Adaptive Frequency Modulation for Satellite Television Systems
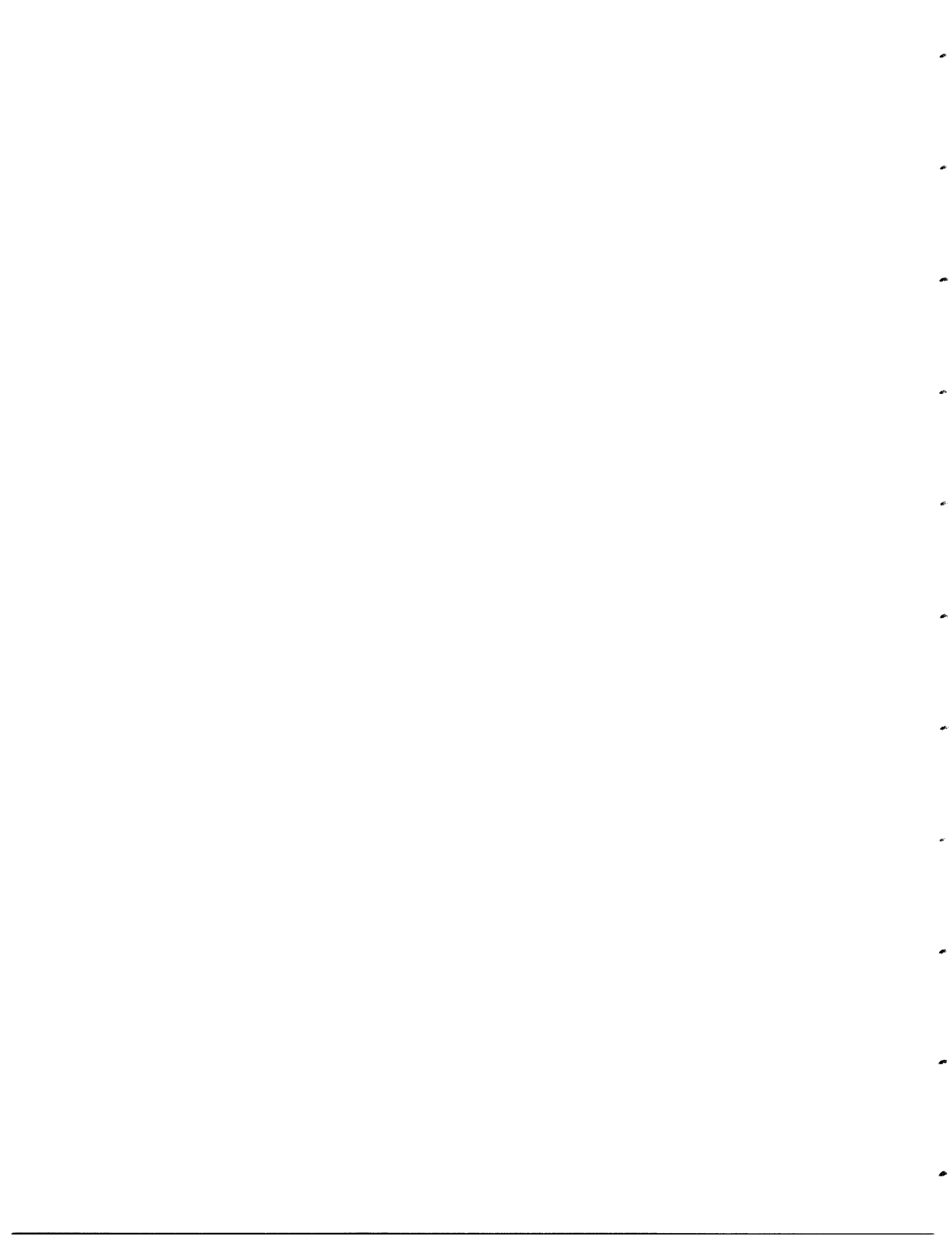
RLE Technical Report No. 554

April 1990

Julien Piot

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

Adaptive Frequency Modulation
for Satellite Television Systems

by

Julien Piot

Submitted to the
Department of Electrical Engineering and Computer Science
on January 10, 1990 in partial fulfillment of the requirements for the Degree of
Doctor of Philosophy in Electrical Engineering

Abstract. Frequency modulation is the first choice coding scheme in many existing applications such as satellite television transmission.

A simple model of image formation predicts large variations in the short-time bandwidth of the modulated signal. Based on this model, the frequency deviation in a small block of the picture is adjusted in order to keep the bandwidth constant. The resulting noise improvement is shown to be significant when a subjective measure of the transmission error is used. This measure is based on noise masking, whose average intensity is related to the block statistics.

In some applications the modulated signal is bandlimited, resulting in distortion in form of envelope and phase distortion. Both terms generate artifacts in the recovered picture, mostly when noise is present in the link. We show through measurements that the peak short-time bandwidth is related to the severity of the distortion, hence justifying the prior approach to adaptation. Improved algorithms are then introduced that minimize the transmission noise, but maintain negligible distortion.

When the information is transmitted in form of multiple components, such as sub-bands, a technique based on a sequential transmission is presented. This technique can also be used to transmit some side information, such as the adaptation function of the modulator. By adjusting the rate of transmission of the various components, the subjective impairment can be minimized. For example, a vertical subband decomposition is shown to be very effective in reducing transmission noise, with a larger improvement than preemphasis techniques. Alternatively, adaptive modulation can be combined with this technique.

Finally, the principle of adaptive frequency modulation is applied to broadcasting and distributing by satellite of television signals. A dual-in-one system, where two NTSC video signals are transmitted through one transponder is demonstrated. Another system proposes direct broadcasting by satellite of high definition television using subband coding and adaptive frequency modulation. A simulation of the two systems demonstrates that high quality transmission is possible in noisy narrow-band channels, using analog modulation.

Thesis Supervisor: Dr. William F. Schreiber

Title: Professor of Electrical Engineering

Contents

List of Figures

List of Tables

Acknowledgments

I wish to express my sincere gratitude to Professor Schreiber for his guidance, support, and encouragement. I am also grateful to my thesis readers, Prof. P. Humblet and Dr. A. Netravali, for their comments and suggestions. I want to thank all the members of the ATRP group for many interesting discussions and enjoyable moments.

Thanks also to my friends in Cambridge, New York, and Atlanta for very nice moments during my stay at M.I.T. In particular, Abeer was a great source of support, and also of fun. Her spirit and humor could not be replaced. Shukran. Thanks to Maurice and Philippe. The Swiss cheese fondue with Mont-sur-Rolle wine were certainly an essential ingredient in completing this thesis. I am also grateful to my family for their constant support and encouragement. Finally, I would like to thank Isabelle for her love and support.

To my parents

Chapter 1

Introduction

Since the National Television Systems Committee adopted the present television standard in 1953, no other coding scheme has been as widely used on the American continent. However, this decade is witnessing a surge of research activity in the field, due largely to the advances of the Japanese high definition television systems and its European and American reactions.

In this changing landscape, satellites are very likely to remain important for distributing and broadcasting television signals. Direct Broadcasting by Satellite (DBS) for example was recently introduced in Europe. Distribution by satellites of NTSC signals, on the other hand, is of great importance now, even for cable companies. The excitement produced by the coming of High Definition Television is motivating studies on how to do a better job in today's satellite systems and design new satellite systems for tomorrow's television.

Even though the trend in communication is to use digital transmission, frequency modulation, an analog modulation scheme, is often preferred for video satellite applications [Gag87].

Digital transmission involves source coding and channel coding. Source coding has the ability to remove most of the redundancy of the video signal, thus reducing the amount of information to be transmitted [JN84] [Ber71]. Channel coding is then required to map this digital information into signals suitable for transmission [Pro83].

However, the channels offered by satellites usually have a very poor signal-to-noise ratio, and a large number of error correction bits must be added in order to communicate reliably [Gal68], thus partially defeating the purpose of source coding [MMHY87].

Frequency modulation, on the other hand, is quite adequate for video applications [Rho85]. Like digital coding, it has the ability to exchange bandwidth with distortion, but its distortion very nicely matches the human visual sensitivity making it much less annoying than quantization noise and other digital artifacts. Of course, the price one pays using analog modulation is the impossibility to regenerate the signal perfectly after transmission, a typical feature of digital coders. For broadcasting and distribution purposes, this is not a major problem as long as this noise is below the level of visibility. This is amply justified since only one unique transmission is going to occur between the transmitting station and the viewers homes.

The performances of a FM coder can be greatly improved by exploiting other properties of the human visual system such as noise masking or reduced chrominance resolution used so far only in digital coders. Reduction of as much as 25 dB of noise perception has been achieved by a digital coder using the masking effect [SB81]. Another form of digital data compression was achieved by relying on the frequency-dependent noise sensibility of the eye using subband coding and DPCM [WO86].

Similar large improvements in the efficiency of analog channel utilization can be achieved by digitally assisting the frequency modulator without losing the advantages of analog communication.

Soft Versus Hard Threshold

Shannon showed that a reliable communication is possible if the channel noise is below a certain level [Gal68]. When the noise is increased above that, an abrupt change of performance occurs, and it is no longer possible to communicate.

This effect, called threshold effect, is also encountered in nonlinear modulation, such as frequency modulation, pulse width modulation, etc. In the case of FM tele-

vision systems, it is manifested by impulse-like degradations, called clicks.

. When digital coding is performed in a satellite channel, forward error correction is used. Block codes or convolutional codes add some redundant information that is used to correct the bits degraded by noise. However, as predicted by Shannon, the performance of these coders drops abruptly when the noise exceeds a certain level. As a consequence, the satellite link is most of the time excessively good, so as to withstand extreme situations such as heavy rains. The additional channel capacity is wasted most of the time, a costly thing to do in terms of the extra power required in the satellite, or the larger antennas required.

Contrastingly, frequency modulation has a softer threshold, in that performance is still satisfactory even when the noise level is above the threshold. As a matter of fact, most of today's programming that uses a satellite link –many do– are affected by clicks. This soft threshold is a big advantage in that the design of the budget link is facilitated, and that no capacity is wasted (up to a certain extent) to handle extreme noise level.

Technological Constraints

Introducing a new modulation method requires a careful evaluation of the existing system constraints such as spectrum division. Spectrum planning for satellites has resulted in a division of the total capacity into smaller bands of fixed bandwidths. These bands can be of 24 MHz, 27 MHz, or 36 MHz, depending on the application. Any new coding scheme will have to fit into this framework, either by transmitting more information, such as enhanced television, in one band, or by putting two or more video signals in one band, without sacrificing quality.

An example of the latter is a dual-in-one coder, a system where two video channels are inserted into one 36 MHz transponder [MMHY87]. However the resulting quality has been suitable only to News Gathering. The digital coder proposed is also very complex.

17

1.1 Outline of the Thesis

The thesis consists of seven chapters, including the introduction and the conclud-ing remarks. They cover the different steps involved in the design of an adaptive modulation system.

Chapter 2 reviews some properties of frequency modulation. It is seen that the noise performance of the transmission link can be improved at the cost of an increased bandwidth. In order to examine this property in the context of image processing, a statistical model of pictures is presented. Finally, the effects of the human visual sys-tem are taken into account in order to define a subjective measure of the transmission noise.

Chapter 3 introduces adaptive frequency modulation as a technique to keep the short-time bandwidth of the modulated signal constant. In order to take into account the nonstationary nature of the picture, a more refined model is proposed, which uses the assumption of stationarity inside blocks of small dimensions. Based on this model, an algorithm to adjust the modulator is derived, and its performance is assessed experimentally. Finally, the subjective noise improvement is evaluated.

Chapter 4 refines the transmission link model in order to include the effects of bandwidth limitation on both the modulated and the demodulated signal. It appears that envelope distortion occurs on the modulated signal, and bandlimiting distortion occurs on the demodulated signal. It is then shown experimentally that consistent quality is possible when the short-time bandwidth is smaller than a fraction of the channel bandwidth. Based on this more complete model, improved algorithms to adjust the modulator are proposed.

Chapter 5 examines the problem of efficiently allocating the resource of the channel between different components of the visual information. In the context of subband decomposition, a sequential transmission of the subbands generates various amount of noise, and of various psychovisual importance, depending on the spatial frequencies of the component. It is shown that by adjusting the rate of transmission of each component, a better subjective balance is possible, yielding substantial improvement.

Finally, chapter 6 describes two-real world applications that demonstrate the potential of adaptive frequency modulation and multirate transmission.

Chapter 2

Background

This chapter introduces the major aspects of frequency modulation. It is shown here that frequency modulation can exchange signal-to-noise ratio with bandwidth, and that this exchange depends on the statistics of the signal source. Elements of visual perception are introduced to define a more meaningful quality criterion.

2.1 Definitions

Frequency modulation can be most easily defined with the help of the complex phasor. A modulated signal $z(t)$ is the real part of its complex phasor $\tilde{z}(t)$:

$$z(t) = Re(\tilde{z}(t)) \tag{2.1}$$

Angle-modulated phasors have constant magnitude while their instantaneous phase $\phi(t)$ varies with time.

$$\tilde{z}(t) = \sqrt{2} A_c e^{j\phi(t)} \tag{2.2}$$

Frequency-modulated signals have the following instantaneous phase $\phi(t)$:

$$\phi(t) = 2\pi F_c t + 2\pi \nu \int_{-\infty}^{t} y(\tau) d\tau \tag{2.3}$$

where F_c is the carrier frequency, ν is the gain of the modulator [Hz/V], and $y(t)$ is the message to be transmitted. The instantaneous frequency $F_i(t)$, which is the

21

derivative of the instantaneous phase, is in this case:

$$F_i(t) = \frac{1}{2\pi}\frac{d\phi(t)}{dt} = F_c + \nu y(t) \tag{2.4}$$

We see here that the instantaneous frequency varies in direct proportion with the signal.

The frequency deviation D is defined by:

$$D = \nu y_0 \tag{2.5}$$

The value y_0 is a known quantity and thus specifying D is the same as specifying the modulator gain ν. In the future, the modulator will be characterized by D instead of ν, and the value y_0 is assumed to be known. The reference value y_0 is usually taken as the maximum value of $y(t)$. In that case, D is the effective peak frequency deviation.

When $y(t)$ is modeled by a Gaussian random variable of mean m_y and variance $\sigma_y{}^2$, y_0 is defined such that:

$$P(y(t) < y_0) = 0.99 \tag{2.6}$$

or, equivalently:

$$y_0 = m_y + 2.5\sigma_y \tag{2.7}$$

This definition, although a little arbitrary, introduces the useful notion of peak deviation for a random message. In the future, we shall assume the signal $y(t)$ to be zero-mean. If not, we can always assume that the mean value is transmitted beforehand and removed from $y(t)$.

Finally, the modulation index β is defined as:

$$\beta = \frac{D}{F_s} \tag{2.8}$$

where F_s is the maximum frequency of $y(t)$ to be transmitted.

Signal Representations

In some instances we use a discrete representation of the video signal $y(t)$, assuming a critical sampling at $F_N = 2F_s$. To stress the fact that a signal is discrete-time, the

square brackets are used. Hence the signal $y(t)$ sampled at $F_N = 2F_s$ is

$$y[n] = y(t)|_{t=nT_N} \qquad (2.9)$$

where $T_N = 1/F_N$.

In the following, F refers to the frequency variable of the Fourier transform, whereas f refers to the frequency variable of the discrete-time Fourier transform.

For raster-scanned signals with line duration H, the relation of the temporal signal $y(t)$ to its two-dimensional discrete representation $y(i,j)$ is given by:

$$y(i,j) = y(iT_N + jH) \qquad (2.10)$$

On other occasions, a three-dimensional discrete representation is used. In that case:

$$y(i,j,k) = y(iT_N + jH + kOH) \qquad (2.11)$$

where O is the number of lines per frame. Progressive scanning is assumed unless specified.

Context will tell what notation is used. The indexes i, j, k will always be used for horizontal, vertical, and temporal directions, respectively.

2.2 Bandwidth of FM Signals

The power spectrum $S_z(F)$ of a stationary process $z(t)$ is the Fourier transform of its autocorrelation function $R_z(\tau)$[Pro83]:

$$R_z(\tau) = E(z(t)z(t - \tau)) \qquad (2.12)$$

$$S_z(F) = \mathcal{F}(R_z(\tau)) \qquad (2.13)$$

There are many criteria used to define the bandwidth of a signal. Bandwidth can be defined on a percentage-of-energy basis using information on its power spectral density. Other possible definitions are the bandwidth necessary to pass the information with negligible distortion, or the frequency separation between carriers, so that negligible interference occurs.

23

The former definition is used in this chapter where it is defined as the band containing 95% of the power of $z(t)$. Clearly this bandwidth depends on the signal $y(t)$ to be transmitted –mostly its power spectrum density and its amplitude distribution– and on the modulator gain, or equivalently the frequency deviation D. A fixed deviation D will yield different bandwidths for different pictures unless they are very similar. Thus the relation between bandwidth and modulation index characterizing an information source will be similar for pictures having similar statistics.

The assumption here is that a picture can be modeled as a stationary process. This model allows us to characterize a source by its average behavior.

2.2.1 Modeling of the Luminance Signal

Television luminance signals can be modeled by stationary processes. A model often used in the literature is the first order Gauss-Markov model. In that case, the spectrum is:

$$S_y(F) = \frac{\sigma_y^2}{\pi F_0} \frac{1}{1 + F/F_0} \qquad (2.14)$$

where F_0 is the process bandwidth and σ_y^2 is the luminance power. In practice, however, the spectrum is limited to a frequency F_s, beyond which the power density is negligible.

The autocorrelation function of the luminance signal is, by inverse Fourier transform:

$$R_y(\tau) = \sigma_y^2 \exp(-F_0\tau) \qquad (2.15)$$

It can be shown that for such a signal $z(t)$ is also a stationary process [Pap83]. For large frequency (i.e. at the tail of the frequency distribution), $S_z(F)$ is approximated by a Gaussian spectrum [Tre68].

$$S_z(F) = \frac{1}{2} \frac{A_c^2/2}{\sqrt{2\pi}\sigma_y\nu} \left[\exp(\frac{-(F - F_c)^2}{2\sigma_y^2\nu^2}) + \exp(\frac{-(F + F_c)^2}{2\sigma_y^2\nu^2}) \right] \qquad (2.16)$$

In order to contain 95 % of the power, the bandwidth is:

$$B_z = 4\nu\sigma_y = 4\frac{\sigma_y}{y_0}D \qquad (2.17)$$

24

or, equivalently:

$$\frac{B_z}{F_s} = 4\frac{\sigma_y}{y_0}\beta \qquad (2.18)$$

Hence, the ratio of modulated bandwidth to basebandwidth is linearly related with the modulation index. It can be noted that only second-order statistics is needed for this evaluation.

Figure 2.2 shows the measured bandwidth required to transmit 95% of the energy as a function of the modulation index β for the picture "Girl", together with the value predicted by the stationary model. Here $y_0/\sigma_y = 2.5$. The curve predicted by the model fits the measurements very well.

2.2.2 Sine Wave Test Signal

For the particular, though unlikely, case where $y(t)$ is a sine wave at frequency F_s , a closed form of $S_z(F)$ exists [Pan65] [AM83].

$$S_z(F) = \frac{A_c^2}{2}\sum_{k=-\infty}^{\infty} J_k^2(\beta)(\delta(F - (F_c + kF_s)) + \delta(F + (F_c + kF_s))) \qquad (2.19)$$

where $J_k(x)$ is the kth Bessel function of the first order. In that case, the bandwidth necessary to pass 95% of the power is:

$$B_z = 2NF_s \qquad (2.20)$$

where N is the smallest positive integer such that:

$$\sum_{k=-N}^{N} J_k^2(\beta) \geq 0.95 \qquad (2.21)$$

The bandwidth necessary to pass 95% of the power was also approximated by Carson [JT37]:

$$B_z = 2(F_s + D) \qquad (2.22)$$

or, equivalently:

$$\frac{B_z}{F_s} = 2 + \beta \qquad (2.23)$$

This value is often used as a simple design rule in FM system. Carson's approximation is displayed on figure 2.3 , together with the exact values. Here $y_0/\sigma_y = \sqrt{2}$.
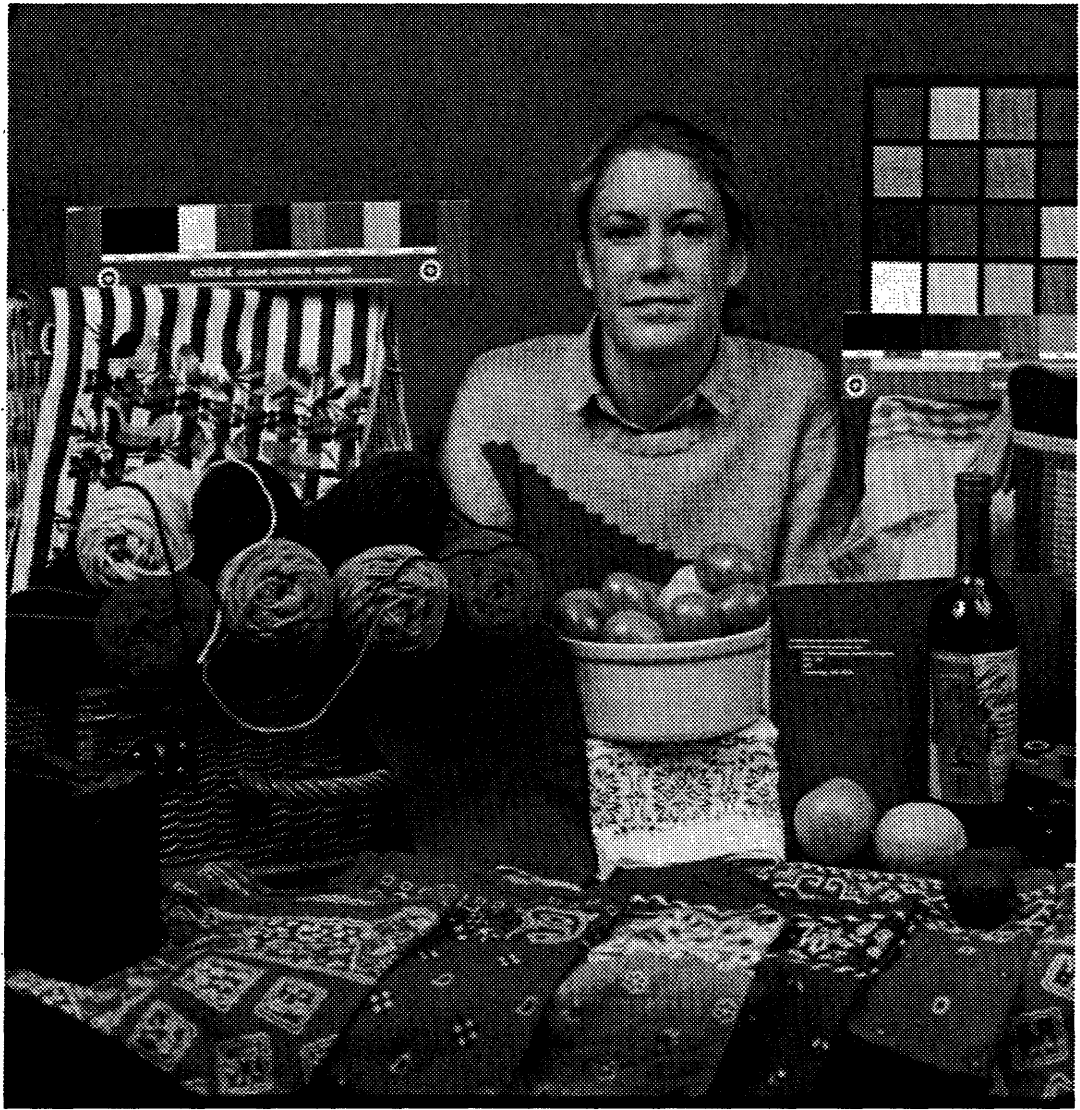
Figure 2.1: Picture "Girl". The size is 512 x 512.

O: measurements
-: Gauss-Markov model
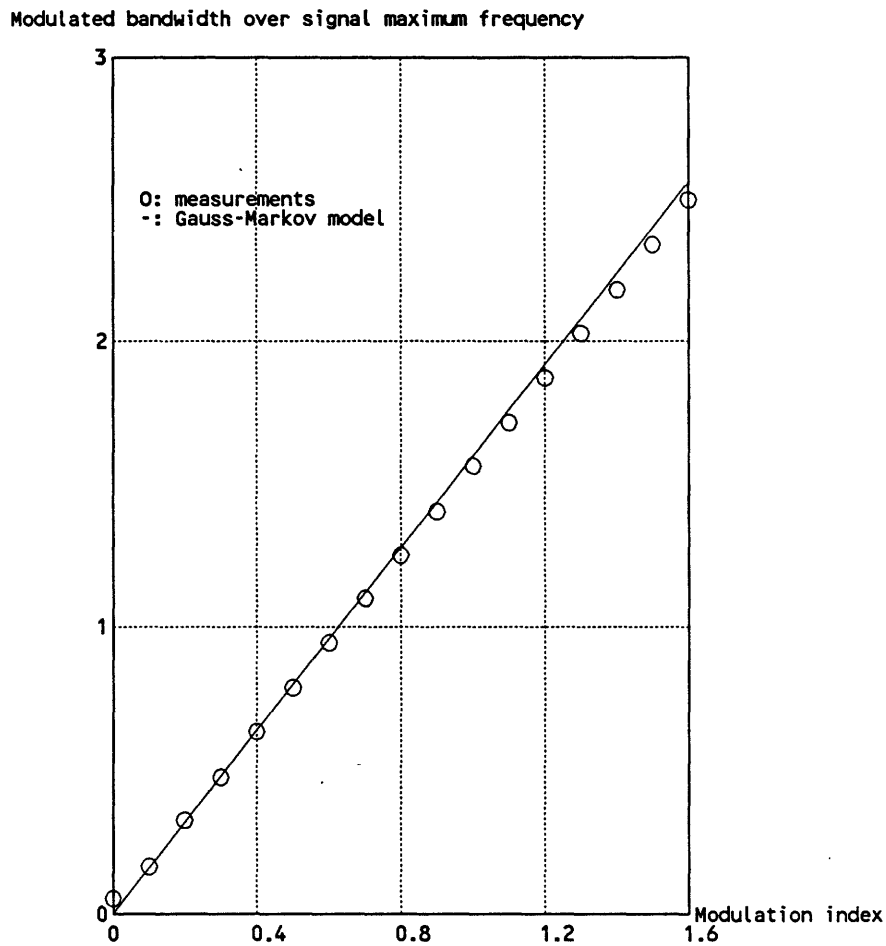
Figure 2.2: Bandwidth of FM signal (95%). Picture "Girl".

