Engineers*, pages 109–117, June 1963.

[AO82] T. Abatzoglou and B. O'Donnell. Minimization by coordinate descent. *Journal of Optimization Theory and Applications*, 36(2):163–174, February 1982.

[Avr76] M. Avriel. *Nonlinear Programming: Analysis and Methods.* Prentice-Hall, 1976.

[Buc81] R. R. Buckley. *Digital Color Image Coding and the Geometry of Color space.* PhD thesis, MIT Dept. of Electr. Eng., June 1981.

[Cho90] W. H. Chou. Methods to improve spatio-temporal adaptive amplitude modulation for video transmission. Master's thesis, MIT Dept. of Electr. Eng., January 1990.

[Cra47] B. H. Crawford. Visual adaptation in relation to brief conditioning stimuli. *Proceedings of the Royal Society of London, Series B*, 134:283–302, 1947.

[Dea60] C. E. Dean. Measurements of the subjective effects of interference in television reception. *Proceedings of the IRE*, pages 1035–1049, June 1960.

[Dix84] R. C. Dixon. *Spread Spectrum Systems.* John Wiley and Sons, New York, 1984.

[FB60] B. L. Fredendall and W. L. Behrend. Picture quality - procedures for evaluating subjective effects of interference. *Proceeding of the IRE*, pages 1030–1034, June 1960.

[FJdF55] A. Fiorentini, M. Jeanne, and G. T. di Francia. Measurements of differential threshold in the presence of a spatial illumination gradient. *Atti. Fond. B. Ronchi*, 10:371–379, 1955.

[Fox] R. Fox. Visual masking. In *Handbook of Sensory Physiology*, vol. 8, ed. R. Held, H. W. Leibowitz, and H. L. Teuber, Springer, New York.

[GF75] W. A. Gardner and L. E. Franks. Characterization of cyclostationary random signal processes. *IEEE Transactions on Informatin Theory*, IT-21(1):4–14, January 1975.

[Gia78] R. S. Giansiracusa. Adaptive filtering for television ghost cancelation. Master's thesis, MIT Dept. of Electr. Eng., June 1978.

[Gir89] B. Girod. The information theoretical significance of spatial and temporal masking in video signals. Presented at SPSE/SPIE Symposium of Electronic Imaging, Juanuary 15-20 1989. Los Angeles, CA.

[JN84] N. S. Jayant and P. Noll. *Digital Coding of Waveforms: Principles and Applications to Speech and Video.* Prentice-Hall, Englewood Cliffs, New Jersey, 1984.

[Kre56] E. R. Kretzmer. Reduced-alphabet representation of television signals. *IRE Convention Record, Part 4*, pages 140–147, 1956.

[LM88] E. Lee and D. G. Messerschmitt. *Digital Communication.* Klumer Academic Publishers, 1988.

[LR67] R. W. Lucky and H. R. Rudin. An automatic equalizer for general-purpose communication channels. *Bell System Technical Journal*, 46(11):2179–2208, November 1967.

[LTKL80] F. Lukas, U. Tulunay-Keesey, and J. Limb. Thresholds at luminance edges under stabilized viewing conditions. *J. Opt. Soc. Am.*, 70(4):418–422, April 1980.

[Luc65] R. W. Lucky. Automatic equalization for digital communication. *Bell System Technical Journal*, 44(4):547–588, April 1965.

[Luc66] R. W. Lucky. Techniques for adaptive equalization of digital communication systems. *Bell System Technical Journal*, 45(2):255–286, February 1966.

[Lue84] D. G. Luenberger. *Linear and Nonlinear Programming.* Addison-Wesley, 1984.

[Orf85] S. J. Orfanidis. *Optimum Signal Processing: An Introduction.* Macmillan Publishing Co., New York, 1985.

[Pro83] J. G. Proakis. *Digital Communications.* McGraw-Hill Book Co., New York, 1983.

[Rat80] J. Ratzel. *The Discrete Representation of Spatially Continuous Images.* PhD thesis, Massachusetts Institute of Technology, June 1980.

[Rob62] L. G. Roberts. Picture coding using pseudo-random noise. *IRE Trans. on Information Theory*, 8:145–154, February 1962.

[SB81] W. F. Schreiber and R. R. Buckley. A two-channel picture coding system: II–adaptive companding and color coding. *IEEE Trans. on Communications*, COM-29(12):1849–1858, December 1981.

[Sch64] L. Schuchman. Dither signals and their effect on quantization noise. *IEEE Trans. on Communication Tech.*, pages 162–165, December 1964.

[Sch88] W. F. Schreiber et. al. Channel-compatible 6-mhz hdtv distribution systems. *SMPTE Television Conference*, Nashville, January 1988.

[SP88] W. F. Schreiber and J. Piot. Video transmission by adaptive frequency modulation. *IEEE Communications Magazine*, M-COM-26(11):68–76, November 1988.

[Spe65] G. Sperling. Temporal and spatial visual masking. i. masking by impulse flashes. *J. Opt. Soc. Am.*, 55:541–559, May 1965.

[ST85] W. F. Schreiber and D. E. Troxel. Transformation between continuous and discrete representations of images: A perceptual approach. *IEEE Trans. on Pat. Anal. and Mac. Int.*, PAMI-7(2):178–186, March 1985.

[Str86] G. Strang. *Introduction to Applied Mathematics.* Wellesley-Cambridge Press, 1986.

[Tro81] D. E. Troxel et. al. A two-channel picture coding system: I–real-time implementation. *IEEE Trans. on Communications*, COM-29(12):1841–1848, December 1981.

[Zar78] A. P. Zarembovitch. Forward estimation adaptive dpcm for image data compression. Master's thesis, MIT Dept. of Electr. Eng., May 1978.