2.3 Noise Analysis

In a first approximation, the effect of the transmission of the modulated signal $z(t)$ is to introduce an additive noise component $n_{tot}(t)$ which can be assumed white and Gaussian over the bandwidth allocated to $z(t)$. The bandwidth allocated to $z(t)$ will be referred to in the following as the I.F. bandwidth (intermediate frequency). I.F. filtering is performed to remove components of the received signal outside the I.F. bandwidth as shown on figure 2.4. It is assumed in this chapter that the I.F. filtering doesn't affect $z(t)$, a valid assumption if the signal bandwidth is smaller than the I.F. bandwidth. The effects of the filtering on $z(t)$ are investigated in chapter 4.
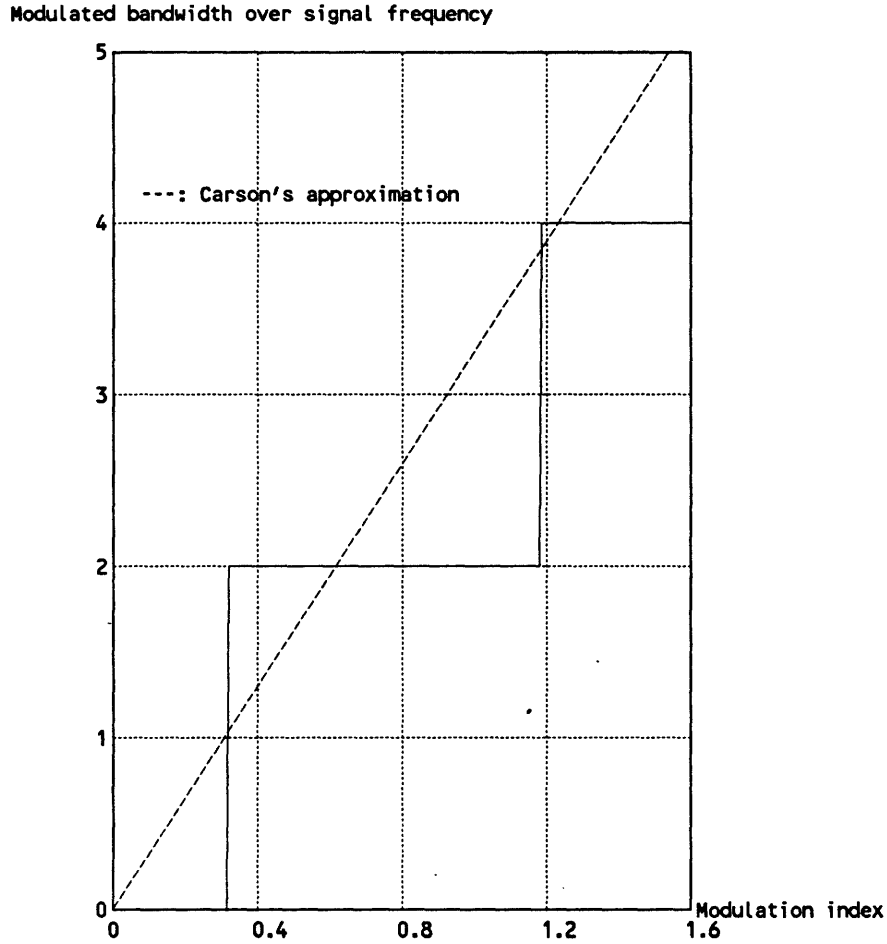
Figure 2.3: Bandwidth of FM signal (95%). Sine wave message.

The power spectrum density of the additive noise is:

$$S_{n_{tot}}(F) = \frac{N_0}{2} \qquad (2.24)$$

We can further assume that the equivalent noise $n(t)$ entering the demodulator after I.F. filtering is zero outside the I.F. bandwidth. By considering the I.F. filter an ideal band-pass filter for the purpose of noise evaluation, very little inaccuracy occurs. The filtered noise can then be modeled as a band-pass process, and, as such, can be expressed by its quadrature and in-phase components [dC84]:

$$n(t) = n_{tot}(t) * h_{IF}(t) \qquad (2.25)$$

$$n(t) = n_{\parallel}(t) \cos(2\pi F_c t) + n_{\perp}(t) \sin(2\pi F_c t) \qquad (2.26)$$

28
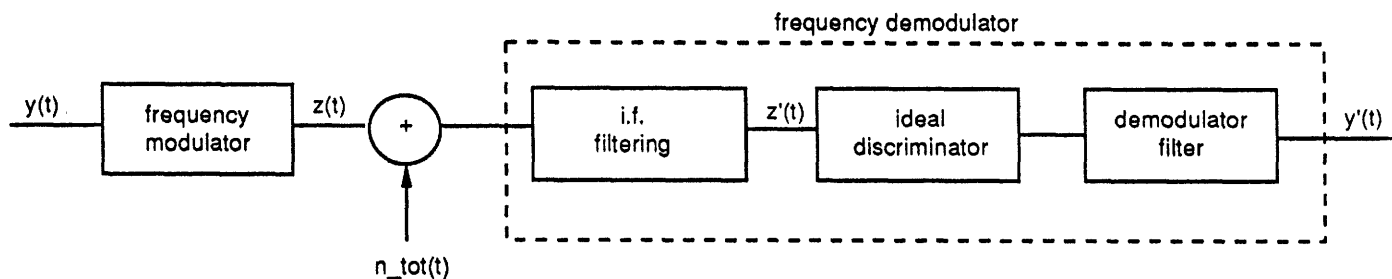
Figure 2.4: Simple model of an FM transmission system

The power spectrum density of $n_\perp(t)$ and $n_{\parallel}(t)$ is:

$$S_{n_{\parallel}}(F) = S_{n_\perp}(F) = \begin{cases} N_0 & ; \ |F| < B_{I.F.} \\ 0 & ; \ \text{otherwise} \end{cases} \qquad (2.27)$$

It is assumed that the magnitude of $n(t)$ is small compared to the magnitude of $z(t)$. This weak-noise analysis avoids the case where the phasor jumps 2π radians when the noise plus signal trajectory encircles the origin [JI65]. This effect –also called the threshold effect– manifests itself by large spikes on the demodulated signal.

2.3.1 Ideal Discrimination

The frequency demodulator shown in figure 2.4 is modeled as an ideal discriminator. The input to the demodulator is $\tilde{z}'(t)$,the noise corrupted and filtered version of $\tilde{z}(t)$. The ideal discriminator performs a Maximum-Likelihood estimation of the phasor angle $\phi(t)$.

$$\hat{\phi}_{\text{ML}} = \arctan\left(\frac{Re(\tilde{z}'(t))}{Im(\tilde{z}'(t))}\right) + i\,2\pi \qquad (2.28)$$

The estimate $y'(t)$ of the transmitted signal $y(t)$ is :

$$y'(t) = \left(\frac{1}{2\pi\nu}\frac{d}{dt}\hat{\phi}_{\text{ML}}(t) \ - \frac{F_c}{\nu}\right) * h_{\text{demod}}(t) \qquad (2.29)$$

where $h_{\text{demod}}(t)$ is an ideal low-pass filter of bandwidth F_s which removes the noise components outside the signal bandwidth. For a sinewave message, F_s is assumed

29

to be the maximum frequency of such a test signal. For the more realistic case of a Gaussian message, F_s is the bandwidth of the video signal required to ensure proper image quality. The message $y(t)$ is also bandlimited to F_s prior to transmission. The value of i in the right-most term in equation (2.28) is chosen to guarantee that no impulse occurs when the phasor crosses π and that the mean value of $y'(t)$ matches that of $y(t)$.

The error $e_n(t)$ due to transmission noise is the difference between the original signal and its estimate.

$$e_n(t) = y(t) - y'(t) \tag{2.30}$$

Using the weak-noise assumption, we can approximate the phase error and then the estimation error.

$$e_n(t) = \left(\frac{1}{2\pi\nu} \frac{d}{dt} \frac{(n_\perp(t))}{A_c} \right) * h_{\text{demod}}(t) \tag{2.31}$$

As $n_\perp(t)$ is a stationary Gaussian process, the estimation error is also stationary and Gaussian. The power spectrum density of $e_n(t)$ under the weak-noise approximation is:

$$S_{e_n}(F) = \begin{cases} \dfrac{1}{(2\pi\nu)^2} \dfrac{(2\pi F)^2 S_{n_\perp}(F)}{A_c^2} & ; \; |F| < F_s \\ 0 & ; \; \text{otherwise} \end{cases} \tag{2.32}$$

We note that the error term due to the noise has a spectrum proportional to the square of the frequency. This noise is usually termed "triangular" noise as its shape is triangular when displayed on a log-log scale, such as that of a spectrum analyzer.

Ideal filtering was preferred in our transmission model over Wiener filtering for $h_{\text{demod}}(t)$. A Wiener filter minimizes the expected error square [Pap83]. It was preferred to leave the signal unchanged (no attenuation) in the baseband so as to avoid any horizontal blurring of the picture.

2.3.2 Signal-to-Noise Ratio

In the previous section, we found a closed form for the estimation error. Here, an expression for the signal-to-noise ratio is derived. To display the results in a meaningful form, some definitions are necessary. The carrier-to-noise ratio is usually defined in

30

the literature as the ratio of the carrier power to the RMS noise power contained in the intermediate frequency band.

$$\text{CNR} = \frac{A_c^2/2}{B_{\text{IF}} N_0} \tag{2.33}$$

The normalized carrier-to-noise ratio CNR_0 is defined for a noise on twice the base bandwidth only:

$$\text{CNR}_0 = \frac{A_c^2/2}{2 F_s N_0} \tag{2.34}$$

The expected value of the error square can be obtained by integrating the error spectrum over the signal baseband.

$$E(e_n(t)^2) = \int_{-F_s}^{F_s} S_{e_n}(F) dF \tag{2.35}$$

where, again, F_s is the maximum frequency of $y(t)$ to be transmitted.

The expected value of the error squared can then be expressed after evaluating (2.35):

$$E(e_n(t)^2) = \frac{F_s^2}{6 \, \text{CNR}_0 \, \nu^2} \tag{2.36}$$

Finally, the signal-to-noise ratio can be expressed as:

$$\text{SNR} = \frac{E(y(t)^2)}{E(e_n(t)^2)} = 6(\frac{\sigma_y}{y_0})^2 \beta^2 \text{CNR}_0 \tag{2.37}$$

For the two sources, it is evaluated to:

$$\text{SNR} = \begin{cases} 3\beta^2 \text{CNR}_0 & \text{; sinewave message} \\ \cong 0.96 \beta^2 \text{CNR}_0 & \text{; Gaussian message} \end{cases} \tag{2.38}$$

We see here that for a given noise spectral density, the signal-to-noise ratio increases with the square of the modulation index β .

Figure 2.5 shows the theoretical prediction of the SNR together with the experimental results obtained by computer simulation for the picture "Girl" and for the sinewave signal. It can be noted that for very low carrier-to-noise ratio, the weak noise approximation no longer holds. For larger CNR , the model matches the measurements very well.

31
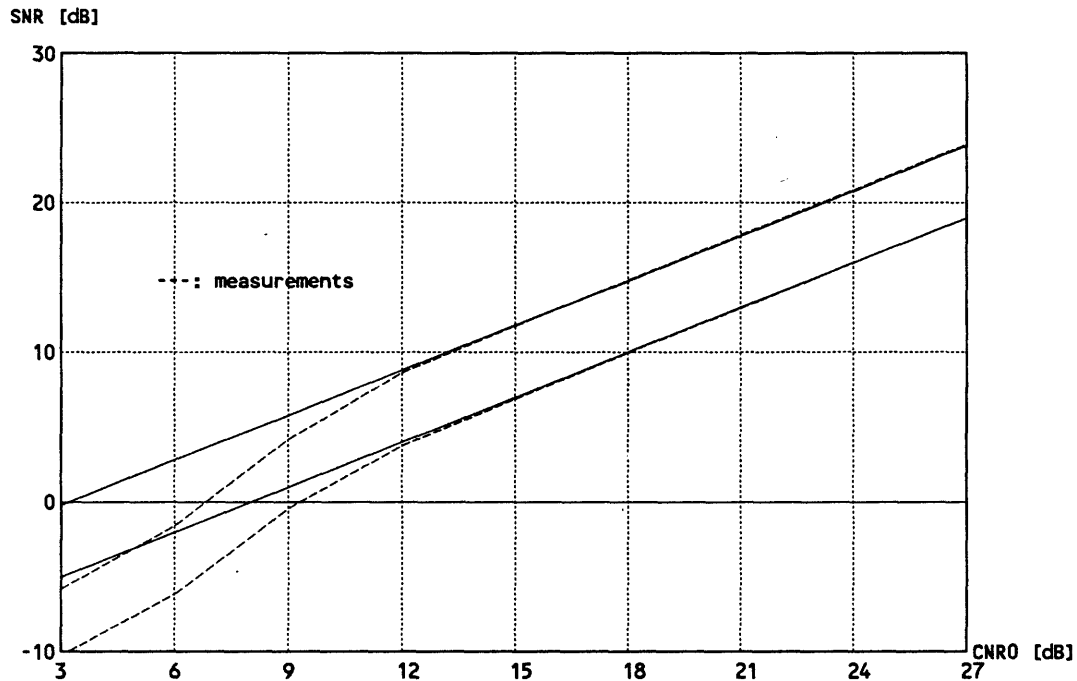
Figure 2.5: Signal-to-noise ratio versus normalized carrier-to-noise ratio. Weak noise approximation. Top: sinewave. Bottom: Gaussian message. The modulation index is 0.4 and the I.F. bandwidth is 4 Fs.

2.4 Noise-Bandwidth Exchange

We have seen in the previous sections that the cost of increasing the modulation index –or equivalently the frequency deviation– is an increase in bandwidth. On the other hand, this increase will yield significantly better noise figure.

Here a measure of how the noise figure relates to the bandwidth is obtained.

2.4.1 SNR Improvement

The normalized carrier-to-noise ratio CNR_0 can be better understood as the SNR of a reference amplitude modulation suppressed carrier system. It can be thought of as SNR_{ref} given the noise conditions of the channel.

32

The SNR of an FM system was derived in equation (2.37). The improvement obtained with frequency modulation is:

$$\frac{\text{SNR}_{\text{FM}}}{\text{CNR}_0} = \begin{cases} 3\beta^2 & \text{; sinewave message} \\ \cong 0.96\beta^2 & \text{; Gaussian message} \end{cases} \tag{2.39}$$

SNR improvement
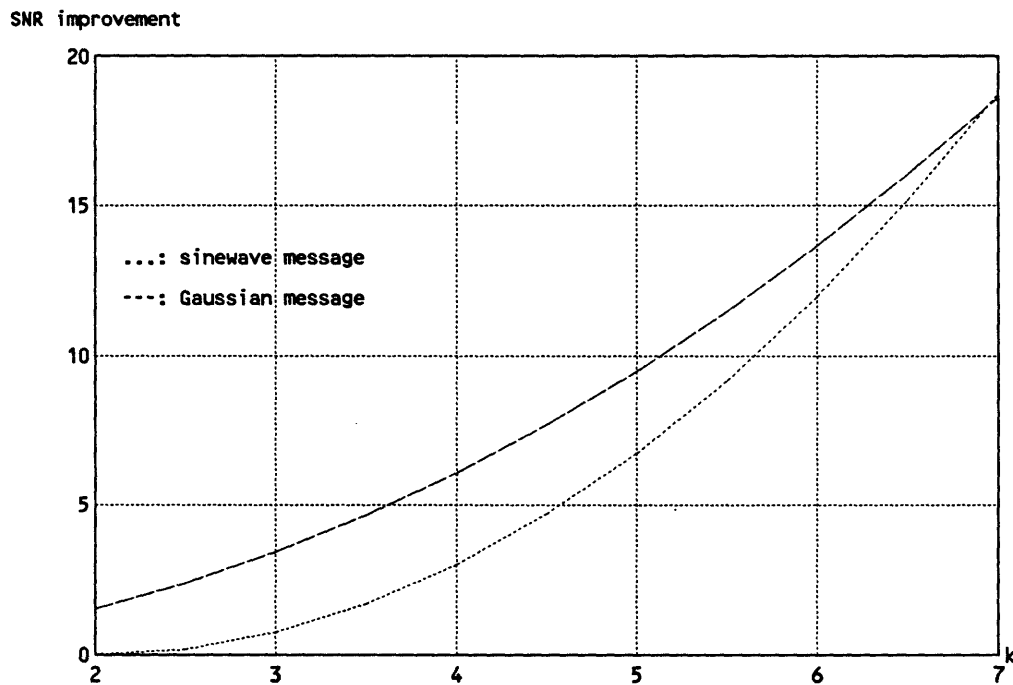
Figure 2.6: Bandwidth-noise exchange function for two signals sources

Let's define the bandwidth expansion factor k as the ratio of modulated bandwidth to basebandwidth.

$$k = \frac{B_z}{F_s} \tag{2.40}$$

where, again, B_z is the 95% bandwidth, and F_s is the bandwidth of $h_{\text{demod}}(t)$. By using the approximations of equations (2.17) and (2.22), we can express the SNR improvement as a function of k.

$$\frac{\text{SNR}_{\text{FM}}}{\text{CNR}_0} = \begin{cases} 3(\frac{k}{2} - 1)^2 & \text{; sinewave message} \\ \frac{6k^2}{16} & \text{; Gaussian message} \end{cases} \tag{2.41}$$

33

Figure 2.6 shows the SNR improvement versus the bandwidth expansion factor for a Gaussian message and the sine wave signal.

2.5 Perception and Noise Evaluation

The design of a coding scheme for transmission or storage usually involves minimizing an associated cost function. For example, a channel coder for digital information seeks to minimize the probability of error on a single digit [Pap83]. Equivalently, a K-L transform coder minimizes the mean error squared by transmitting only the coefficients of the Karhunen-Loeve transform with largest variance [Pra78].

The mean square error is not an adequate criterion for coding images, as it doesn't reflect well the perceived image quality. Defining a valid subjective measure of quality is a very complex task [Gir88]. However, one can improve the simple-minded mean square error criterion by incorporating elements of perception in the definition. Two elements will be used here, noise weighting and noise masking.

It is not intended to use sophisticated models of vision. Rather, we seek to introduce tools to better describe the performances of a system. It is hoped that these tools will fit in a more general communication engineering frame while still describing the requirements of visual information processing. Specifically, we shall define and evaluate the weighting factor and the masking factor.

2.5.1 Noise Weighting

In this section, the weighting function of the human visual system is evaluated for normal viewing conditions. The autocorrelation of the noise is then computed, from which the weighted noise mean square can be derived.

The response of the human visual system to different stimuli at different spatial frequencies is not uniform. Similarly, the response to a secondary stimulus (error signal) associated with a main stimulus (image) very much depends on the frequency content of the secondary stimulus. For example, high-frequency interference is less annoying than low-frequency interference.

In order to simulate the weighting effect of a viewer, the CCIR Rep. 637-2 defines the following low pass function to measure the weighted noise of 525-line television signals [Rho85]:

$$W_c(F) = \frac{1 + j2\pi FT_1}{(1 + j2\pi FT_2)(1 + j2\pi FT_3)} \qquad (2.42)$$

where $T_1 = 2.56\mu s.$, $T_2 = 3.7\mu s.$, and $T_3 = 730ns$. It is defined for normal television viewing conditions.

An equivalent weighting network for discrete pictures is defined below.

From the previous continuous-time weighting network, we aim to deduce a discrete-time equivalent network. The NTSC luminance bandwidth, for which this recommendation was intended, is 4.2 MHz. A discrete representation of a picture is obtained by sampling at the critical sampling frequency $F_N = 8.4MHz$. This maps the one-dimensional raster-scanned signal into a three-dimensional discrete signal. It should be noted that very often sampling is performed at a higher frequency. For example, in the CCIR recommendation 601, it is performed at four times the color subcarrier (14.24 MHz). Pictures obtained in this way are horizontally oversampled. Such is not the case of the picture used in the MIT computer facility. In the following, critical sampling is assumed.

From the frequency response prototype in continuous time, it is desired to design a discrete-time filter with approximately same response. Different methods have been tried, and the most satisfactory is the frequency sampling method. It amounts to specifiy the discrete-time Fourier transform of our desired filter to equal the corresponding response of the prototype filter at a number of discrete frequencies.

$$W(f)|_{f=FT_N} = W_c(F); \quad F = 0, \pm F_N/N, \pm 2F_N/N, ..., F_N/2 \qquad (2.43)$$

where N is a large even number and $W(f)$ is the discrete-time Fourier transform of $w[n]$.

$$w[n] = \text{IDTFT}_N(W(f)) \qquad (2.44)$$

The weighting function can be used for horizontal or vertical weighting of a sampled picture with square pixels.

35

In the following, we shall use the separable approximation which allows us to express a function of space as a product of single-variable functions. For example, the two-dimensionnal weighting function is the product of horizontal and vertical weighting:

$$w(i,j) = w[i]w[j] \tag{2.45}$$

In a similar manner, the two-dimensional correlation of the error signal can also be expressed in a separable form.

$$R_{e_n}(i,j) = R_i[i]R_j[j] \tag{2.46}$$

First the continuous-time autocorrelation of the noise is evaluated, and then that of its discrete representation is derived.

Using equation (2.32), we find the time autocorrelation of the noise:

$$R_{e_n}(\tau) = \frac{F_s^2}{(2\pi^2\nu)^2\mathrm{CNR}_0} \, g(2F_s\tau) \tag{2.47}$$

where $g(x)$ is the second derivative of $-\mathrm{sinc}(x) = -\sin(\pi x)/(\pi x)$, as shown on figure 2.7.

$$g(x) = \sin(\pi x)(\frac{\pi}{x} - \frac{2}{\pi x^3}) + \cos(\pi x)(\frac{2}{x^2}) \tag{2.48}$$

The expected value of the error squared can be obtained by evaluating $R_{e_n}(0)$. The limit of $g(x)$ as x goes to 0 is yields a result consistent with equation (2.35).

We see that $g(x)$ is negative for $x = \pm 1$. It demonstrates that $e_n(\tau)$ is very likely to change sign from one sample to the next when critical sampling is performed, illustrating the high-frequency structure of the noise. This contrasts with the $\mathrm{sinc}(x)$ function, the normalized autocorrelation of an ideally low-pass white noise, which is zero for $x = \pm 1$.

When a two-dimensional discrete representation of the signal is used, the horizontal autocorrelation of the noise is:

$$R_i[i] = \frac{F_s^2}{(2\pi^2\nu)^2\mathrm{CNR}_0} \, g(t)|_{t=iT_N} \tag{2.49}$$

36
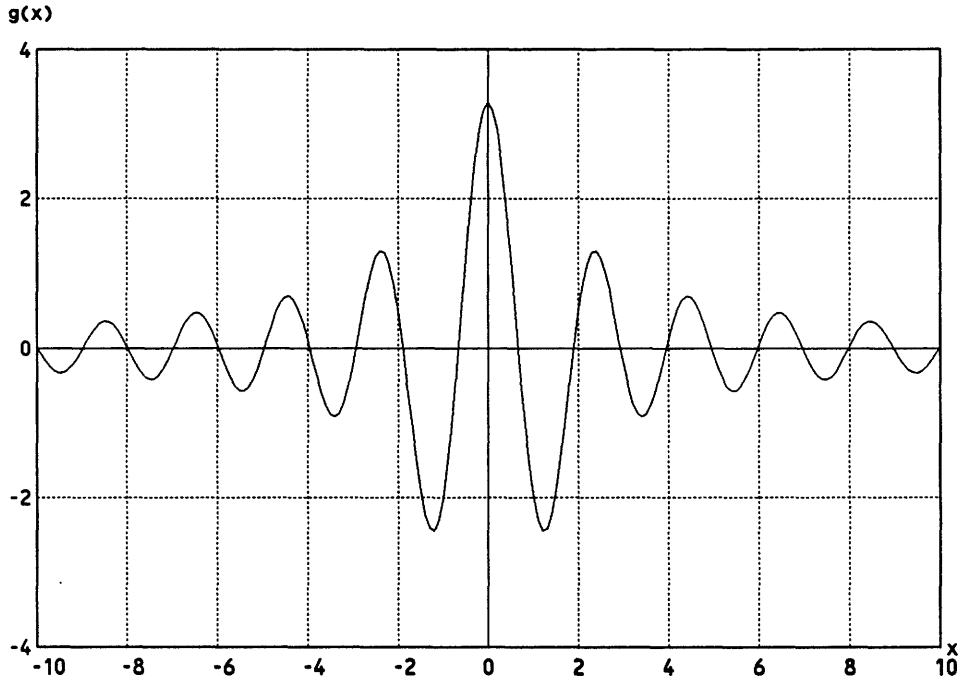
Figure 2.7: The function g(x).

By evaluating the normalized correlation $g(x)/g(0)$, we notice that it never exceeds 1% for $x > 100$. For a raster-scanned display, this means that the noise correlation from a pixel at one line to the next line is approximately zero, i.e. there is nearly no noise correlation from one line to the next. Hence, we find the vertical correlation of the error signal from line n to $n + j$:

$$R_j[j] = \frac{F_s^2}{6\mathrm{CNR}_0\nu^2} \, \delta[j] \tag{2.50}$$

where $\delta[n]$ is the unit sample sequence [OS75].

The mean-square error of the weighted noise $e_{nw}(t)$ can now be evaluated from the autocorrelation Fourier transform.

$$E(e_{nw}(i,j)^2) = \frac{N_0}{(2\pi\nu)^2 A_c^2} \int_{-1/2}^{1/2} (2\pi f_i)^2 \mid W(f_i) \mid^2 df_i \int_{-1/2}^{1/2} \mid W(f_j) \mid^2 df_j \tag{2.51}$$

where f_i, and f_j are frequency variables related to the indexes i, j respectively.

37

Finally, the weighting factor is defined. It is the ratio of the perceived noise power to the noise power. It reflects the subjective attenuation due to the human visual system, and depends on the type of interference. Its evaluation gives:

$$\zeta_{w_{tot}} = \frac{E(e_{nw}^2(i,j))}{E(e_n^2(i,j))} = \zeta_{wi}\,\zeta_{wj} = 0.014 \tag{2.52}$$

which is, expressed in decibels:

$$10\log(\zeta_{wi}) = -10.4\,\text{dB} \tag{2.53}$$

$$10\log(\zeta_{wj}) = -6.2\,\text{dB} \tag{2.54}$$

where ζ_{wi} and ζ_{wj} are the weighting factors in the horizontal and vertical directions, respectively. The perceived impairment is reduced significantly due to the high-frequency structure of the noise, mostly in its horizontal direction. This aspect of frequency modulation is one of the reasons for its popularity in television applications. However, it appears that frequency modulation doesn't perform as well in the vertical direction as it does in the horizontal one.

2.5.2   Noise Masking

Masking denotes the reduced sensitivity to a secondary stimulus (error signal) due to the presence of a large primary stimulus. This phenomenon is related to the saturation effect of the discharge pattern of sensitive cells in the retina. This can be noticed on noisy television signals: noise is mostly perceptible on large blank areas. Masking manifests itself mostly by hiding noise close to edges (luminance jumps) and in texture-like part of the picture. It has been used in image coding, both for DPCM [NP77], and PCM [SB81] coders. The reader is referred to [NP77] for a more complete overview of the bibliography.

A model is introduced here to quantify the masking effect. The main goal of this model is to provide us with an analytical tool to evaluate system performance.

The model assumes that the error signal at pixel (i, j) is attenuated by a masking function $M_a(i, j)$ [Tom86].

$$e_{nm}(i,j) = e_n(i,j)M_a(i,j) \tag{2.55}$$

38

where $e_{nm}(i,j)$ is a discrete representation of the masked error signal at pixel (i,j). The attenuation $M_a(i,j)$ is a function of the activity function $A(i,j)$ [NP77]:

$$A(i,j) = \sum_{n=i-l}^{i+l} \sum_{t=j-k}^{j+k} \alpha^{\|(n,t)-(i,j)\|} (1/2(|m_{nt}^H| + |m_{nt}^V|)) \qquad (2.56)$$

which is a local average of the slopes in the neighborhood of (i,j). α is taken here to be 0.35 and k=l=2.

$$m_{nt}^H = x(n,t) - x(n-1,t) \qquad (2.57)$$

$$m_{nt}^V = x(n,t) - x(n,t-1) \qquad (2.58)$$

The masking function is related to the activity function through the polynomial:

$$M_a(i,j) = \frac{1}{1 + KA(i,j)} \qquad (2.59)$$

where K=0.04. Figure 2.8 shows the masking function for the picture "Girl". Close to an edge, our model predicts a maximum attenuation of 20 dB of the perceived error signal.

The masked error is obviously no longer stationary. However, for a given picture, we can derive the sample mean of the masked error square:

$$\overline{(e_{nm}(i,j)^2)} = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} e_{nm}^2(i,j) \qquad (2.60)$$

The masking factor is the ratio of the masked noise to noise mean square values. It reflects the subjective attenuation of the noise power due to masking. It depends on the statistics of the picture. For a specified part of the picture, it can be evaluated:

$$\zeta_m = \frac{\overline{e_{mn}(i,j)^2}}{\overline{e_n(i,j)^2}} \qquad (2.61)$$

where the sample mean is computed on that part of the picture.

For a stationary error signal, the average masking factor can be estimated by:

$$\zeta_m = \frac{1}{MN} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} M_a(i,j)^2 \qquad (2.62)$$

where M, respectively N, is the vertical, horizontal, size of the picture, or the region of interest.

Figure 2.8: The masking function for the picture "Girl". The masking is unity in white regions.

2.5.3 Perceived SNR

Image quality, as perceived by the viewer, is not reflected by the signal-to-noise ratio calculated over the entire picture, but, rather, by evaluating it on small parts of the picture. Viewers usually object to the worst part of the picture [Tom86].

Hence, when the statistics of noise, or of other form of distortion, varies over the picture, a segmented evaluation of the signal-to-noise ratio is preferred. The same applies to the evaluation of the masking factor, which then describes the reduction of noise sensitivity in the small area of the picture for which it was computed. Texture-like parts of the picture will benefit a lot from this effect, while blank areas obviously will not benefit from it at all.

The perceived SNR for a segment of picture is:

$$\text{perceived SNR} = \frac{\text{SNR}}{\zeta_w \zeta_m} \tag{2.63}$$

where ζ_m is computed on that part of the picture.

To illustrate this concept, an example is developed.

It is desired to transmit "Girl" over a channel presenting a normalized carrier-to-noise ratio CNR_0 of 15 dB, using frequency modulation. From different bandwidth considerations, the modulation index β is set to 0.5.

We find that the SNR after demodulation is 9 dB. Taking weighting into account, the perceived SNR is 25 dB.

Next, we evaluate the factor due to masking. A small section of "Girl" is used for the purpose of demonstration, as illustrated in figure 2.9, where the masking function is shown next to the transmitted picture. The expected masking factor is 0 dB on the panel behind the bottle, whereas on the texture behind this panel it is about -7 dB. On an edge, the factor is -14 dB (a conservative value).

Finally, we find for the different areas:

$$\text{perceived SNR} = \begin{cases} 25\,\text{dB} & ; \text{panel} \\ 32\,\text{dB} & ; \text{texture} \\ 39\,\text{dB} & ; \text{edge} \end{cases} \tag{2.64}$$

41

Figure 2.9: Illustration of masking and noise perception. Left: original, middle: after transmission, right: masking function.

We can draw two conclusions from this example. First, some elements of perception, although very rudimentary, can help us to better evaluate a system performance. Second, it appears that frequency modulation doesn't perform uniformly over the picture, leaving room for improvement.

Chapter 3

Adaptive Frequency Modulation

In the previous chapter, we have evaluated the noise improvement for a given bandwidth increase. The information to be transmitted, a sequence of still pictures, was assumed to be stationary in the wide sense: the first- and second-order statistics do not vary over the picture [JI65]. It was also found that noise perception is not uniform across the picture.

In this chapter, a scheme is presented to increase performance at little cost: an increased complexity of the encoder and decoder and the transmission of additional side information.

3.1   Decomposition into Lows and Highs

The Gauss-Markov model of a raster scanned video signal is very approximate. The local mean of the luminance signal not only varies with scene content, but also it varies in different parts of the picture [HC76]. Similarly, a local measure of the variance ranges from zero on a uniform background to a large value in textures.

In order to match the Gaussian image model better, the local mean of the picture is removed. For the modulated signal, this amounts to reducing very-low-frequency variations of the carrier frequency. The local mean is obtained by low-pass filtering the picture.

The low-pass picture is downsampled by M_l x N_l in order to reduce significantly

this information, which will also be transmitted. The resulting component is called the Lows. An approximation of the low-pass picture is then obtained by interpolating the Lows up by the same factor. A component free of low-frequency variations, called the Highs, is computed by taking the difference of the original picture and its reconstructed low-pass version. Figure 3.1 shows how this decomposition is performed. The decomposition has proved to be fruitful for a digital coder [SB81]. Figure 3.2 shows the two components for "Girl", where $M_l = N_l = 4$.

To compute the local mean, low-pass filtering is performed with a separable two-dimensional Gaussian prefilter.

$$h_{av}(i,j) = h_a[i] \cdot h_a[j] \tag{3.1}$$

where

$$h_a[i] = \frac{1}{\sqrt{2\pi}\sigma_f} \exp(\frac{i^2}{2\sigma_f^2}) \tag{3.2}$$

The standard deviation σ_f of the filter impulse response is set to 1.6 when the down-sampling factor is 4 x 4 and to 3.2 when the downsampling factor is 8 x 8. Reconstruction after downsampling is performed with a postfilter of same form, but with σ_f equal to 2.4, and 4.8, respectively. This class of filters is common in image processing [ST85]. In this chapter a downsampling factor of 4 x 4 wil be used.
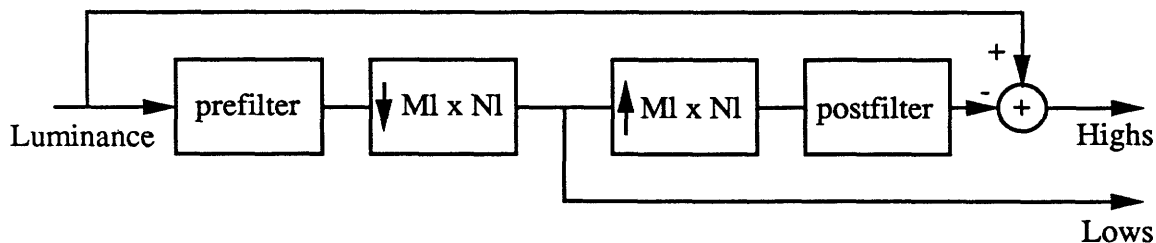
Figure 3.1: Decomposition of the picture into Lows and Highs components

(a)

(b)

Figure 3.2: Picture "Girl": (a) Lows component. (b) Highs component.

3.2 Frequency Modulation of a Discrete Signal

A discrete representation of the signal is essential to perform operations such as component separation, as it permits the use of VLSI technology. Although the components defined above are discrete both in space and amplitude, the effect of amplitude quantization is ignored. Mathematical operation in the encoder and decoder can be performed with a sufficient precision so as to yield no noticeable alteration of the image.

In this section, the necessary mapping from a discrete signal to a continuous signal is reviewed, and the results are specialized to the FM case. Let's suppose that the encoder generates the samples of component l at a rate r_l samples/sec. Let us assume that a channel is available with a bandwidth limited to B. We propose to find the performance of frequency modulation for such a discrete source.

Two additional blocks are needed to perform frequency modulation of a discrete signal. First, an interpolator maps the stream of samples $x_l[n]$ into a continuous signal $y(t)$ which is fed to the modulator. Second, a combination of a matched filter and a sampler maps the continuous signal $y'(t)$ obtained after transmission back into a digital stream $x_l'[n]$. This is shown in figure 3.3.
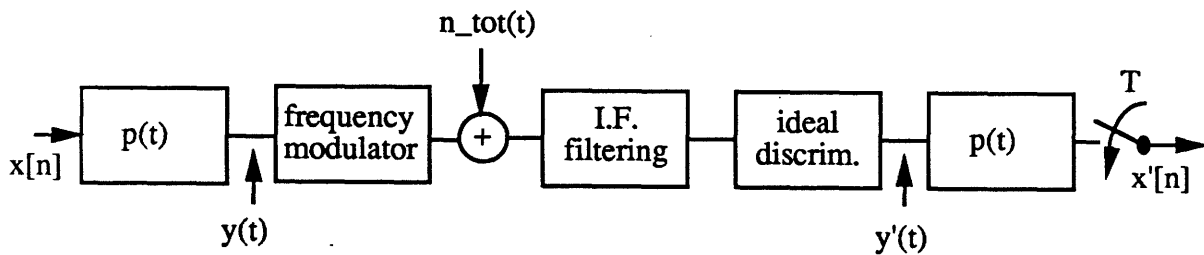
Figure 3.3: FM system for a sampled signal

Interpolation is performed with pulse $p_l(t)$:

$$y(t) = \sum_n x_l[n] p_l(t - nT_l - \theta) \tag{3.3}$$

46

where $T_l = 1/r_l$. The initial phase θ is a random variable uniformly distributed between 0 and π.

The pulse $p_l(t)$ used here has a raised-cosine spectrum [Pro83]. Other pulses are possible.

The receiving filter is matched to the transmitted pulse $p_l(t)$ and to the transmission noise. To simplify the discussion, we shall consider matching to a white noise. In this case, matched filtering amounts to specify the demodulator filter described in section 2.3.1 as having impulse response $p_l(t)$. A small gain would be possible by using a receiving filter matched to the triangular noise, or by using optimal joint pre- and postfilter [Mal86]. This is investigated in chapter 5.

$$h_{\text{demod}}(t) = p_l(t) \tag{3.4}$$

In our case, we have:

$$p_l(t) * p_l(t) = \frac{\cos 2\pi \lambda t}{1 - (4\lambda t)^2} \text{sinc}(r_l t) \tag{3.5}$$

The pulse $p_l(t)$ convolved with itself satisfies the Nyquist criterion, which states that the Fourier transform of $p_l(t) * p_l(t)$ must have a central symmetry around $F = r_l/2$ to guarantee no intersymbol interference. The parameter λ controls the amount of roll-off. For $\lambda = 0$, we are in the perfectly bandlimited case, as described in chapter 2. For larger λ, the bandwidth of $y(t)$ increases, but the trailing and leading oscillations of $p_l(t)$ decay more rapidly than for $\lambda = 0$. This is an important feature to guarantee robustness when timing jitter is present in the receiver [Pro83].

In order to predict B_x, the bandwidth of the frequency-modulated signal, the second-order statistics of $y(t)$ are needed. By modeling the discrete component l by a Gauss-Markov process with correlation coefficient ρ_l, we define:

$$R_{x_l}[n] = \sigma_l^2 \, \rho_l^{|n|} \tag{3.6}$$

Note that if the samples of $x_l[n]$ are transmitted in a random order, as is done in some scrambling schemes, the correlation coefficient is zero. The scrambled picture looks like white noise.

47
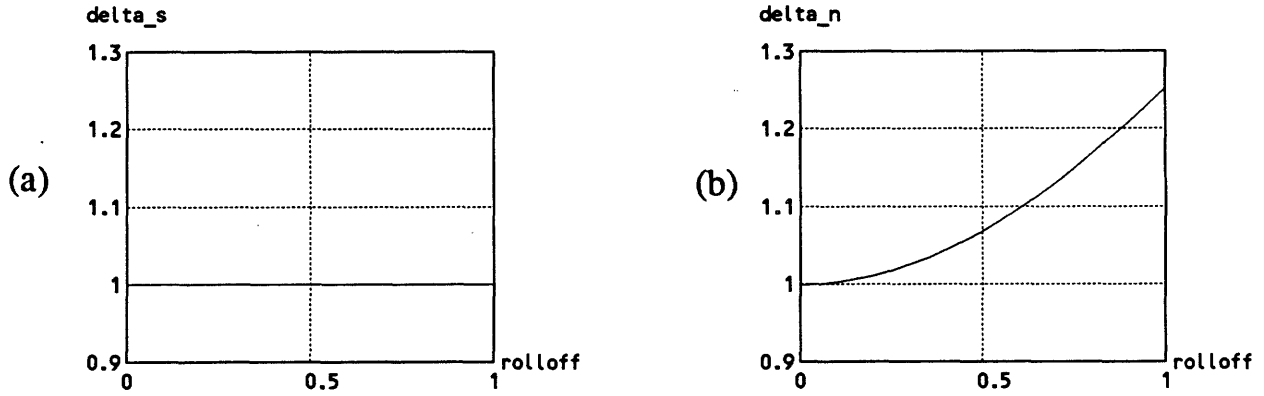
Figure 3.4: (a) Standard-deviation ratio of interpolated signal $y(t)$ to discrete signal $x_l[n]$. (b) Standard-deviation ratio of demodulated noise with roll-off λ to demodulated noise in the ideal bandlimited case.

The spectral density of $y(t)$ is:

$$S_y(F) = \mid P_l(F) \mid^2 \frac{\sigma_l^2(1 - \rho_l^2)}{1 + \rho_l^2 - 2\rho_l \cos(2\pi F T_l)} \tag{3.7}$$

We find the ratio Δ_s^2 of the variance of $y(t)$ to the variance of $x_l[n]$.

$$\Delta_s^2 = \frac{\sigma_y^2}{\sigma_l^2} = \frac{1}{\sigma_l^2} \int_{-\infty}^{+\infty} \mid P_l(F) \mid^2 \frac{\sigma_l^2(1 - \rho_l^2)}{1 + \rho_l^2 - 2\rho_l \cos(2\pi F T_l)} df \tag{3.8}$$

Figure 3.4 displays this ratio as a function of λ. Hence, the ratio k of modulated bandwidth to half sample-rate is:

$$k = \frac{B_z}{r_l/2} = 4 \frac{\Delta_s \sigma_l}{x_0} \beta \tag{3.9}$$

where x_0 is a known quantity used to define D:

$$D = \nu x_0 \tag{3.10}$$

The modulation index β is defined, consistently with (2.8), as:

$$\beta = \frac{D}{r_l/2} \tag{3.11}$$

Equation (3.8) should be compared with (2.18). It can be shown that Δ_s is always equal to unity. FM transmission of samples of a bandlimited process using interpolation $p_l(t)$ requires the same bandwidth as if it were directly transmitted in a continuous form with the same peak frequency deviation.

The last block of figure 3.3 consists of a matched filter and a sampler. The sampling process is assumed ideal. We shall investigate the influence of $p_l(t)$ on noise performance.

The noise spectrum is:

$$S_{e_n}(F) = \frac{1}{(2\pi\nu)^2} \frac{(2\pi F)^2 S_{n_\perp}(F)}{A_c^2} \mid P_l(F) \mid^2 \qquad (3.12)$$

We find the expected error squared on $x_l[n]$:

$$E(e_n^2[n]) = E((x_l'[n] - x_l[n])^2) = E((y'(t) - y(t))^2) = \int_{-r_l}^{+r_l} S_{e_n}(F)dF = \Delta_n^2 \frac{F_s^2}{6\,\mathrm{CNR_0}\,\nu^2}$$
$$(3.13)$$

The term Δ_n^2 denotes the ratio of noise variance with rolloff parameter λ to the noise variance in the ideal bandlimited case. Evaluation of equation (3.13) shows a modest increase of demodulated noise as the rolloff parameter λ increases, as illustrated in figure 3.4.

Finally, the signal-to-noise ratio for a sampled source is:

$$\mathrm{SNR} = \frac{E(x_l^2[n])}{E(e_n^2[n])} = \frac{6}{\Delta_n^2} \left(\frac{\sigma_l}{x_0}\right)^2 \beta^2 \,\mathrm{CNR_0} \qquad (3.14)$$

where

$$\mathrm{CNR_0} = \frac{A_c^2/2}{r_l N_0} \qquad (3.15)$$

The SNR improvement is:

$$\frac{\mathrm{SNR_{FM}}}{\mathrm{CNR_0}} = \frac{6k^2}{16\Delta_s^2\Delta_n^2} \qquad (3.16)$$

When comparing (3.16) with (2.41), we see that the performance of a FM system for a sampled source is slightly inferior to that of the original continuous source. The change of performance depends on the rolloff parameter of the interpolation function. The receiver filter degrades the noise performance a little. In the following , the ideal interpolation pulse $p_l(t)$ with $\lambda = 0$ is used, unless otherwise specified. In practical applications, when timing jitter is a problem, the rollof λ can be increased.

49

3.3 Adaptive Modulation of the Highs

In the section 3.1, we transformed the luminance signal into a zero-mean discrete signal, called the Highs. The Lows component is transmitted separately as shown in chapter 6. However, we cannot assume that the statistics of the Highs are stationary. A good approximation is to consider a picture as made of many small blocks inside which the statistics can be assumed constant. Every block is modeled by a multivariate Gaussian model whose parameters can be evaluated.

If the statistics of all blocks are known prior to transmission, it is possible to improve the performance of our modulation scheme by adjusting the modulation index from block to block. The index in each block is transmitted to the receiver. By doing so, we must however respect a constraint on the usage of the channel. A logical constraint is to restrict the bandwidth of the modulated signal to be smaller than a preassigned value B, the channel bandwidth. More precisely, the short-time bandwidth b_z, where the bandwidth is estimated on a small period of time, must be limited. By using short-time bandwidth estimate, we can ensure that critical cases are handled equally well by a fixed bandwidth limited channel.

As noise performance improves for larger frequency deviation, the problem of optimum adaptive modulation can then be stated as follows:

- In block (m, n), choose the frequency deviation $D(m, n)$ so as to:
  Maximize $D(m, n)$ with constraint that the short-time bandwidth b_z computed over the block region doesn't exceeds B more than 5 % of the time.

Varying the frequency deviation amounts to varying the modulator gain, as D is defined as the deviation for a given input y_0. An optimal solution to this problem is very hard to treat analytically. It also depends on the algorithm and the window used for the short-time estimate. An approximate solution is developed here.

A block diagram of the adaptive frequency modulator is shown in figure 3.5.
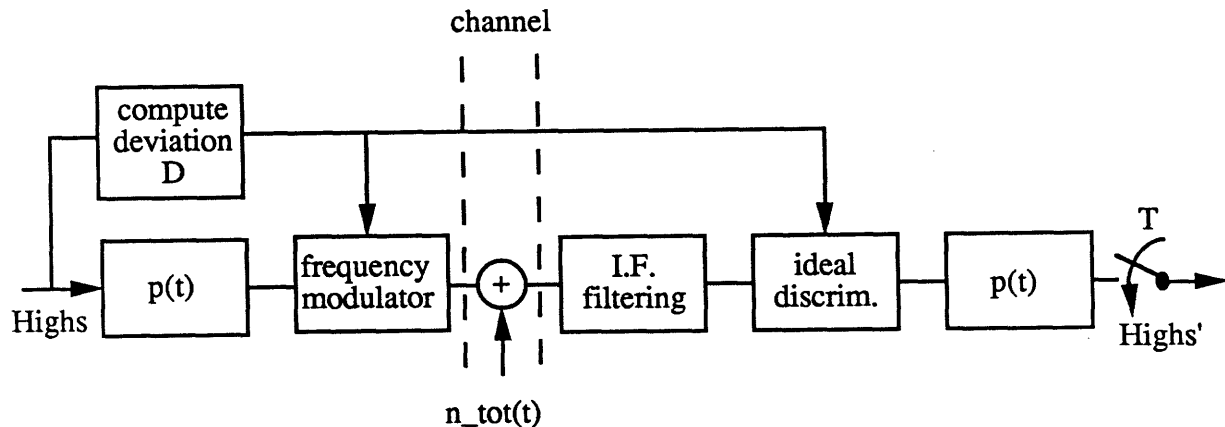
Figure 3.5: Block diagram of adaptive frequency system

3.3.1 Derivation of the Deviation Function

In block (m, n) of size M times N, the standard deviation of a pixel in the block can be estimated:

$$\hat{\sigma}(m, n) = (\frac{1}{MN} \sum_{i=mM}^{(mM)+M-1} \sum_{j=nN}^{(nN)+N-1} y(i, j)^2)^{\frac{1}{2}} \qquad (3.17)$$

In a window centered in the block (m, n) and for a given frequency deviation, a measure of the short-time bandwidth is obtained using equation (2.17), which states that the FM bandwidth is linearly related to the signal standard deviation for Gaussian processes.

Under the assumption that the Highs can be modeled by a Gaussian process, the deviation $D(m, n)$ in block (m, n) yielding a short-time bandwidth equal to B is then:

$$D(m, n) = B \frac{y_0}{4\hat{\sigma}(m, n)} \qquad (3.18)$$

where y_0 is a known constant, usually the peak value of $y(t)$. In order to avoid abrupt change of noise performance from block to block, the deviation is varied smoothly using bilinear interpolation between block values. This avoids an artifact known as blocking effect. Because of occasional excessive bandwidth values, the index in (3.18)

51

is decreased by about 15% in practical cases.

$$D(m,n) = \alpha B \frac{y_0}{4\hat{\sigma}(m,n)} \tag{3.19}$$

where $\alpha = 0.86$ for "Girl".

Alternatively, one can define a nominal deviation D_0 based on a specified standard deviation σ_{max}:

$$D_0 = \alpha B \frac{y_0}{4\sigma_{max}} \tag{3.20}$$

Then,

$$D(m,n) = f_a(m,n)D_0 \tag{3.21}$$

where $f_a(m,n)$ is the adaptation factor for the block (m,n).

$$f_a(m,n) = \frac{\sigma_{max}}{\hat{\sigma}(m,n)} \tag{3.22}$$

The adaptation factors comprises the extra information necessary to control the receiver.

In order to verify that the short-time bandwidth is effectively linearly related to a local measure of the standard deviation, an experiment is performed. It compares what the Gaussian model predicts and what we measure.

Bandwidth measurement relies on spectrum estimation techniques. Power spectrum estimates are affected by uncertainty and also bias. Specifically, a smoothing of the spectrum is obtained by windowing the estimated correlation function. The choice of the window will determine the trade-off between variance and bias of the power estimate. The popular Hanning window was selected for this application.

Another important effect of windowing is that a smaller window decreases the frequency resolution of the spectrum, a result well-known as the uncertainty principle [dC84].

Figure 3.6 shows a typical scanned line of the Highs component $y[i]$. It is quite apparent here that the signal statistics vary a lot over the picture width. The measured short-time bandwidth of the modulated signal $z(t)$ is displayed below. Every block n is aligned with the corresponding segment of $y[i]$.

52
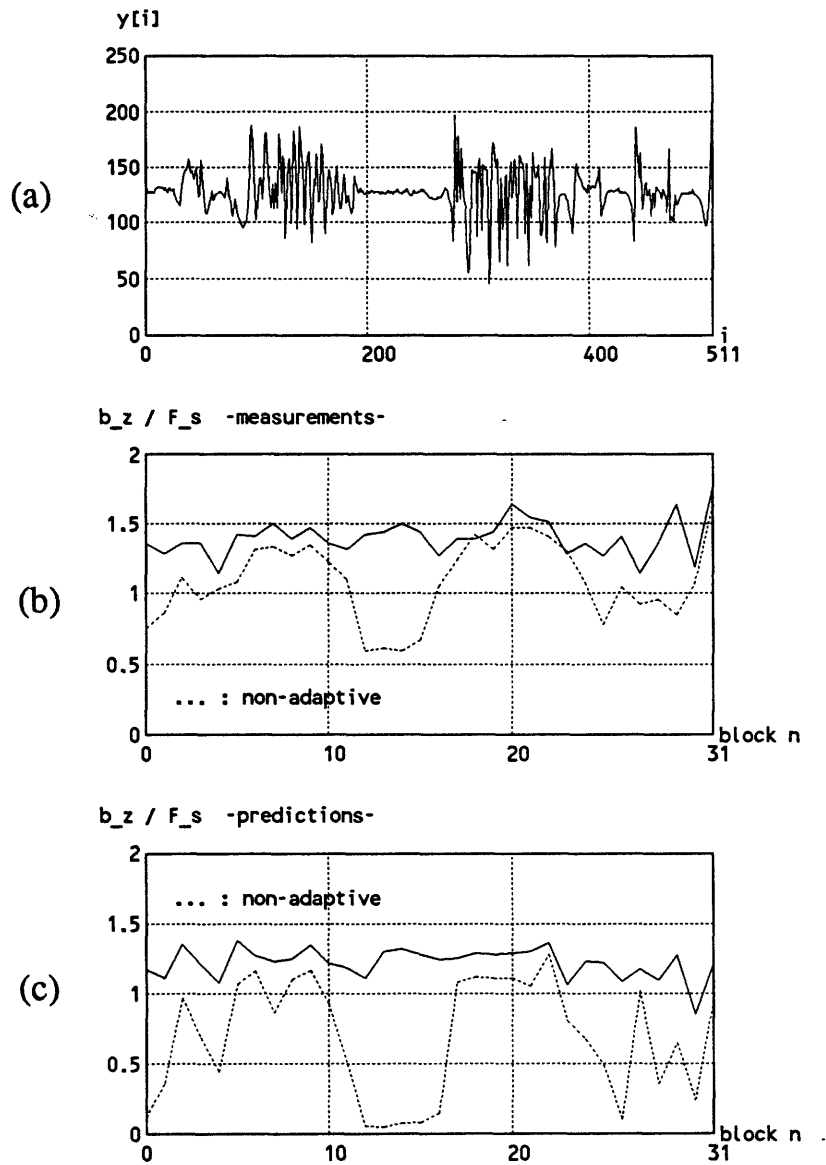
Figure 3.6: Evolution of the short-time bandwidth for adaptive and non-adaptive FM. (a) Scanned line of the Highs component. (b) Normalized short-time bandwidth of $z(t)$ -measurements-. Window is of type Hanning and of length $16/F_s$ (c) Normalized short-time bandwidth of $z(t)$ -prediction using the Gaussian model-.

In the non-adaptive system, the bandwidth fluctuates a lot, with large values in "busy" segments of the video signal. By contrast, and as expected, these fluctuations are very small for the adaptive system. The peak values of both systems are about the same. The block size used for computing the deviation is $N = M = 4$.

The measurements should be compared with the theoretical predictions using the Gaussian model. The former presents systematically larger values than the latter. It is attributed to the windowing of the correlation function which enlarge the spectrum because of the convolution with the window Fourier transform. Also, the low frequency resolution due to smoothing in (b) makes it impossible to measure small bandwidths accurately, hence the large differences between (b) and (c) for small bandwidth values.

Taking into account the large uncertainty of both measurements and predictions, we can be satisfied with the Gaussian model to predict short-time bandwidth. The solution proposed in section 3.3.1, although not optimal, is adequate for the purpose of adaptive modulation under a limited short-time bandwidth criterion. This is demonstrated by figure 3.7, where the histogram of the short-time bandwidth b_z is computed for "Girl" with $k = B/F_s = 1.5$. Ideally, this function should be a Dirac function at $b_z = 1.5F_s$. The variance of the estimates and the modeling errors are responsible for the bell shape. The probability of exceeding the nominal value B is very small. In reality, it is even smaller when the bias of the estimates due to smoothing is taken into account. The window size for the spectrum estimate is $16/F_s$.

3.4 Noise Analysis

In the previous section, a decomposition of the picture into blocks made possible an adaptive modulation scheme. In this section, the noise improvement due to adaptive modulation is evaluated. The transmission of the side information, namely the Lows and the adaptation factors, is not taken into account in this section. It is covered in chapter 5 and 6.

The result of a computer simulation of the system is shown on figure 3.9 for

p(b_z / F_z)
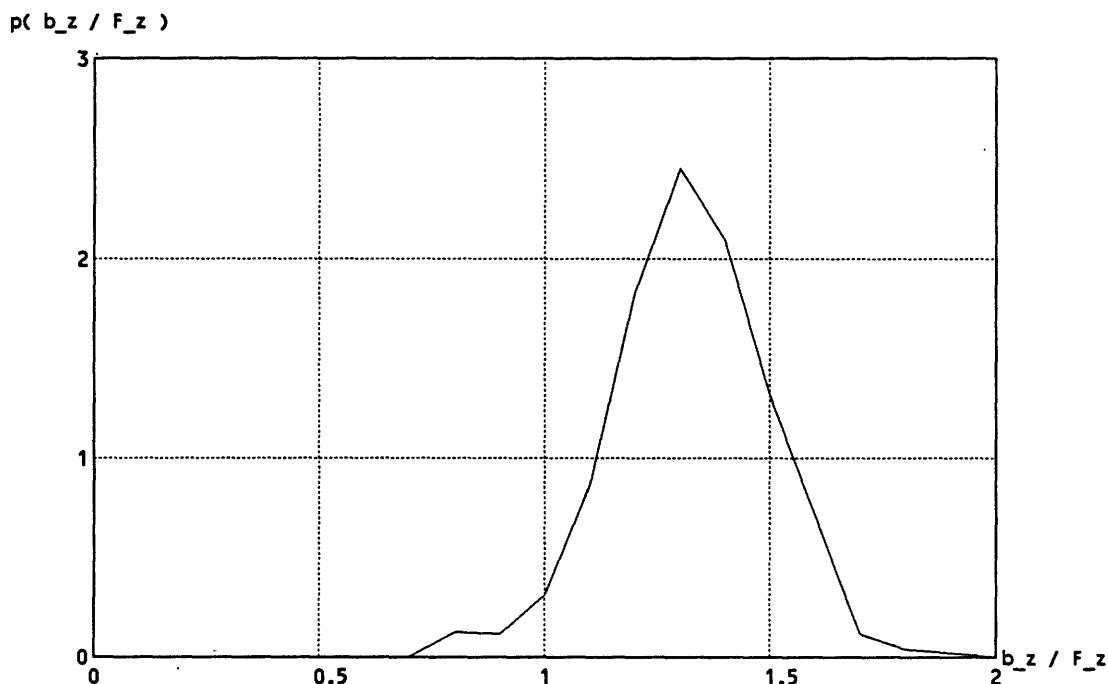
Figure 3.7: Histogram of the short-time bandwidth of the adaptively modulated signal for k = 1.5.

$\text{Max}(b_z) = B = 1.5F_s$. It should be compared with figure 3.10 where a regular frequency modulator of fixed deviation is used with the same peak bandwidth to signal maximum frequency ratio. A large decrease of noise visibility can be noted for the adaptive system. The block size for the adaptation is $N = M = 4$.

3.4.1 Block Process

A doubly stochastic model for image formation is proposed here for the Highs component. It is based on the assumption that signal statistics are constant in a block of small dimensions. Blocks of size 4 x 4 will be considered in the examples.

In a given block, the process is assumed to be a zero-mean Gauss-Markov process. The correlation coefficient is assumed constant, but the standard deviation of the signal in a block is modeled by a random variable, whose probability density function

55

Figure 3.8: Original segment (256 x 256) of "Girl".

Figure 3.9: Segment (256 x 256) of "Girl" after adaptive frequency modulation with bandwidth expansion factor of 1.5. The normalized carrier-to-noise ratio is 16 dB.

Figure 3.10: Segment (256 x 256) of "Girl" after fixed deviation frequency modulation with bandwidth expansion factor of 1.5. The normalized carrier-to-noise ratio is 16 dB.

$p_\sigma(\sigma_0)$ can be evaluated. The ergodic assumption allows us to identify an ensemble mean through a sample measurement. The pdf function $p_\sigma(\sigma_0)$ is shown for a class of picture, including "Girl", in figure 3.11. We see that with a very high probability the standard deviation is below 5, illustrating the large number of occurrence of relatively blank areas. The correlation coefficient varies from block to block, but this variation doesn't affect the analysis.
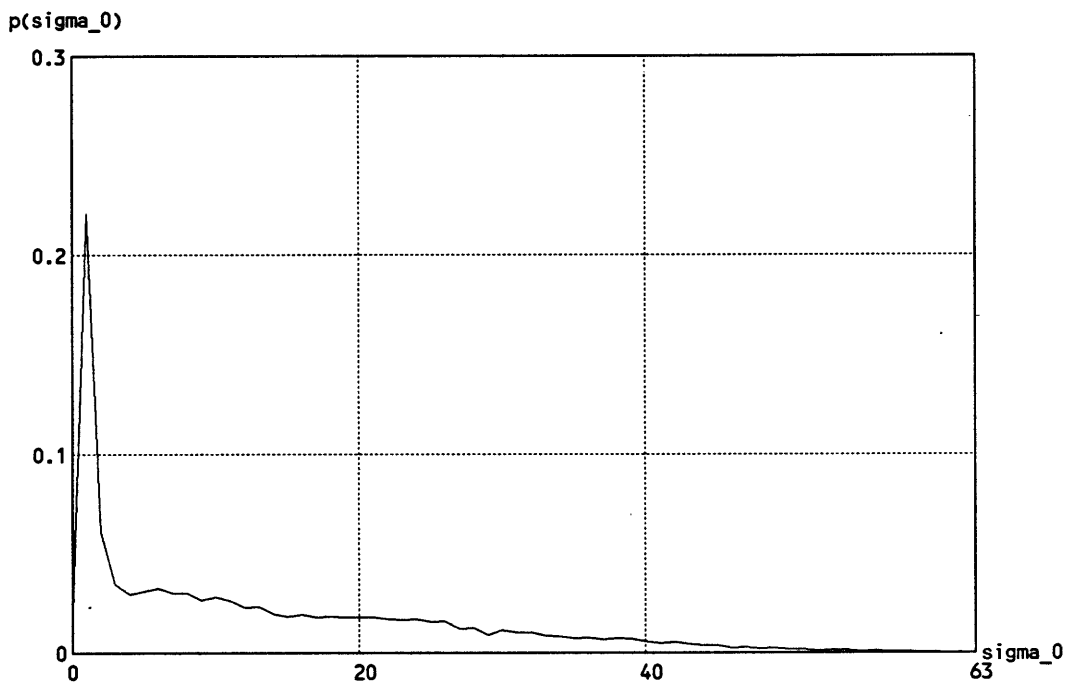
Figure 3.11: Probability distribution of the standard deviation in a block

3.4.2   Noise Improvement

Let's define the peak standard deviation σ_{max} such that this value is exceeded with a probability inferior to 0.05. Let's also define k_{peak} as the peak bandwidth expansion factor:

$$k_{peak} = \frac{\text{Max}(b_z)}{F_s} \qquad (3.23)$$

59

Then:

$$\text{SNR} = \frac{6}{16} k_{\text{peak}}^2 \left(\frac{\sigma_y}{\sigma_{max}}\right)^2 \text{CNR}_0 \tag{3.24}$$

Let's also define SSNR, the segmented signal-to-noise ratio, where both signal and noise are evaluated inside a given block only.

The SSNR in a block with standard deviation $\sigma = \sigma_{\text{max}}$ is, using (2.41):

$$\text{SSNR} = \frac{\sigma_{\text{max}}^2}{\sigma_n^2} = \frac{6}{16} \alpha^2 k_{\text{peak}}^2 \text{CNR}_0 \tag{3.25}$$

In a block with a standard deviation $\sigma = \sigma_0$, the SSNR obtained with adaptive modulation is:

$$\text{SSNR} = \frac{\sigma_0^2}{\sigma_n^2} = \frac{6}{16} \alpha^2 k_{\text{peak}}^2 \text{CNR}_0 \tag{3.26}$$

The segmented signal-to-noise ratio with adaptive modulation is the same in all blocks.

If no adaptive modulation is applied and for the same block:

$$\text{SSNR} = \frac{\sigma_0^2}{\sigma_n^2} = \frac{6}{16} k_{\text{peak}}^2 \text{CNR}_0 \frac{\sigma_0^2}{\sigma_{\text{max}}^2} \tag{3.27}$$

Hence, the coding gain G_{adap} due to adaptive modulation for a given σ is:

$$G_{\text{adap}} \,|\, (\sigma = \sigma_0) = \alpha^2 \frac{\sigma_{\text{max}}^2}{\sigma_0^2} \tag{3.28}$$

The average improvement is obtained by averaging over all possible values of σ_0.

$$G_{\text{adap}} = E(G_{\text{adap}} \,|\, (\sigma = \sigma_0)) = \alpha^2 \int_0^\infty \frac{\sigma_{\text{max}}^2}{\sigma_0^2} p_\sigma(\sigma_0) d\sigma_0 \tag{3.29}$$

For the picture "Girl", $\alpha = 0.86$ and $\sigma_{\text{max}} = 38$. Evaluation of (3.29) gives:

$$10 \log(G_{\text{adap}}) = 5 \text{ dB} \tag{3.30}$$

Table 3.1 presents the SNR computed on the whole picture for adaptive and non-adaptive FM with $k_{\text{peak}} = 1.5$ and $\text{CNR}_0 = 16 \,\text{dB}$. For comparison, a regular FM scheme without removal of the Lows but with the same peak bandwidth $\text{Max}(b_z) = 1.5 F_s$ is also computed. There is a good agreement between the measurements and the predictions.

The improvement obtained by adaptive modulation is only 5 dB. Visual inspection of figures 3.9 and 3.10 suggests that this improvement is larger. We shall see in the next section that this improvement, when masked noise is considered, is much larger.

60

	Predicted SNR [dB]	Measured SNR [dB]
Non-adaptive -luminance-	10.55	10.59
Non-adaptive -Highs-	9.38	9.37
Adaptive -Highs-	14.34	14.31

Table 3.1: SNR of different FM systems for "Girl". $Max(b_z) = 1.5F_s$. CNR0 = 16 dB.

3.4.3 Masked Noise Improvement

Adaptation of the frequency deviation in every block decreases the noise in blocks of small variance. On the other hand, blocks of large activity do not benefit from adaptation. However, they are more likely to benefit from masking. The doubly stochastic model of a block will be used here to evaluate the combined improvement of adaptation and masking.

Evaluation of the masking function at one point requires knowledge of the samples around this point in order to evaluate the signal activity. Based on the definition (2.56), we need 36 pixels values:

$$M(i,j) = M(y[i-3][j-3], y[i-2][j-3], ..., y[i+1][j+2], y[i+2][j+2]) \quad (3.31)$$

Using the statistical description of section 3.4.1, the joint probability density function is Gaussian with known correlation and mean. For a given standard deviation $\sigma = \sigma_0$, the expected masking factor is:

$$E(\zeta_m | \sigma = \sigma_0) = \int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} M(Y_0, Y_1, ..., Y_{35}) \, p_{y_0,y_1,...,y_{35}|\sigma_0}(Y_0, Y_1, ..., Y_{35} \mid \sigma_0) dY_0...dY_{35}$$

$$(3.32)$$

where $p_{y_{00},y_{01},...y_{(M-1)(N-1)}|\sigma_0}(Y_{00}, Y_{01}, ..., Y_{(M-1)(N-1)} \mid \sigma_0)$ is the joint Gaussian density of a separable two-dimensional Gauss-Markov process of correlation coefficient ρ and variance σ_0^2 [Pra78].

In order to evaluate the expected masking factor numerically conditioned on the standard deviation σ_0, a two-dimensional Gauss-Markov field is generated. The mask-

61

ing function is then evaluated, from which we get the average factor. The average masking factor $E(\zeta_m|\sigma_0)$ as a function of σ_0 is displayed in figure 3.12.
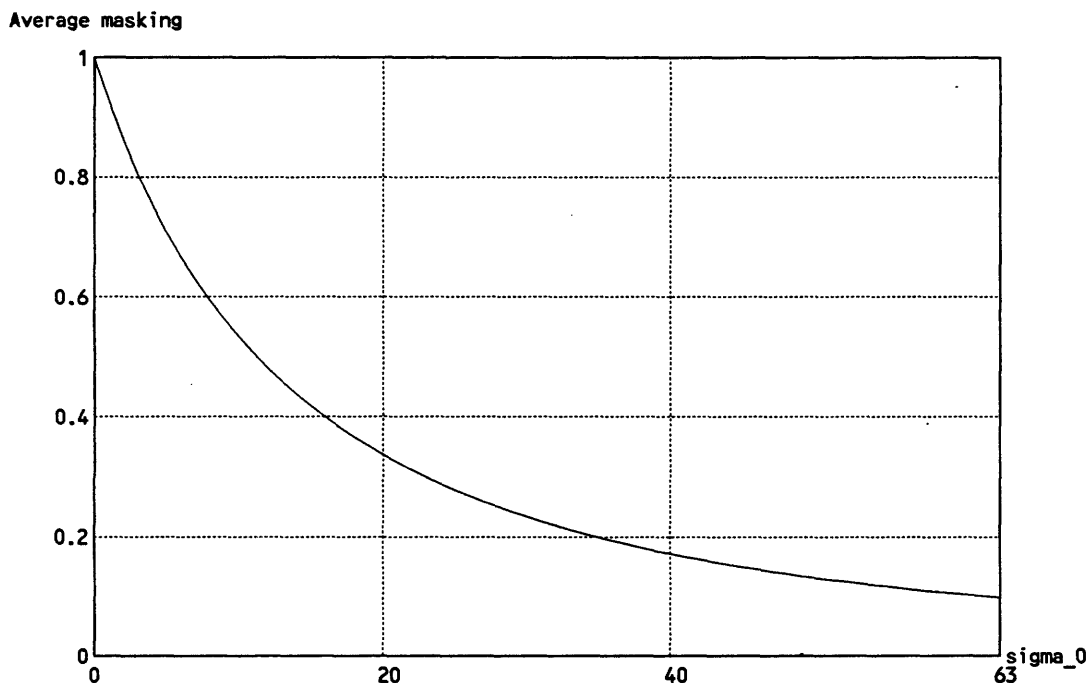
Average masking

Figure 3.12: Average masking factor for a separable Gauss-Markov process with correlation coefficient 0.97.

Finally, the combined coding gain due to adaptive modulation and masking is computed:

$$G_{\text{adap-m}} = \alpha^2 \int_0^\infty \frac{\sigma_{\text{max}}^2}{\sigma_0^2} E(\zeta_m|\sigma_0) \, p_\sigma(\sigma_0) d\sigma_0 \qquad (3.33)$$

Evaluation of (3.33) for "Girl" gives:

$$10\log(G_{\text{adap-m}}) = 16.4 \text{ dB} \qquad (3.34)$$

We see that the combined effect of adaptive modulation and masking is much larger than adaptive modulation alone. A coding gain of 16.4 dB for the same bandwidth makes this scheme very attractive.

The analysis that led to this result is based on an average masking measure. The variation of the masking measure for a given block variance and, most importantly,

62

	Predicted SNR [dB]	Measured SNR [dB]
Adaptive	14.34	14.31
Adaptive -masked noise-	25.8	23.6

Table 3.2: SNR of adaptive FM system for "Girl". $\text{Max}(b_z) = 1.5F_s$. $\text{CNR}_0 = 16$ dB.

the modeling errors, lead to slightly different experimental values. The predicted and measured SNR for "Girl" are shown on table 3.4.3. The prediction is off by 2 dB.

The main purpose of the Gaussian model is to highlight the combined effect of adaptive modulation and masking in the context of a bandwidth-limited channel. The accuracy of the model permits very reasonable prediction of the system behavior. More accurate models could be used. For example, the non-stationary image model assumes a constant correlation coefficient. Of course, ARMA parametric model would show variations of the correlation parameter, a fact well known in speech coding. However a variable correlation coefficient would only help predict the masking effect better, as the bandwidth depends only on the standard deviation in our model.

The major source of discrepancy in the model is the presence of large edges. Edges are badly modeled by a Gaussian process. This makes it difficult to predict the masking factor using a stochastic approach.

3.5 Summary

In this chapter, a new television transmission scheme is presented. It capitalizes on the inability of a human viewer to perceive noise in busy areas of the picture.

First a decomposition of the picture into Lows and Highs components is performed. The Lows component does not benefit from masking due to its low-pass nature, and has to be transmitted separately. Adaptive frequency modulation of the Highs permits a substantial improvement in noise rejection.

By adjusting the frequency deviation in each block, without exceeding the avail-

able bandwidth, an improvement of about 5 dB is obtained. As this improvement is larger in blank .areas, where viewers are most sensitive to noise, the subjective improvement is much larger. Using the masking function and a doubly stochastic model, it is evaluated as 16 dB.

Finally, a block model of the Highs components permits a reasonable prediction of the modulated signal bandwidth resulting in a possible algorithm for the adaptation of the frequency deviation. This algorithm aims at maximizing the short-time bandwidth at any given time without exceeding a given value. When the bandwidth exceeds this value, distortion in the transmission channel might arise. This is the topic of chapter 4.

Chapter 4

Transmission Distortion and Improved Adaptation Algorithms

Recent technical advances are making point-to-point communication more effective. For example, optical fibers technology allows for very large data rates, as more fibers can be connected when required. With radio wave transmission, such a situation doesn't arise. Whereas a fiber can guide the wave perfectly, antennas are bound to leak some power to neighboring sites and induce interference. As a result, the number of transmitters cannot be indefinitely increased. In this respect, it appears that the capacity of satellite links is not unlimited and some sharing of the resources must take place.

Bandwidth limitation is a typical strategy to split the information capacity. Users are allocated a frequency range and have to guarantee that very little of their signal power spreads over other frequency bands. This bandwidth limitation is performed with sharp band-pass filters. A similar restriction on the bandwidth was imposed in chapter 3. It was required that 95 percent of the power be inside a frequency band, and, when short-time spectrum estimation is considered, the worst-case situation should still satisfy the constraint.

In this chapter, the effects of bandwidth limitation on the demodulated signal are described and related to our previous approach. Alternate methods of computing the adaptation factors are then described, based on measures of the distortion that takes

65

place in the transmission link. These algorithms outperform the method described in chapter 3 by maximizing the frequency deviation without producing additional visible distortion.

Three types of impairments are generated in the channel. In order to distinguish them better in the following, we try to use a consistent terminology. Additive noise, or simply noise, usually refers to an independent error term. On the other hand, distortion will mostly refer to a dependent error term that is generated through bandlimiting of the modulated signal. Finally, clicks specifically refer to impulses generated when the noise is large compared to the modulated signal envelope.

4.1   A Satellite Link Model

In this section, it is assumed that spectrum truncation is the dominant effect of the satellite link on $z(t)$. Hence, the distortion on the phasor is a linear frequency distortion. It generates a non-linear distortion of the demodulated signal, called bandlimiting distortion, and an additional linear distortion term. Finally, the fluctuations of envelope are considered, since clicks occurrence depends on them.

A complete satellite link involves many stages of processing [PA86]. In the uplink, the frequency modulated signal is filtered so as not to interfere with other signals in adjacent bands. The received signals in each transponder are then amplified through a traveling wave tube amplifier (TWTA) which typically presents some nonlinearity. The downlink involves another bandpass filtering operation, and, finally, demodulation takes place. The combined system is complex and many types of degradations can result. For example, interference between adjacent channel occurs when filtering is insufficient. Also, intermodulation is created when two signals are combined in the nonlinear amplifier. Finally, and most importantly, truncation of the FM spectrum yields distortion in the demodulated signal.

Many of these defects can be ignored in a first approximation. First, interference can be reduced by imposing a better filtering on $z(t)$. Second, the effects of the nonlinearities in the transponder can be neglected, when only one modulated signal

66

is processed in the amplifier, as is the case in the systems proposed here. It appears that the main degradation is due to truncation of the modulated signal spectrum, and a single band-pass operation can accurately model the whole transmission link. This conclusion validates the FM transmission model in figure 2.4. However, in this chapter, the effects of the IF filtering on the demodulated signal are no longer ignored.

4.1.1 Spectrum Truncation

In this section, the band-pass filtering operation is approximated by a low-pass filtering operation, and an equivalent low-pass filter is derived. The combined band-pass filtering that takes place in the transmission link is denoted $h_{\text{IF}}(t)$.

The received phasor $\tilde{z}'(t)$ is, when no transmission noise is present:

$$\tilde{z}'(t) = \int_0^\infty h_{\text{IF}}(\tau)\tilde{z}(t-\tau)d\tau \tag{4.1}$$

Defining the equivalent low-pass process $\breve{z}(t)$:

$$\breve{z}(t) = \tilde{z}(t)\exp(-j2\pi f_c t) \leftrightarrow S_{\breve{z}}(f) = S_{\tilde{z}}(f+f_c), \tag{4.2}$$

we can express equation (4.1) as:

$$\tilde{z}'(t) = \breve{z}'(t)\exp(j2\pi f_c t), \tag{4.3}$$

where

$$\breve{z}'(t) = \int_0^\infty \breve{h}(\tau)\breve{z}(t-\tau)d\tau \tag{4.4}$$

and

$$\breve{h}(t) = h_{\text{IF}}(t)\exp(-j2\pi f_c t) \leftrightarrow \breve{H}(f) = H_{\text{IF}}(f+f_c) \tag{4.5}$$

It is seen here that the problem can be treated as if heterodyned down to DC.

The filter $\breve{h}(t)$ is complex-valued, as its Fourier transform no longer satisfies the complex-conjugate symmetry property. $\breve{H}(f)$ consists of two pass bands, one centered at $f = 0$, and the other centered at $f = -2f_c$. However, for large carrier frequency, $S_{\breve{z}}(f)$ is close to zero at negative frequencies, and removing the second pass band of $\breve{H}(f)$ centered at $f = -2f_c$ contributes to very little inaccuracy.

Hence, we can approximate $\breve{h}(t)$ by $h^+(t)$, defined by:

$$H^+(f) = \begin{cases} \breve{H}(f) & ; \ f > -f_c \\ 0 & ; \ \text{otherwise} \end{cases} \tag{4.6}$$

When $H^+(f)$ is conjugate symmetrical around $f = 0$, $h^+(t)$ is real. The class of filter described by a real low-pass filter might seem restrictive. If required, more general filters can be obtained by letting $h^+(t)$ be complex. In practice, however, the symmetry condition is often satisfied, and a real filter can be used.

The low-pass approximation is implemented in the computer simulation. band-pass processes are not well suited to a computer representation.

4.1.2 Phase and Envelope Distortion

In the previous section, it was found that the transmission link behaves like a low-pass filter, which can usually be taken as having a real impulse response. In this section, the effect of the filtering on the instantaneous phase and envelope is introduced, and the resulting distortions on the transmitted waveform are described. We use the word distortion here to stress that it depends on the signal. It is shown that two types of distortion appear. One is linear frequency distortion, which does not depend on the frequency deviation, while the other is nonlinear, and depends on the deviation.

First, the phasor $\breve{z}(t)$ is expressed in form of phase and envelope:

$$\breve{z}(t) = r(t) \exp(j\phi(t)) \tag{4.7}$$

The instantaneous phase of $\breve{z}(t)$ before transmission is $\phi(t)$:

$$\phi(t) = \frac{2\pi D}{y_0} \int_{-\infty}^{t} y(\tau) d\tau \tag{4.8}$$

The instantaneous envelope $r(t)$ is constant:

$$r(t) = A_c \tag{4.9}$$

The phase $\phi'(t)$ of $\breve{z}'(t)$ after transmission can be expressed as:

$$\phi'(t) = Im[\log(\frac{h^+(t) * \breve{z}(t)}{H_{\text{IF}}(f_c)}) + \log(H_{\text{IF}}(f_c))] = \theta(t) - \phi_0 \tag{4.10}$$

68

It consists of a time-varying term $\theta(t)$ and a fixed offset ϕ_0, which is the phase shift of $H_{\text{IF}}(f)$ at $f = f_c$. As the bias ϕ_0 can be removed, it is ignored in the following. Equation (4.10) simply expresses mathematically that $\phi'(t)$ is the instantaneous phase of the filtered phasor.

The time-varying term $\theta(t)$ can be expanded as a sum of two terms, where the first term is linear in $\phi(t)$ and the second is nonlinear [BR68].

$$\theta(t) = \Phi(t) + Im[\log(\frac{\int_0^\infty h^+(\tau) \exp j(\phi(t - \tau) - \Phi(\tau))d\tau}{H_{\text{IF}}(f_c)})] \qquad (4.11)$$

with

$$\Phi(t) = h^+(t) * \phi(t) \qquad (4.12)$$

The linear term $\Phi(t)$ is simply the convolution of $\phi(t)$ with $h^+(t)$.

Equation (4.11) is a convenient way to expand the received phase $\phi'(t)$ into two terms, only one of which depends on the frequency deviation.

The second term of (4.11) translates into a nonlinear distortion term $e_d(t)$ of the demodulated signal. It is a function of $y(t)$, of the frequency deviation D, and of the filter characteristic.

$$e_d(t) = \frac{1}{2\pi\nu} \frac{d}{dt}(\phi(t) - (h^+(t) * \phi(t))) \qquad (4.13)$$

The nonlinear distortion $e_d(t)$ will be referred to as bandlimiting distortion in the following.

Similarly, the envelope of the received phasor is:

$$r'(t) = \exp(Re[\log(h^+(t) * \breve{z}(t))]) \qquad (4.14)$$

When the frequency deviation is large, some power is lost in the transmission process. This is manifested by a drop of the envelope. Actually, the envelope $r'(t)$ can be zero when the signal spectrum is completely outside the channel pass band, in which case no signal is transmitted. Obviously, the instantaneous phase is undetermined in this case.

In the following, we characterize $e_d(t)$ as a function of $y(t)$ and D. The effects of the linear filtering are also investigated. The envelope distortion is shown in figure 4.1, together with the bandlimiting distortion to illustrate the effect of spectrum truncation.
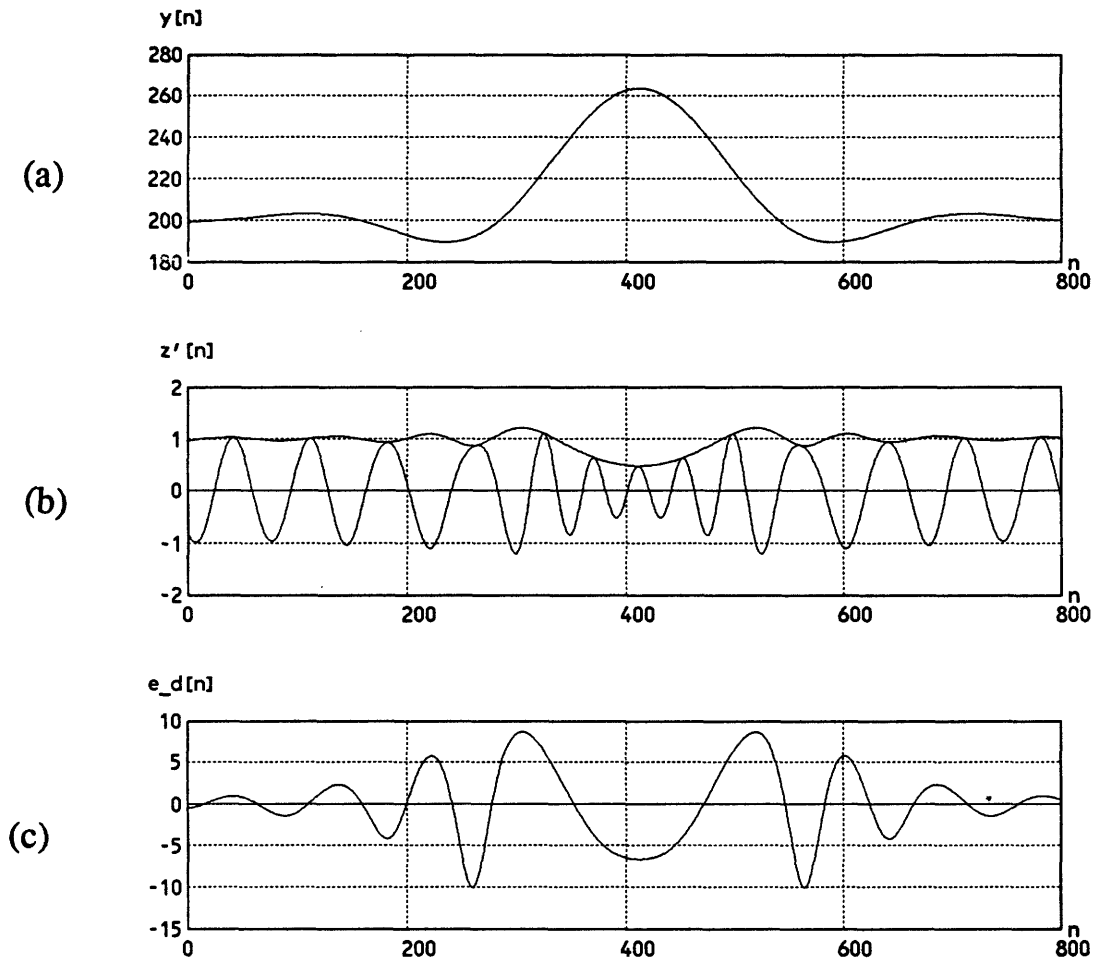
Figure 4.1: Illustration of envelope distortion and bandlimiting distortion. (a) bandlimited impulse. (b) modulated signal shown with envelope. (c) bandlimiting error.

4.1.3 Limitations Due to Linear Filtering

In the previous section, it was shown that one effect of bandlimiting $z(t)$ on $y'(t)$ was equivalent to filtering the message $y(t)$ with $h^+(t)$. In this section, we investigate this

effect on $y(t)$, where $y(t)$ is a continuous-time signal, generated from a discrete source, as shown in section 3.2. The discrete source is a component of the picture sequence. In this case, the channel must satisfy the Nyquist criterion in order to guarantee no intersymbol interference. One necessary condition is:

$$B_{\text{IF}} > \text{Max}(r_l) \tag{4.15}$$

where r_l is the transmission rate [samples/sec] of the component l.

Similarly, when the equivalent low-pass filter is considered:

$$B^+ > \frac{1}{2}\,\text{Max}(r_l) \tag{4.16}$$

where B^+ is the bandwidth of $h^+(t)$, which is half that of $h_{\text{IF}}(t)$.

It appears that the linear distortion on the demodulated signal imposes an upper bound on the transmission rate.

The visual effect of linear filtering is assessed in a series of experiments. We cover two cases; one is filtering of the Highs component, and the other is filtering of the subbands issued from the Highs, when subband decomposition is performed before transmission. Subband decomposition is introduced at this point to be more complete about the effect of linear filtering. It is further described in chapter 4. Suffice it to say here that the decomposition is reversible when the analysis filters, which split the signal into subbands, and the synthesis filters, which recombine the subbands, respect the perfect reconstruction condition [Vet87]. It will be shown that in order to do so, the Nyquist criterion must be satisfied. In the following example, a four-by-four decomposition is performed to generate 16 discrete time components.

Linear Filtering of the Highs

The ideal situation described in section 3.2 no longer holds when linear distortion occurs. Obviously, the bandwidth B^+ is equal to or larger than half the rate r, but some attenuation and phase distortion can be present in the pass band. This in turn violates Nyquist criteria for zero intersymbol interference. A zero-forcing equalizer can take care of removing the interference at the expense of an increased noise [Pro83].

The question arises as to whether this equalizer is needed. The perceptual effect of linear filtering is investigated in order to answer this question.

In figure 4.2, spectrum truncation is performed with a Butterworth filter of 3rd and 11th order. The cut-off frequency is set so as to simulate a very-narrow-band FM case, namely the channel bandwidth B_{IF} is equal to the rate of the source. Neither the phase distortion, nor the magnitude distortion, can be noticed. In this case, a zero-forcing equalizer turns out not to be necessary.

Figure 4.2: Effect of linear distortion on the Highs component. From left to right: original, Butterworth 3rd order, Butterworth 11th order. $B^+ = F_s$.

Linear Filtering of the Subbands

Filtering the Highs component proved to be non critical, as only very high horizontal frequencies of the reconstructed picture are altered by the operation. Such is not the case when subband transmission is considered. The intersymbol interference in this case violates the perfect-reconstruction property of the subband analysis and synthesis filters. Its effect is very detrimental, as shown in figure 4.3 where the IF filtering is performed on the phasor modulated by the subbands. In this respect, it appears that when a subband decomposition is performed, the Nyquist criteria should be satisfied. In practice, a zero forcing equalizer is inserted before the subband

synthesis. Figure 4.3(c) shows the reconstructed picture when the Nyquist criterion is satisfied. Still, some degradation can be observed. It is due to the filters used in the analysis and synthesis operation, which do not exactly satisfy the perfect-reconstruction condition, and to the filter $h^+(t)$, which does not exactly satisfy the Nyquist condition. Nevertheless, a large improvement can be seen, when the perfect-reconstruction condition is close to being satisfied. The approximation of an ideal low-pass filter impulse response used in figure 4.3 is obtained by windowing a sinc function with a Hamming window.

Figure 4.3: Effect of linear distortion on the four by four subbands issued from the Highs component. From left to right: Butterworth 3rd order, Butterworth 11th order, approximation of an ideal low-pass filter. $B^+ = F_s$.

4.2   Bandlimiting Distortion

The bandlimiting distortion $e_d(t)$ is the nonlinear term of the error due to spectrum truncation, and is a complex function of the signal, the frequency deviation, and the channel characteristic. It can be derived by evaluating equation (4.11), or from its series expansion [BR68]. In our case, it is simply obtained by direct computation of the phasor transient, and evaluation of its instantaneous phase. If $y'(t)$ is the

73

demodulated signal obtained by computer simulation when no noise is present, the bandlimiting distortion is:

$$e_d(t) = y'(t) - (y(t) * h^+(t)) \qquad (4.17)$$

A complete description of $e_d(t)$ is beyond the goal of this research, and also of limited interest. What is needed here is a simple measure of $e_d(t)$ which can be useful to the designer of a communication system.

Two approaches are possible to define a meaningful measure of distortion. The first one investigates the distortion when the message is a well defined test signal. The second one characterizes the distortion signal by some statistics, such as RMS value, or peak value, etc.. In the latter case, it is evaluated with a message consisting of a typical signal, such as a possible realization of a Gaussian process, or a set of test pictures. These two approaches are investigated in the context of the bandlimiting distortion.

4.2.1 Sinewave Test Signal

The most common test signal is a sinewave. In this case, the harmonic distortion measure is commonly used. It is the ratio of the power of a given harmonic to the power of the fundamental frequency. This test signal is of limited interest in the field of image processing, but is often used to characterize audio systems. In our case, this measure will be used as a first experiment to verify the low-pass complex-envelope model and its computer simulation.

Figure 4.4 compares the third-harmonic distortion generated in a first-order Butterworth filter from laboratory measurements [Rut68] and from the computer simulation. The filter used in the laboratory is a band-pass filter centered at 70 MHz and of bandwidth 2.45 MHz. A very good agreement can be observed, although the computer model uses a symmetrical low-pass Butterworth filter. A band-pass Butterworth filter is not conjugate symmetrical with respect to its center frequency. It should be noted that alternative methods exist to compute $e_d(t)$ in this case based on a Fourier series approach [Rut68].
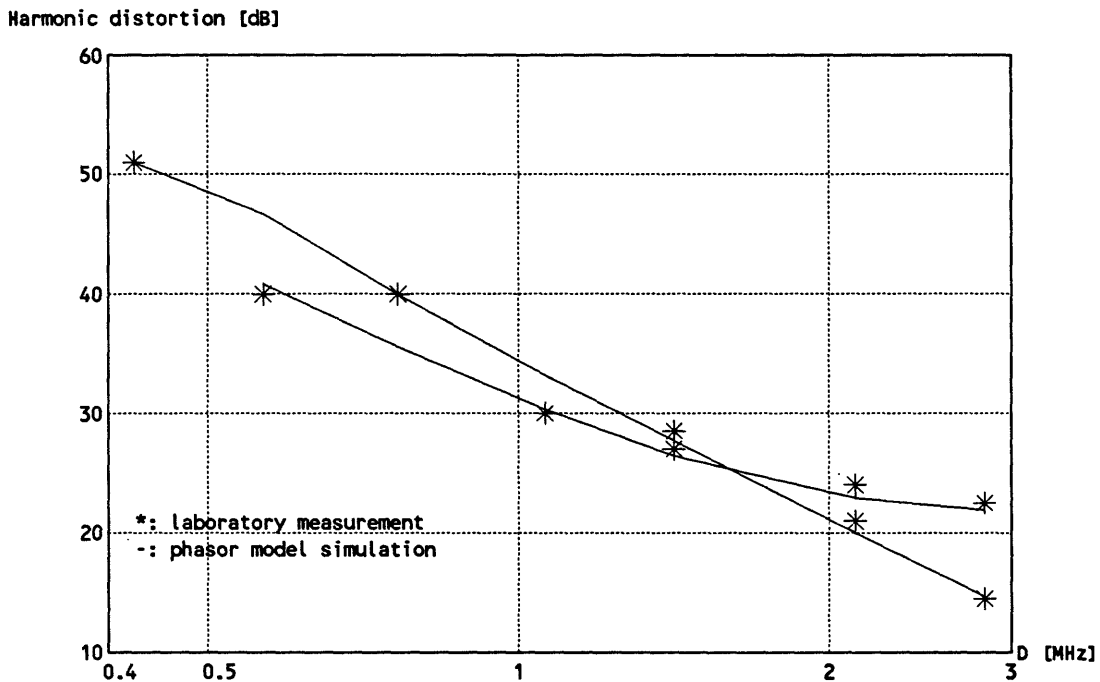
Figure 4.4: Third-harmonic distortion ratio of a first-order Butterworth filter for sinewaves at two different frequencies.

The distortion mechanism in this case can be best explained by the attenuation and phase shift imposed on each frequency lines of the phasor spectrum.

4.2.2 Impulse Test Signal

Another important class of test signal is impulse-like waveforms [Wea71], such as the sinc function. The sinc function is the perfectly bandlimited impulse:

$$y(t) = A_p \frac{\sin(2\pi F_s t)}{2\pi F_s t} \tag{4.18}$$

Impulse test signals are better suited to television engineering as they describe some typical critical pictures such as bars.

In this case, a measure based on the peak value of the distortion is found more appropriate than a mean-square description. Typically, viewers object to the worst part of the picture.

In the following, the behavior of the test signal in different channels is investigated.

75

The influence of the frequency deviation is investigated for two ratios of channel bandwidth to signal basebandwidth, and for three types of channels, namely 3rd order Butterworth, 11th order Butterworth, and finally an approximation of an ideal low-pass filter.

The effective peak-to-peak frequency deviation is :

$$Ptp(F_i(t)) = \frac{A_p D_0}{y_0}$$ (4.19)

where $Ptp()$ is a function that returns the peak-to-peak value of a function.

The measurements are reported in figure 4.5. A typical feature of the distortion behavior is that it increases abruptly when the peak-to-peak effective frequency deviation approaches the channel bandwidth. The jump occurs for a ratio $Ptp(F_i(t))/B_{\text{IF}}$ of about one, for channels with a steep transition between pass and stop band, and a slightly larger ratio when the transition is smoother. This behavior can be understood by realizing that the carrier gets attenuated when its instantaneous frequency is equal to or larger than the cut-off frequency.

It appears that the concept of instantaneous frequency can give us some insight about the filtering mechanism. If the frequency were to be varied very slowly, the carrier could be seen as a locally fixed frequency sinewave. When the sinewave frequency is in the transition bandwidth, i.e. when the peak-to-peak frequency deviation equals the channel bandwidth, some attenuation takes place, and the transmission is no longer ideal. This interpretation is known as the quasi-static approximation. This approximation is true in our case, as $y(t)$ is about constant as we move away from the origin.

As shown in figure 4.5, the bandlimiting distortion doesn't get exceedingly large as long as the peak-to-peak deviation is below the channel bandwidth. In this case and in our experience, the distortion was never visible. However, in practice, it is not possible to increase D close to its critical value. This becomes apparent when the behavior of the carrier envelope is investigated.

The minimum value of the envelope $Min(r(t))$ is shown in figure 4.6 as a function of the ratio of peak-to-peak deviation to channel bandwidth. For large deviations, a

76

Figure 4.5: Peak bandlimiting distortion for impulse bandlimited at F_s as a function of the ratio of peak-to-peak deviation to channel bandwidth. (a) 3rd order Butterworth. (b) 11th order Butterworth. (c) approximation of ideal low-pass filter.

drop of the envelope is observed. The sudden increase of distortion that was observed in figure 4.5 occurs for deviation values where the envelope has dropped to zero. It appears here that the demodulator cannot track the carrier phase for such large deviations, because the phasor magnitude is zero.

Envelope distortion is closely related to bandlimiting distortion as both arise because of the phasor filtering. In the following, both envelope and bandlimiting distortion are measured.

4.2.3 Interaction of Noise and Envelope Distortion

Envelope distortion is especially detrimental when the effect of noise is combined with that of spectrum truncation. This is the reason why it is not possible to increase the peak-to-peak deviation close to its critical value.

In chapter 2, it was assumed that the noise level was small compared to that of the carrier. When the weak-noise assumption no longer holds, a drop of performance occurs, as measured in figure 2.5. The loss of performance is due to spurious noise impulses, usually referred to as clicks. Clicks appear when the channel noise $n(t)$ deviates the tip of the phasor in such a way that encirclement of the origin takes place. A good description of the phenomenon can be found in [JI65]. An approximation of the upper bound of the probability of a click occurrence is [JI65]:

$$P(\text{click}) = \frac{1}{\sqrt{3}} \frac{B_{\text{IF}}}{F_s} \; Q(\frac{A_c}{\sqrt{N_0 B_{\text{IF}}}}) \qquad (4.20)$$

where the function $Q(x)$ is:

$$Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp(\frac{-u^2}{2}) du \qquad (4.21)$$

The function $Q(x)$ is extremely sensitive to its argument, as demonstrated in table 4.1, where the probability of a click is computed for different carrier-to-noise ratio. It appears that a carrier-to-noise ratio of about 12 dB is required when a negligible amount of clicks can be tolerated. This amounts to:

$$A_c > 5.6 \; \sqrt{N_0 B_{\text{IF}}} \qquad (4.22)$$

78

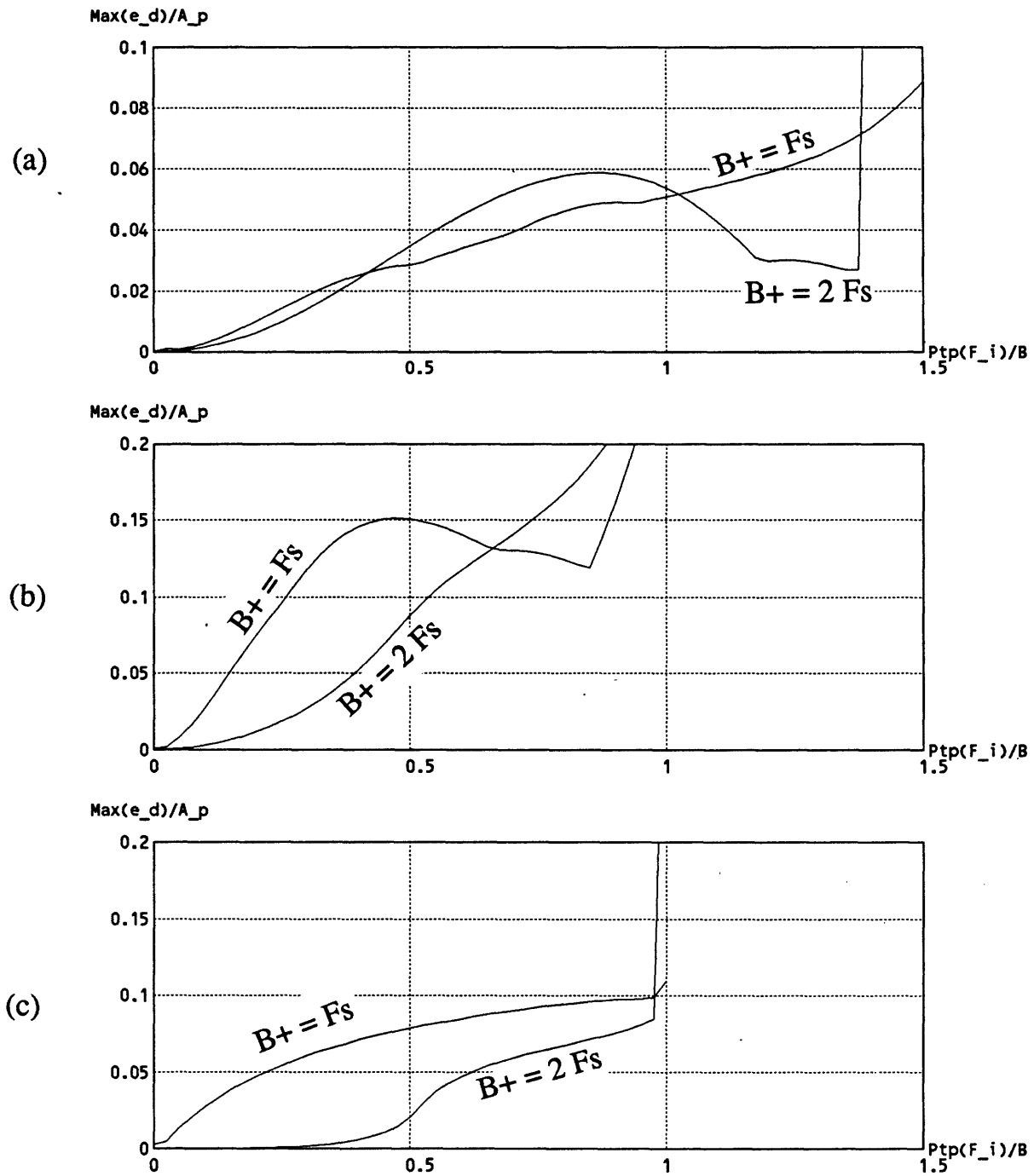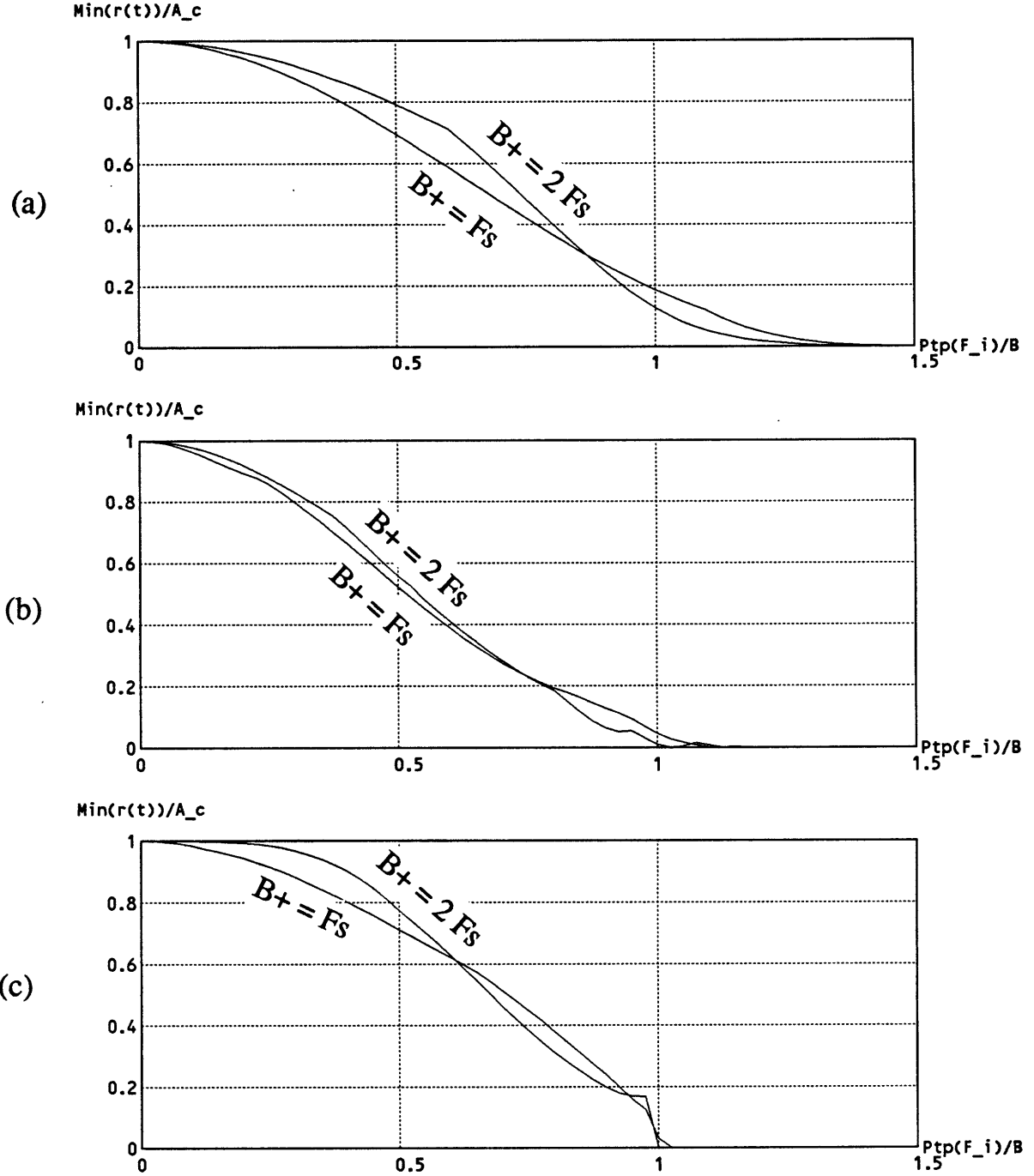Figure 4.6: Minimum of envelope for impulse bandlimited at F_s as a function of the ratio of peak-to-peak deviation to channel bandwidth. (a) 3rd order Butterworth. (b) 11th order Butterworth. (c) approximation of ideal low-pass filter.

CNR [dB]	P(click)	# clicks in a 500x500 frame
12	10^{-8}	0.003
11	$3.5\ 10^{-7}$	0.09
10	$4.6\ 10^{-6}$	1.1
9	$3.5\ 10^{-5}$	9

Table 4.1: Click occurrence for different CNR. $B_{IF} = 2F_s$

Similarly, if clicks are to be negligible in any segment of the picture including those with large envelope distortion, the instantaneous envelope satisfies:

$$\text{Min}(r(t)) > 5.6\ \sqrt{N_0 B_{\text{IF}}} \tag{4.23}$$

Hence, if the constraint is to guarantee rare occurrence of clicks, the deviation must be set so that the inequality (4.23) is satisfied. Interestingly, the maximum allowable deviation depends on the noise condition of the channel: the cleaner the channel, the larger the possible deviation. The effect of excessive deviation is illustrated in figure 4.7 where the deviation is increased from line to line. For a ratio $Ptp(F_i(t))/B_{\text{IF}}$ larger than 0.7, clicks degrade the picture when the CNR is 12 dB. Similarly, for a CNR of 9 dB, a value of $Ptp(F_i(t))/B_{\text{IF}}$ of 0.5 only produces occasional clicks. It should be noted that clicks occur close to the edge, where the envelope drop takes place. Hence, envelope distortion is not directly detrimental to the picture quality, but when noise is present in the channel, it has to be taken into account in the design.

4.2.4 Pictures as Test Signals

In the previous section, it was demonstrated that for some test signals, both envelope distortion and bandlimiting distortion occurs as the frequency deviation gets larger. In this section, a similar measurement is performed for pictures where the peak short time bandwidth of the modulated signal is varied.

In the following experiments, the picture "Girl" is used. It is considered a typical picture as it contains a large variety of texture, edges, and different sizes of objects. In
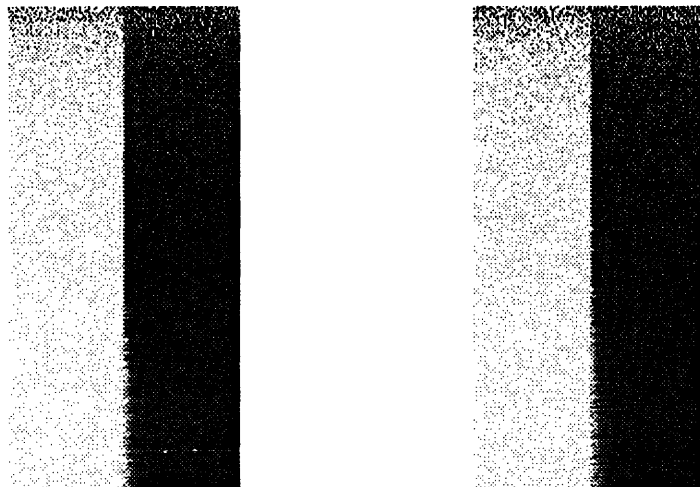
Figure 4.7: Effect of envelope drop on click occurrence. (a) CNR = 12 dB. (b) CNR = 9 dB. $Ptp(F_i(t))$ varies from line to line from 0 to B_{IF}. $B_{IF} = 2F_s$

turn, we shall investigate the luminance component, the Highs component, different subbands obtained from the Highs, and, finally, a white noise field. The peak value of bandlimiting distortion and envelope distortion are evaluated for two ratios of channel bandwidth to signal bandwidth, namely $B^+ = F_s$, for very narrow band FM, and $B^+ = 2F_s$, for narrow band FM. The channel is modeled by a 11th order Butterworth filter. The peak value of the distortion is the largest value of $e_d(t)$ that appears in the test sequence. All the test sequences are of the same size.

In section 3.4.1, a statistical model for pictures was introduced. The model describes the pixels in a block of small size as a Gaussian random process, whose standard deviation σ is itself a random variable. This model can be used to predict the short-time bandwidth of the modulated signal in a compact way. In particular, the peak short time bandwidth can be estimated by:

$$\text{Max}(b_z) = \nu \, \sigma_{\max} = \frac{D}{y_0} \sigma_{\max} \tag{4.24}$$

where $\sigma_{\max}$ is defined by:

$$P(\sigma < \sigma_{\max}) = 0.99 \tag{4.25}$$

81

Hence, when the pdf of σ is estimated, a compact representation of the modulated signal bandwidth is available. It is shown in the following that the distortion behavior as a function of $\text{Max}(b_z)$ is similar for a large class of signals. This substantiates the assumption that second-order statistics alone can predict the system performance.

In this series of experiments, we have measured the peak value of the distortion and the minimum value of the envelope. This approach provides us with a gross characterization of the distortion process, but the measurements are found useful as a design guideline.

In figure 4.8, the peak distortion is shown for the luminance signal. The luminance is preemphasized before being transmitted. The preemphasis function found in MAC systems [Rho85] is used for this purpose. Preemphasis is used here as a mean to balance the three impairments encountered in bandlimited FM transmission, namely additive noise, bandlimiting distortion, and clicks. Because the low frequencies are attenuated by the preemphasis network, the carrier frequency remains in the middle region of the channel pass band. Slow frequency shifts of the carrier close to the transition band are thus avoided, together with the increased occurrence of clicks. In a way, preemphasis favors bandlimiting distortion around edges in order keep clicks occurrence unlikely. This preemphasis network does not change noise visibility.

It can be seen that a significant drop of envelope takes place before bandlimiting distortion becomes critical. The effect of the link can be ignored when the peak short time bandwidth is smaller than $0.3B_{\text{IF}}$. This corresponds to an effective peak-to-peak frequency deviation of about $0.6B_{\text{IF}}$.

A very similar distortion behavior is shown in figure 4.9 for the Highs component. In this respect, preemphasis, or equivalently deemphasis of the low frequencies, is very similar to removing the Lows, as far as distortion is concerned.

The influence of other statistics, such as the correlation function and higher order statistics, can be seen in figure 4.10, 4.11, 4.12. They show the peak distortion for three subbands obtained from the Highs. Subband decomposition is explained in more detail in chapter 5. Subbands tend to have a "whiter", more uniform, spectrum. The power spectrum density of each subband at high frequencies is larger, and produces

abrupt frequency changes of the carrier, which are adversely affected by the channel.

Smaller bandwidths are necessary to produce a given drop of the envelope. This illustrates the more critical nature of some subbands in term of inducing distortion, mostly those corresponding to high horizontal frequencies of the original images such as those shown in figure 4.11 and 4.12. For most subbands, the effects of the link are small, if the peak short time bandwidth is below $0.2B_{IF}$.

The subbands are obtained from a 4 x 4 decomposition. The first index refers to the horizontal frequency range, while the second refers to the vertical frequency range. Each index goes from 0 to 3, from the lowest to the highest frequency range.

Finally, the same measurements are performed on a white noise sequence, in which case the correlation between pels is zero. This signal is obviously critical in term of distortion, as the carrier frequency varies very abruptly from sample to sample. As expected, the envelope drops significantly for relatively moderate bandwidths, as shown in figure 4.13. In this case, the peak short time bandwidth should be smaller than $0.25B_{IF}$ in order to prevent large envelope drops.

4.3   Adaptation Algorithms

In this section, two new algorithms to compute the adaptation factors are presented. In one algorithm, the frequency deviation in each block is found so that the effective peak frequency deviation in each block is constant. The second algorithm is iterative in that the deviation is iteratively adjusted in each block until the distortion reaches a given value. Both algorithms are related to the method presented in chapter 3. Here a maximum occupancy of the channel bandwidth is not obtained through a prediction of the signal bandwidth, but, rather, by considering the effect of spectrum truncation. In this respect, improved performance can be expected without increasing the transmission distortion.
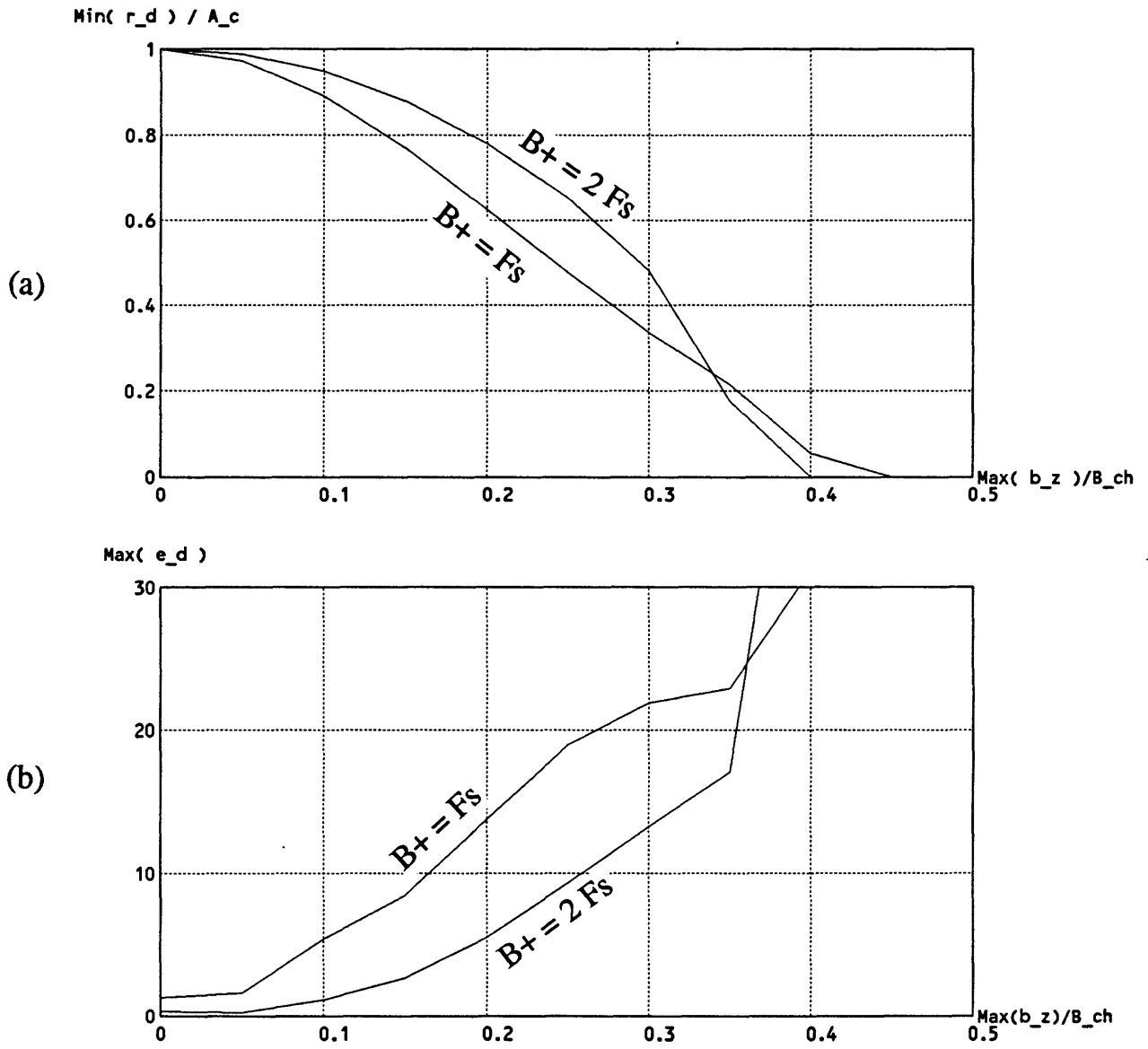
Figure 4.8: Peak distortion measured for the preemphasized luminance signal.

(a)

(b)
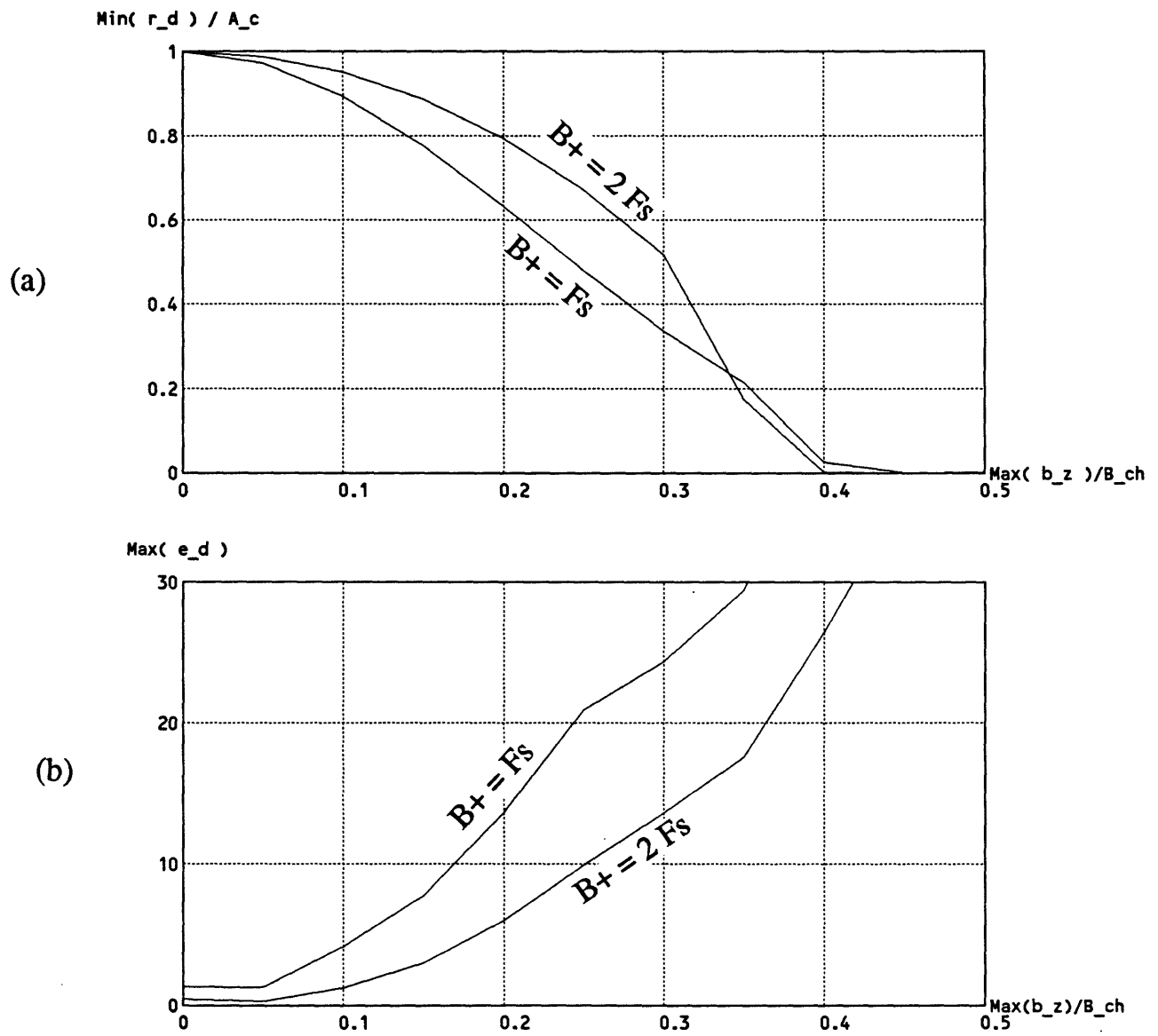
Figure 4.9: Peak distortion measured for the Highs component.

Min(r_d) / A_c

(a)

Max(b_z)/B_ch

Max(e_d)

(b)

Max(b_z)/B_ch
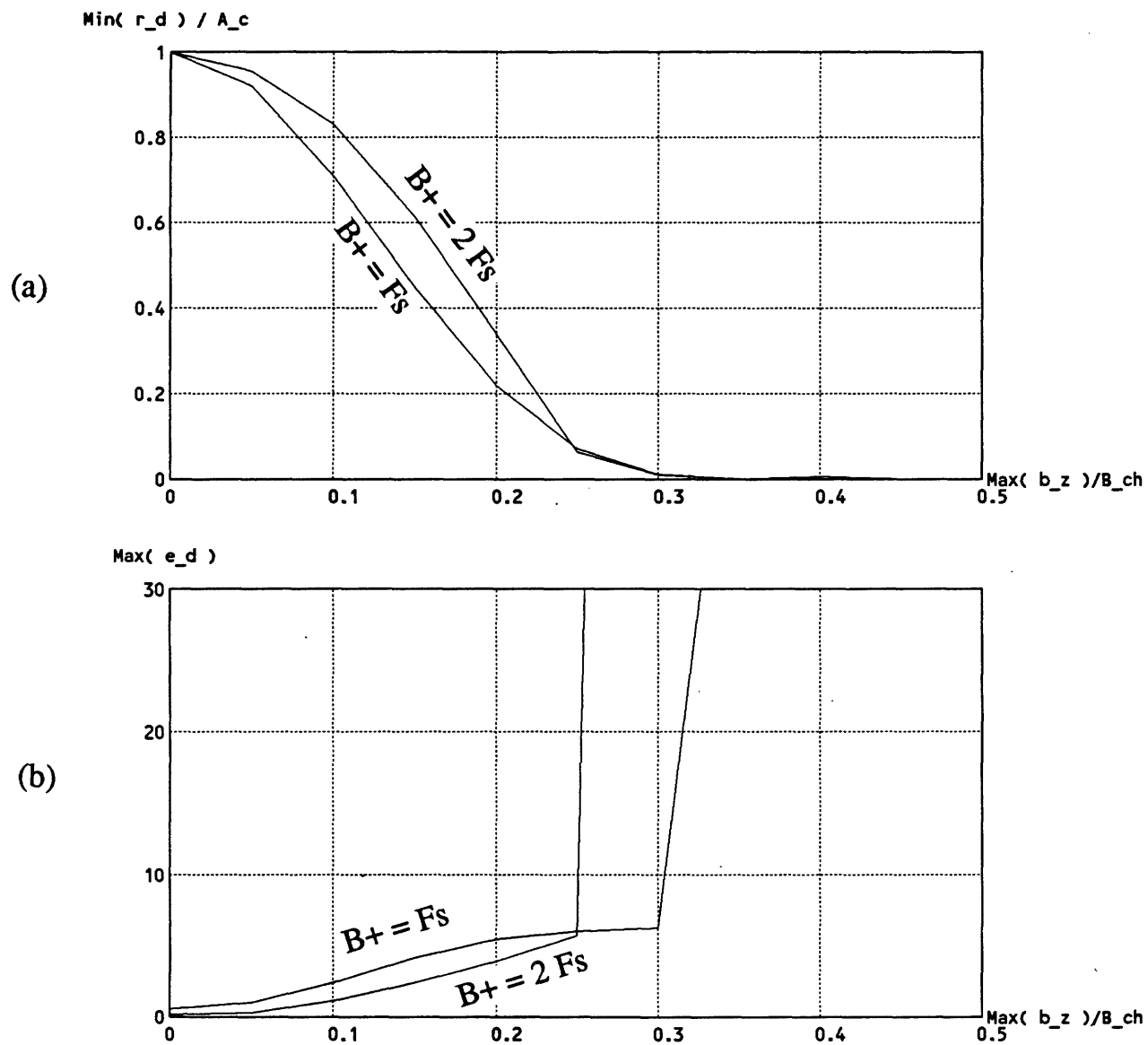
Figure 4.10: Peak distortion measured for the subband -0-3.

86

(a)

(b)
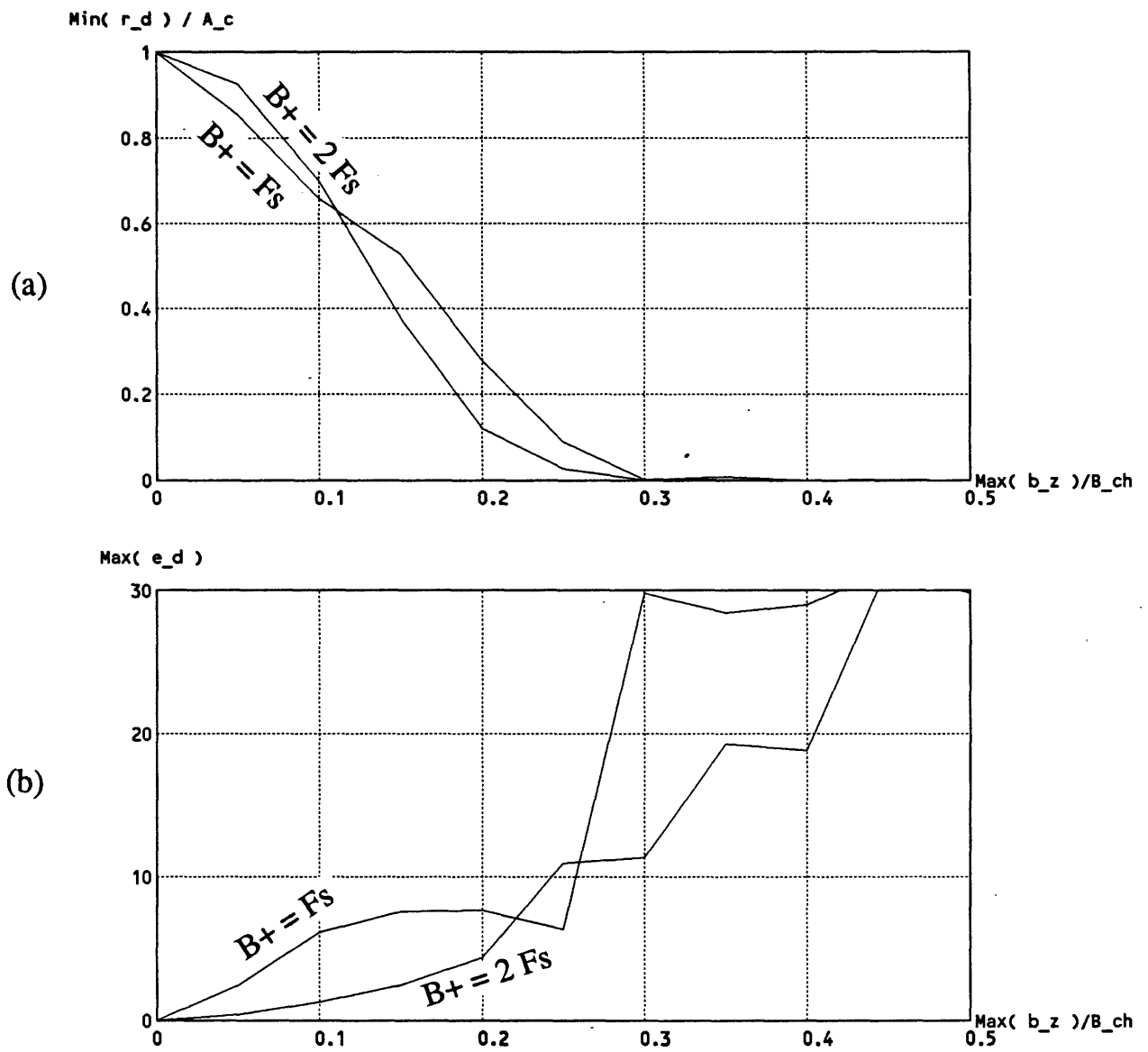
Figure 4.11: Peak distortion measured for the subband -3-0.
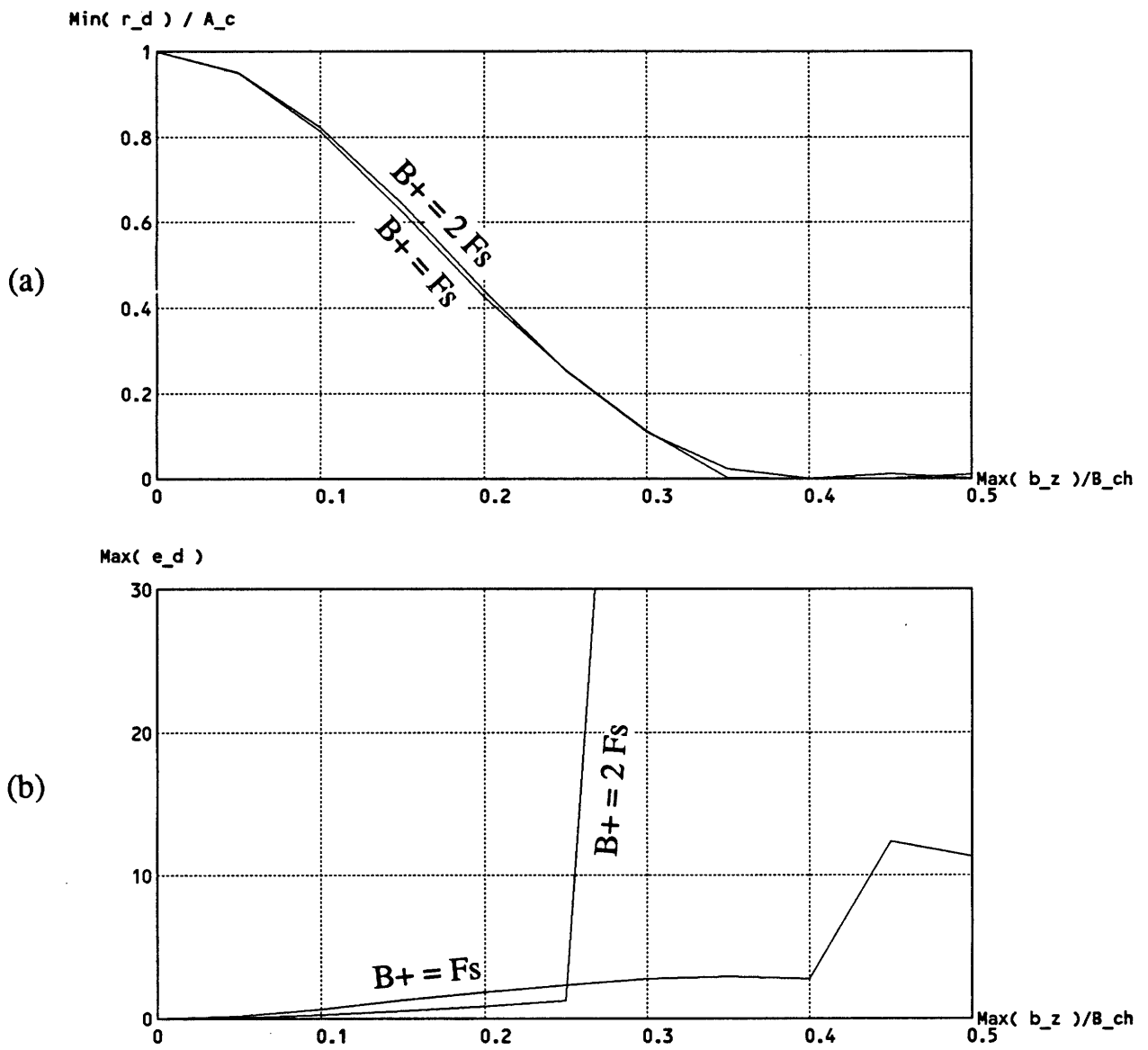
Figure 4.12: Peak distortion measured for the subband -3-3.

(a)

Min(r_d) / A_c

$B+ = 2\,Fs$

$B+ = Fs$

Max(b_z)/B_ch

(b)

Max(e_d)

$B+ = Fs$

$B+ = 2\,Fs$

Max(b_z)/B_ch
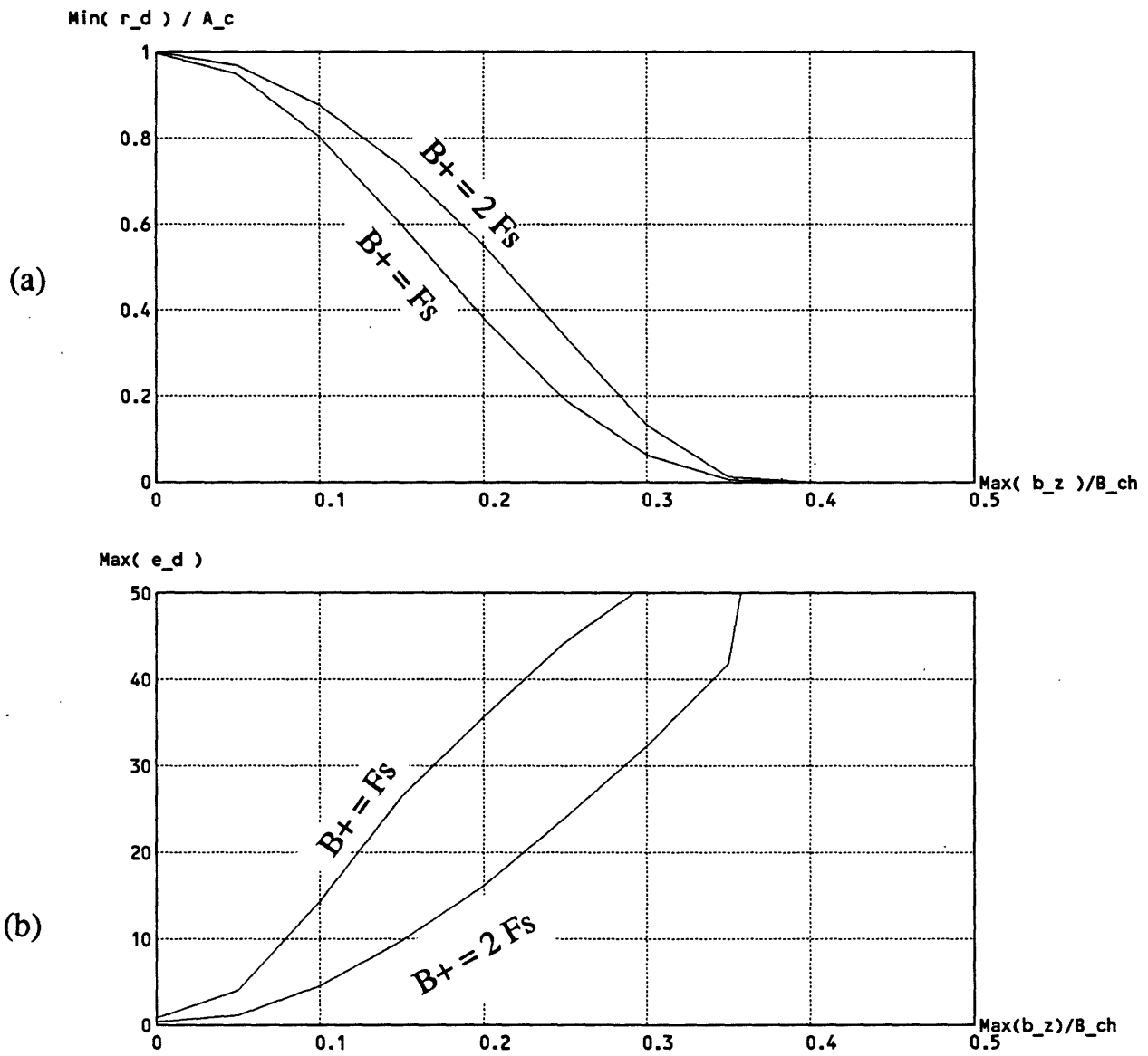
Figure 4.13: Peak distortion measured for a white noise field.

4.3.1 Effective Peak Frequency Deviation Algorithm

In the previous section, it was shown for both a test signal and typical picture material that if the effective peak frequency deviation is only a fraction of the channel bandwidth B^+, then the bandlimiting distortion is well behaved and not visible. Hence, a possible adaptation algorithm that ensures limited distortion is to bound the effective peak frequency deviation (EPFD) in a small block of the picture. This contrasts with the constant-banwidth algorithm where short-time bandwidth measures are used. Hence:

$$D(m,n) = f_a(m,n)D_0 \qquad (4.26)$$

with

$$D_0 = \nu y_0 \qquad (4.27)$$

Constant effective peak deviation in a block is obtained for:

$$f_a(m,n) = \alpha \frac{y_0}{\text{Max}(|y(t)|)} \qquad (4.28)$$

where α is introduced to correct for some possible overshoot due to the bilinear interpolation of $D(m,n)$. A value of 0.86 for α was found to be effective. Here $\text{Max}(|y(t)|)$ is taken for discrete value of t corresponding to pixels in the block (m,n).

The peak effective deviation in block (m,n) is:

$$\text{Max}(|\Delta F(t)|) = \text{Max}(|F_i(t) - F_c|) = D(m,n) \frac{\text{Max}(|y(t)|)}{y_0} = \alpha D_0 \qquad (4.29)$$

Hence, by setting the deviation D_0 smaller than the channel bandwidth B^+, $\text{Max}(|\Delta F(t)|)$ will also be smaller than B^+, resulting in a well behaved distortion function. Typically, y_0 is the largest value of $y(t)$, making $f_a(m,n)$ always larger, or equal to one.

This algorithm is compared to the constant-banwidth algorithm in figure 4.15, where the bandlimiting distortion is also shown. Both peak effective deviation and peak bandwidth are as large as possible, without producing visible distortion. In the peak deviation algorithm, the distortion is on average larger, demonstrating the effectiveness of the adaptation method. Similarly, the noise performance of the EPFD algorithm is superior.
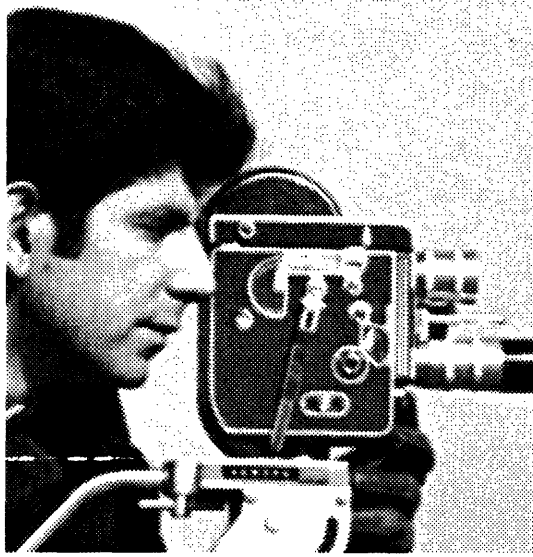
Figure 4.14: Original picture "cman"

The previously described algorithms are capable of very large noise reduction in blank areas. In blank areas, the Highs component is very small, and the frequency deviation can be increased substantially without increasing the distortion induced in the channel, yielding a large noise reduction. This reduction is so large that the deviation is usually not increased up to its maximum possible value. The deviation in our simulations was increased by a factor of 32 at most, without any loss of performance.

In busy areas, such as close to edges, the deviation cannot be increased without producing artifacts. The noise is masked by the picture content is these areas, but some noise can still be seen close to sharp edges because of the limited effect of masking. The question addressed in the next section is how to increase the frequency deviation further in busy areas while keeping the distortion bounded.

4.3.2 Iterative Algorithm

An iterative system for the computation of the adaptation factors $f_a(m, n)$ is proposed in figure 4.16. On the transmitter side, the signal as obtained after transmission is reconstructed with help of a channel model -a simple band-pass filter- and an adaptive

91
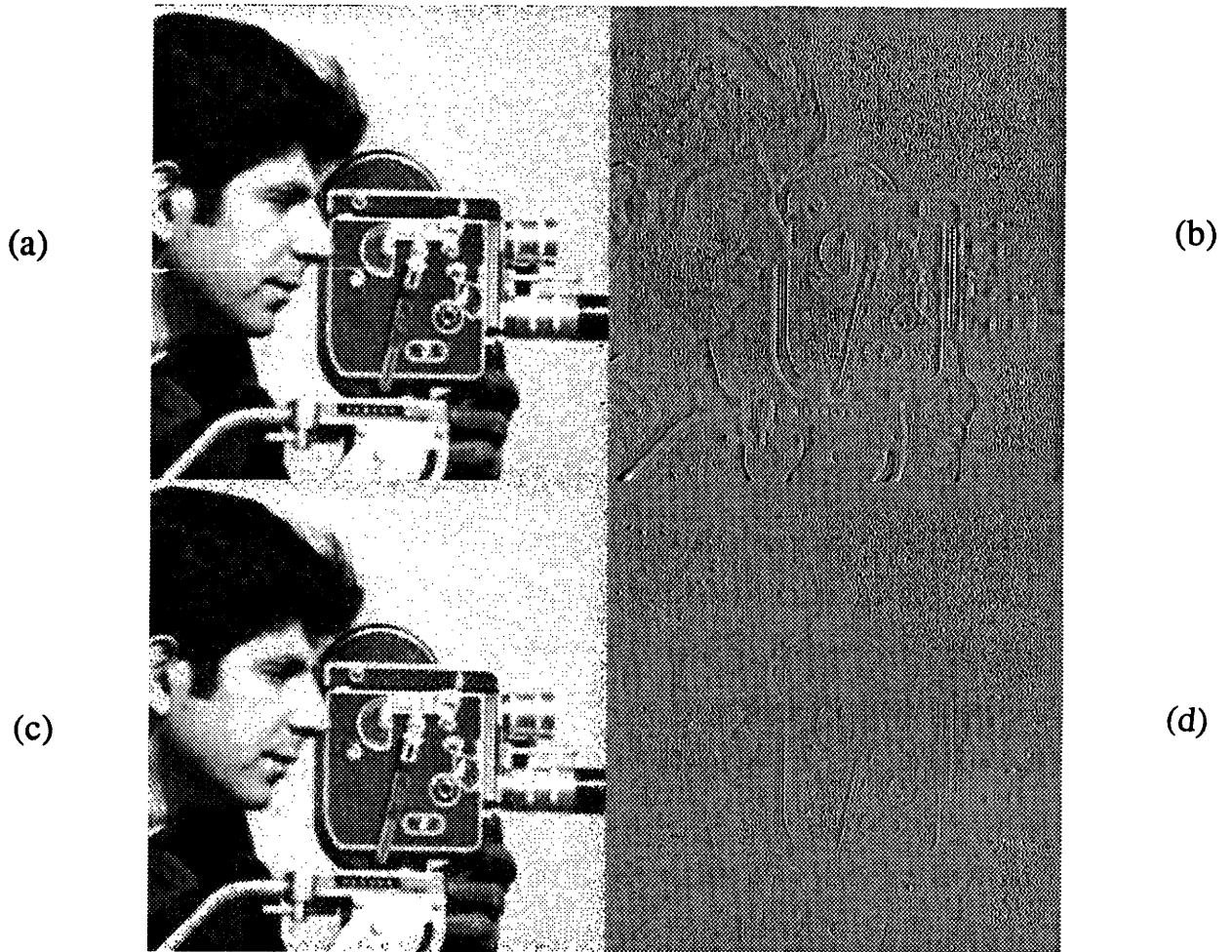
Figure 4.15: Comparison of two algorithms. (a) Effective peak frequency deviation algorithm. $D_0 = 0.8B^+$ (b) Corresponding bandlimiting distortion. (c) Constant-banwidth algorithm. Max(b_z) $= 0.4B^+$ (d) Corresponding bandlimiting distortion. $B^+ = F_s$, CNR $= 9$ dB. The bandlimiting distortion is shown scaled up by ten.

frequency demodulator. Estimation of the bandlimiting distortion takes place, which shows whether larger or smaller deviation is required.
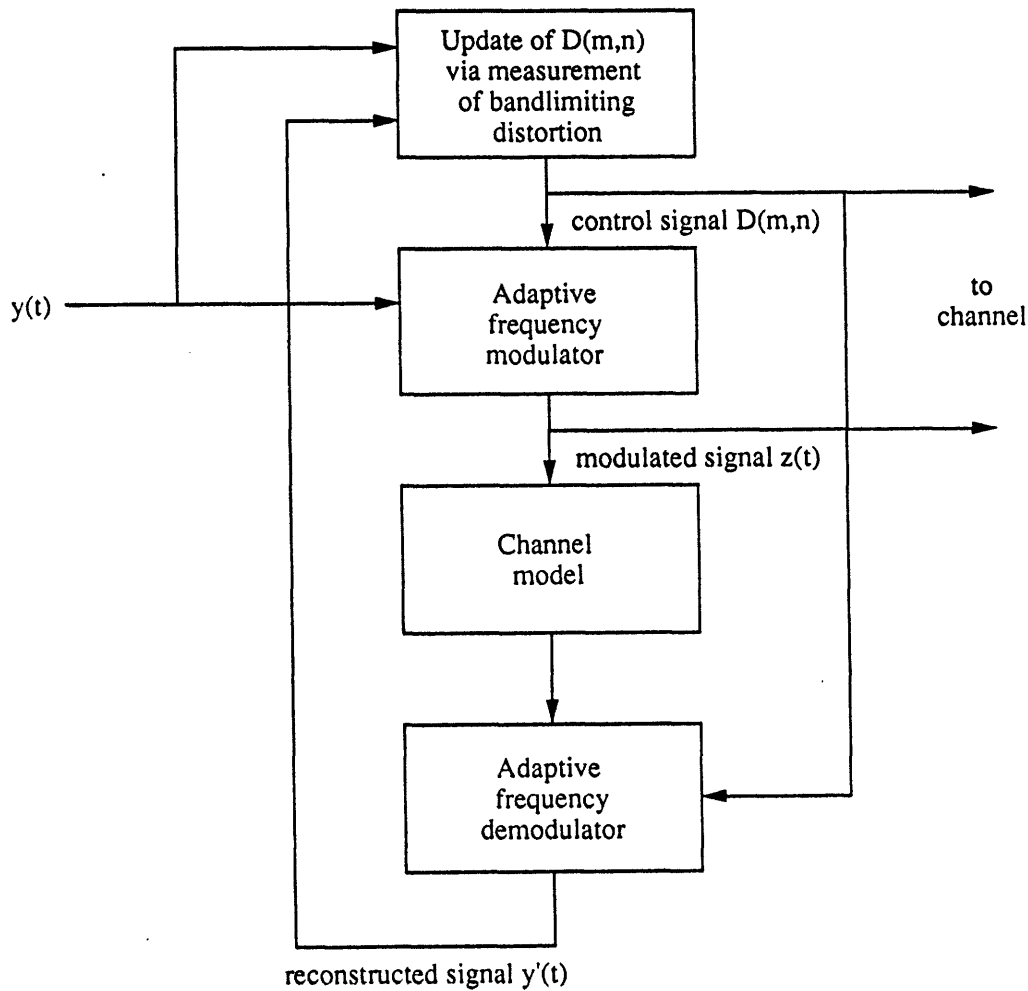
Figure 4.16: Iterative system for the update of the deviation function

Update of the Deviation

The flow chart that controls the update of the deviation in each block is shown in figure 4.17. First, an initial guess of the deviation function is performed. One of the two previously described algorithms can be used. A measure of the peak value of the bandlimiting distortion in each block follows. The peak value is used here because it better reflects the psychovisual importance of the distortion, and because it better indicates when overdeviation occurs.

93
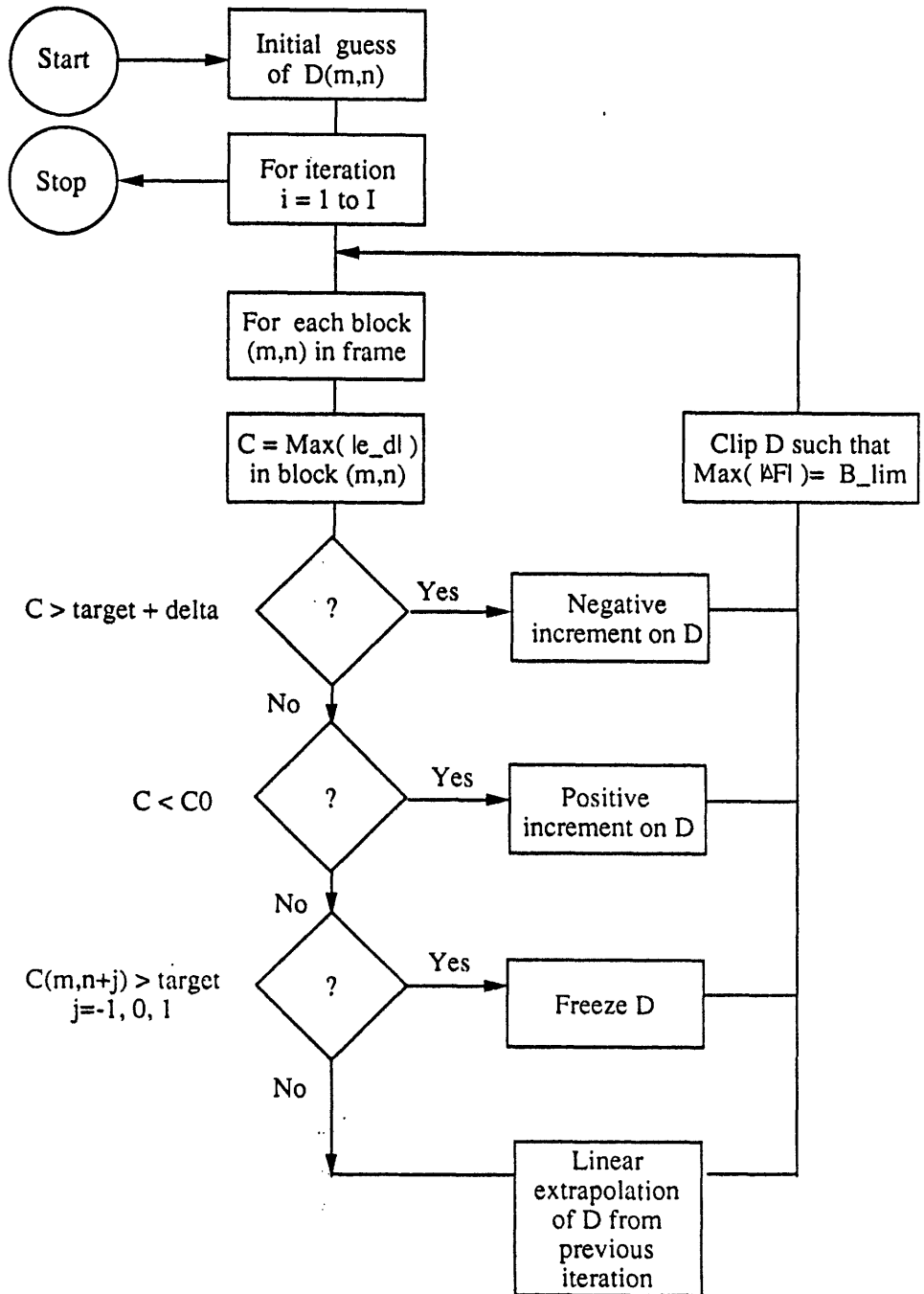
Figure 4.17: Flow-chart of update computation

Finally, computation of the deviation update takes place. For very small distortion, the deviation is incremented. For large distortion, a negative increment is applied. In other cases, the deviation is updated by linear extrapolation from the two previous iterations, until the distortion gets into the desired range. The desired distortion value is called the target and is a value small enough so that it is not visible and that the envelope doesn't drop down to the point where clicks occur in a typical noisy environment. A target value of 20, out of a luminance dynamic range of 256, was typically used in the simulation.

Linear extrapolation follows the quasi-Newton method [Str86]. In a first approximation, the distortion in block (m, n) depends on the deviation D in block (m, n) only. The variation, if small, of the deviation in neighboring blocks doesn't change much the distortion value in block (m, n). If T is the target distortion, D_0 the desired deviation, and C the distortion measured in block (m, n), the function $C(D)$ satisfies the following equation:

$$C(D_0) = T \tag{4.30}$$

The function $C(D)$ is nonlinear, and also varies for each block. This is shown in figure 4.18, where the distortion is displayed for many blocks, and the deviation is progressively increased. For some blocks, large distortion is measured as the activity in this block is large. In other blocks, there is very little distortion. A closed form of the function $C = C(D)$ doesn't exist, but equation (4.30) can be solved iteratively. Given measurements of $C_j = C(D_j)$ for $j = 0, 1, ..i$, the update D_{i+1} of the deviation solves the linearized equation:

$$C_i + \frac{1}{\mu}C_i'(D_{i+1} - D_i) = T \tag{4.31}$$

where μ is a factor controlling the rate of convergence and the stability of the method. A value of μ of 0.2 is typically used. When C_i is larger than T, but not too large, a larger value of μ is used to accelerate the decrease of deviation D. This favors smaller distortion values, in order to stay away from the threshold region, where the weak-noise assumption no longer holds.

MODULATION INDEX

BANDLIMITING DISTORTION

HORIZONTAL AXIS
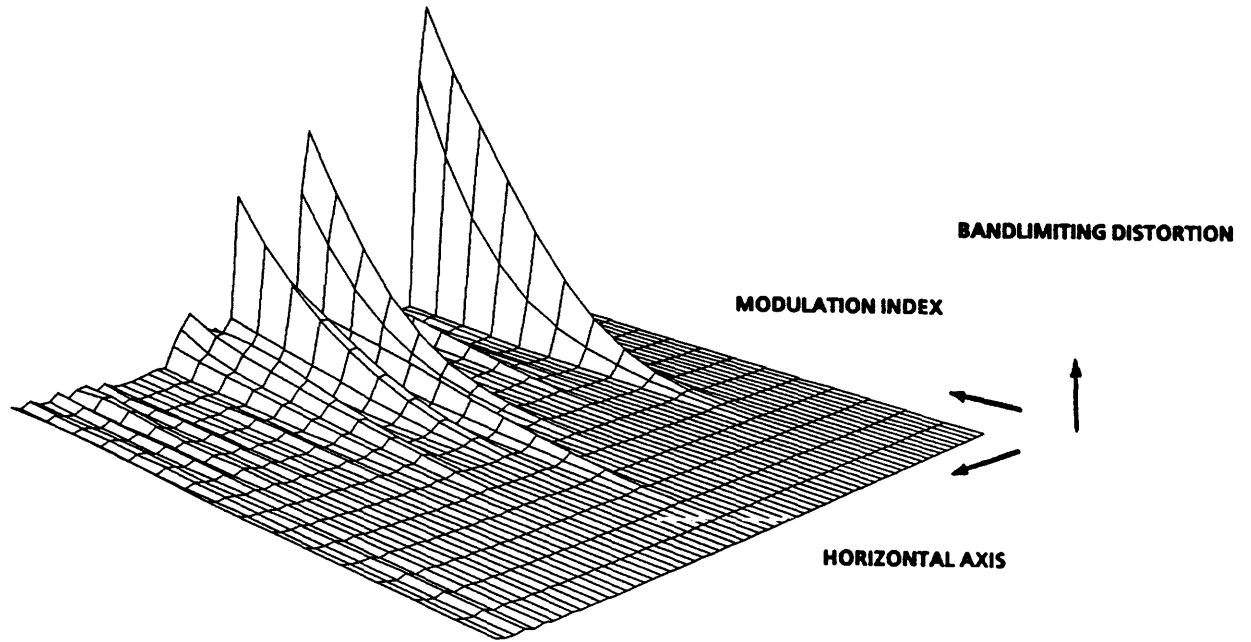
Figure 4.18: Bandlimiting distortion for many blocks as the deviation is increased.

The derivative C' of C with respect to D is approximated at the ith iteration by:

$$C_i' = \frac{C_i - C_{i-1}}{D_i - D_{i-1}} \qquad (4.32)$$

Low-Frequency Deemphasis

In order to limit the size of the Lows component, the bandwidth of the low-pass filter used to generate the Lows is very small. As a result, some relatively low frequencies are still present in the Highs component in the form of slowly changing flat field. These flat fields are not likely to produce significant distortion when spectrum truncation is applied. The associated modulated signal is a locally constant frequency sinewave, and the filtering involved in the channel introduces a constant phase shift, which does not produce distortion on the demodulated signal. However, when the peak effective frequency deviation is larger than the channel bandwidth B^+, the attenuation that takes place in the channel produces an envelope drop, which can be very detrimental when noise is also present. In order to prevent this, the deviation in each block

96

is clipped so that the peak effective frequency deviation never exceeds B_{lim} where $B_{\text{lim}} < B^+$ as shown in figure 4.17.

The limitation on the peak effective deviation is counterproductive, as the idea of the iterative scheme is to allow for even larger deviation when the distortion is small enough. One way to go around the problem is to allow for a deemphasis of the low frequencies before feeding the signal to the modulator. By imposing a smaller than unity gain G_0 at low frequencies on the deemphasis circuit, the peak effective frequency deviation that takes place in the flat fields is reduced by G_0. Hence, the deviation D is clipped so that the effective frequency deviation when no deemphasis is applied never exceeds $B_{\text{lim}} = B^+/G_0$. In that case, the effective peak deviation in flat fields, and when deemphasis is applied, is still below B^+, while allowing for larger deviation in regions where distortion is more likely to occur.

On the receiver side, a corresponding emphasis of the low frequencies is performed. As the noise spectrum density is very small at low frequencies, the increase of the noise mean square value is very small. In our simulation, the following emphasis circuit has been used with success:

$$H_{\text{emphasis}}(s) = G_0 \frac{1 + sT_1}{1 + sT_2} \qquad (4.33)$$

where $T_1 = 0.53\mu s$, $T_2 = 0.27\mu s$, and $G_0 = 0.7071$.

Performance of the Algorithm

The iterative algorithm converges very rapidly, typically requiring only two to three iterations. Measurement of the peak value of the bandlimiting distortion, rather than its RMS value, improved both speed and stability, when compared with a system proposed earlier [SP88].

Figure 4.19 shows the different iterations, together with the corresponding peak bandlimiting distortion in each block. The initial guess, based on the EPFD algorithm, produces distortion close to edges and on some background texture. After three iterations, the distortion is more uniform, showing the effectiveness of the method to increase the modulator gain until a level of distortion is attained.
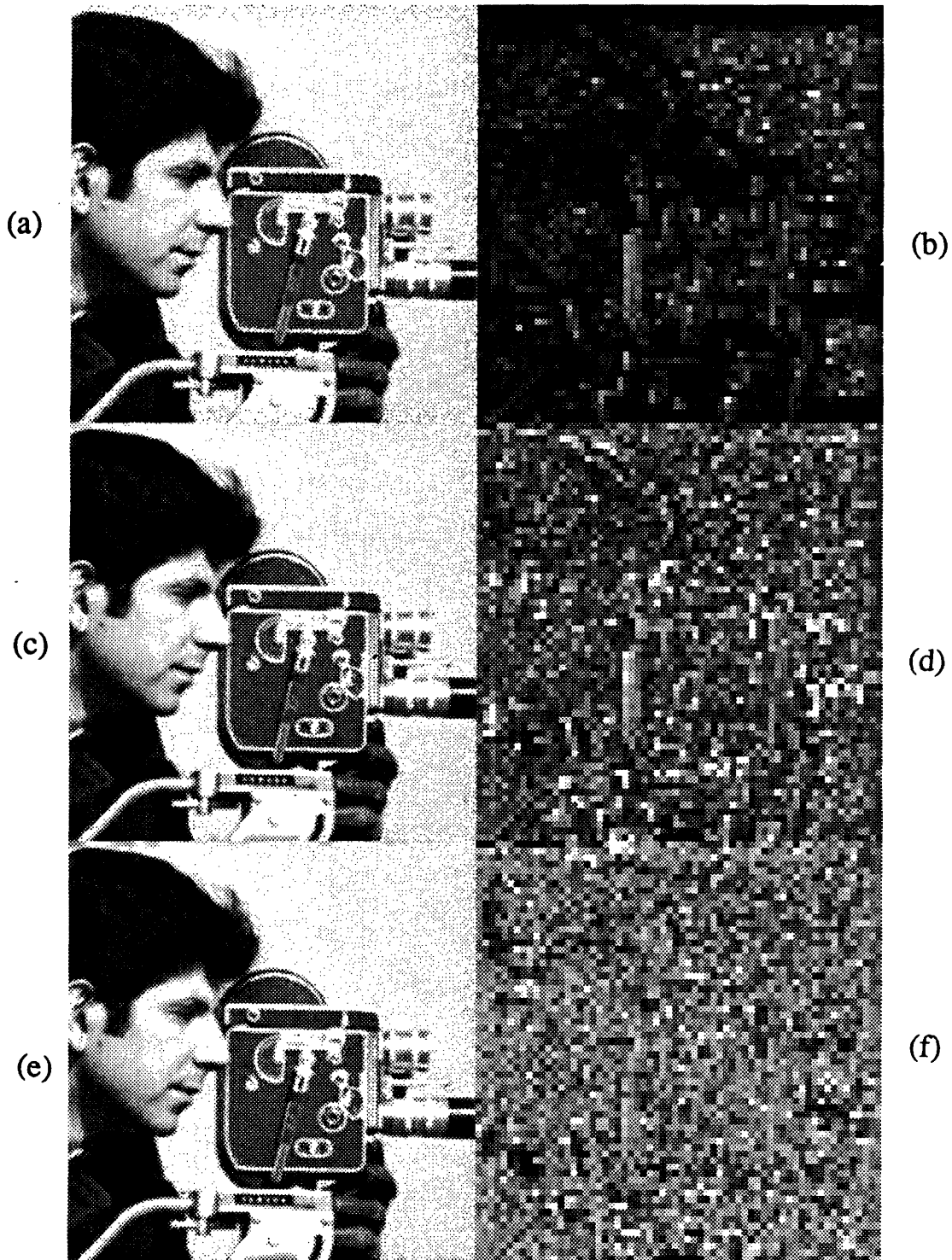
Figure 4.19: Illustration of iterative algorithm. (a) initial guess. (b) bandlimiting distortion of initial guess. (c) first iteration. (d) corresponding bandlimiting distortion. (e) third iteration. (f) corresponding distortion. The block size is 4 x 4. The channel is 11rd Butterworth of bandwidth $B^+ = F_s$, and the CNR is 9 dB.

Algorithm	SNR -masked noise- [dB]
PB	19.5
PEFD	21.7
Iterative -initial guess-	22.1
Iterative -first iteration-	26
Iterative -third iteration-	25.8

Table 4.2: Comparison of three algorithms.

Figure 4.20 shows the histogram of the distortion obtained after three iterations. The target distortion T was set to 20. The measured histogram shows a large peak at about 15, with a few blocks exceeding 20.
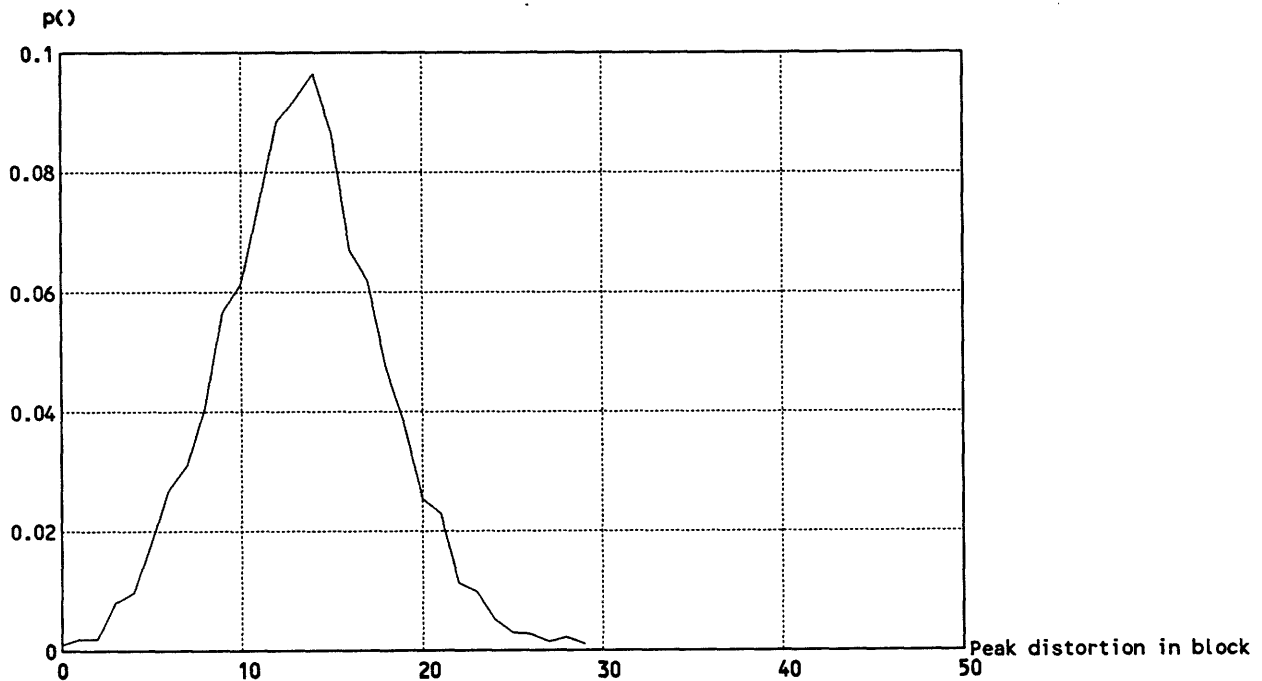
Figure 4.20: Histogram of peak bandlimiting distortion on third iteration.

Finally, the three algorithms are compared in table 4.2. The channel model is a 11th order Butterworth filter of bandwidth $B^+ = F_s$, and the CNR is 9 dB. It appears that the largest improvement of the iterative algorithm comes after the first

iteration. The following iterations improve the distribution of distortion across the picture, as shown in figure 4.19 but don't contribute much to a noise improvement. An improvement of about 4 dB is obtained with the iterative method, when compared with the two other methods.

4.3.3 Frame-Iterative Algorithm

The iterative algorithm can improve the noise performance of adaptive modulation, but the encoder is much more complex, as several iterations must be completed before the frame is finally transmitted. The complexity can be decreased, however, by iterating the deviation from frame to frame.

In this scheme, encoding of the present frame is made based on measured distortion from the past two frames. When there is no motion in the scene, the results are similar to those of the iterative algorithm. When motion is present, two artifacts are likely to appear. First, noise lags behind moving edges. The encoder doesn't increase the deviation at the prior location of the edge, and noise is not necessarily masked. Second, clicks, or clipping, occur on moving edges. The encoder doesn't foresee the presence of an edge, and increases the deviation. However, the deviation is too large for the edge, resulting in an envelope drop, clicks, or clipping. Of course, a motion-compensated algorithm would take care of these artifacts, but the additional complexity defeats the purpose of the frame-iterative algorithm.

It appears that for static regions, the frame-iterative algorithm performs better than than the EPFD algorithm. For moving regions, however, performance drops. This is investigated for a panning sequence at 4 pels/frame and 8 pels/frame, i.e. at medium to large motion. In figure 4.21, the original and the coded sequence using the PEFD algorithm are shown, while in figure 4.22, the two coded sequences using the frame-iterative algorithm are shown. The channel is the same as that used in figure 4.19. The artifacts associated with the frame-iterative algorithm are not negligible. It appears here that the PEFD algorithm gives more consistent results when various amounts of motion are expected.
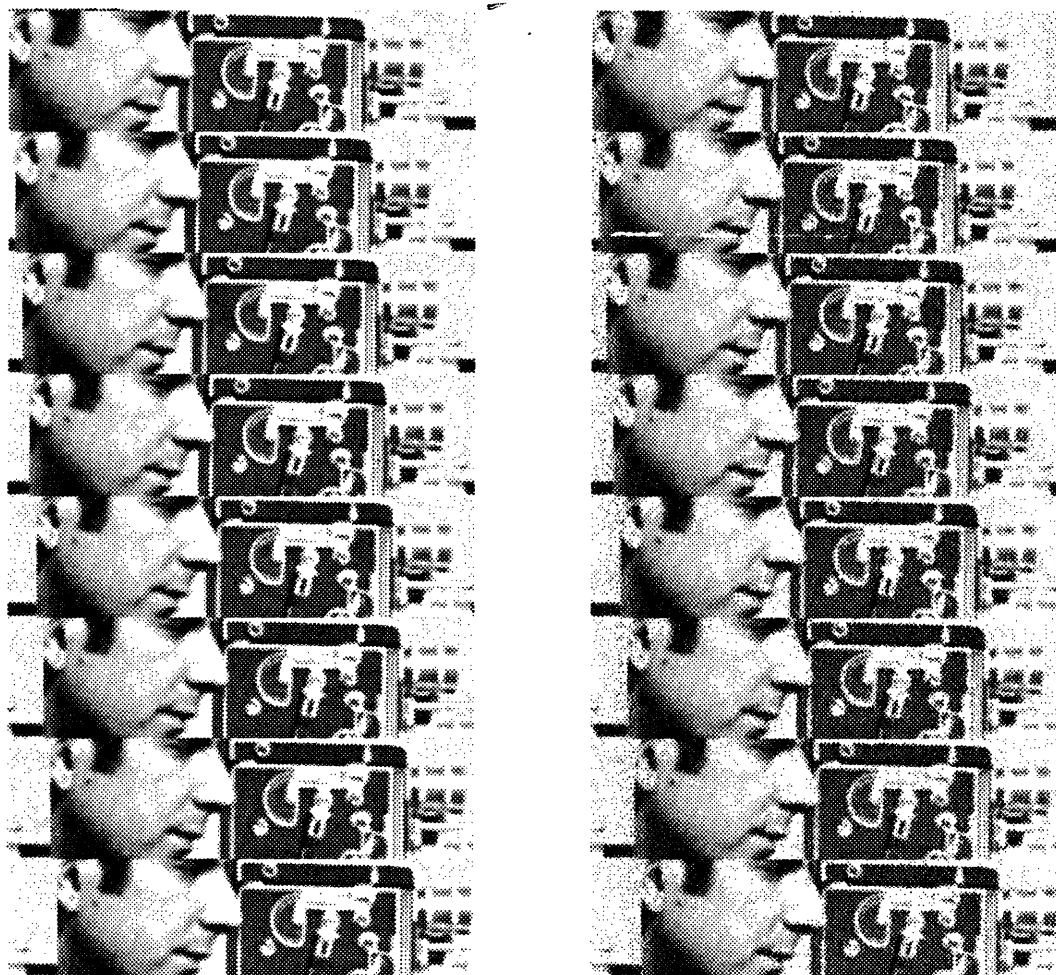
<center>(a)</center> <center>(b)</center>

Figure 4.21: Coded sequence with the PEFD algorithm. (a) original. (b) PEFD algorithm.
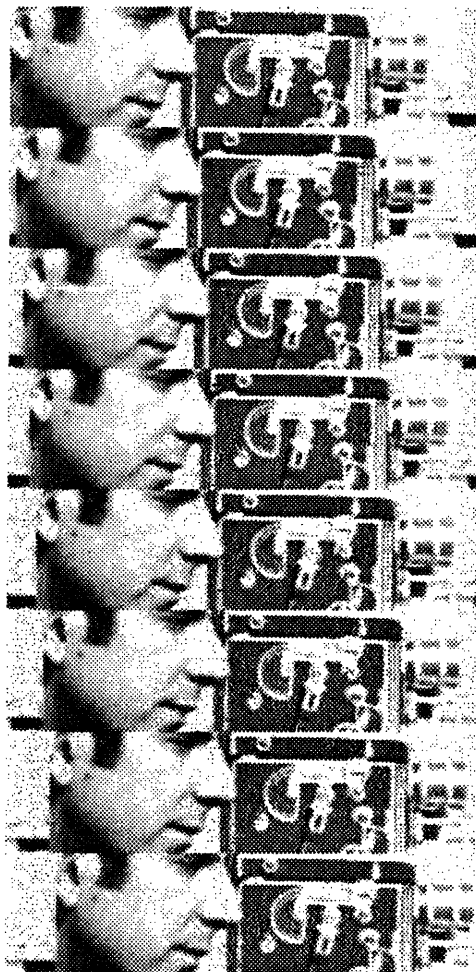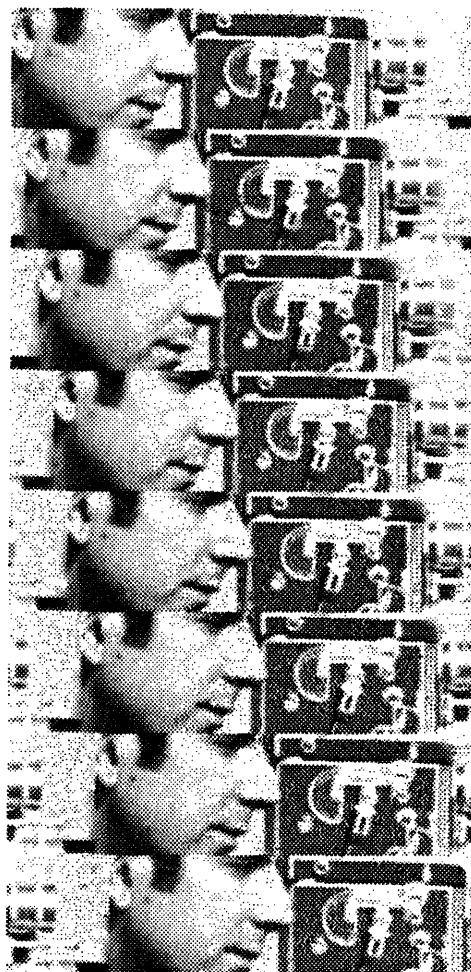
Figure 4.22: Coded sequence with the frame-iterative algorithm. (a) panning at 4 pels/frame. (b) panning at 8 pels/frame.

4.4 Summary

In this chapter, a more realistic model of a satellite transmission link is presented, and its effects on the demodulated signal are investigated. Three major impairments arise due to bandpass filtering. First, the demodulated signal is filtered by the channel equivalent low-pass system. This effect can be canceled in most instances. Second, some nonlinear distortion, called bandlimiting distortion, is generated, which depends on the frequency deviation and the channel characteristics. Third, the instantaneous envelope of the modulated signal drops, when much power is outside the channel pass band. The last two impairments impose a limitation on the deviation when consistent picture quality is required. Measurements show that if the peak modulated bandwidth is smaller than one third the channel bandwidth, the effect of the link can be ignored. Similarly, for a test signal, the same holds true when the peak-to-peak effective deviation is smaller than the channel bandwidth.

Three algorithms based on distortion limitations are then introduced. The effective peak frequency deviation algorithm (EPFD) maximizes the deviation while keeping the peak effective deviation in a block smaller than a tolerated value. The iterative and frame-iterative algorithms increase the deviation in each block until a certain level of distortion is reached in that block. A coding gain improvement of up to 6 dB is obtained when compared with the constant-bandwidth algorithm described in chapter 3.

Chapter 5

Multirate Frequency Modulation

In chapter 2, it was shown that the triangular noise spectrum very well matches the frequency response of an observer, which is essentially a low-pass response. However, the noise spectrum is uniform in the two other directions, namely vertical and temporal, and is more visible. The question addressed in this chapter is how to process the signal in different frequency bands differently so as to minimize the perceived transmission noise.

A popular technique that achieves this is preemphasis. It consists of preemphasizing high frequencies before transmission, and deemphasizing them after transmission so as to get the original back. In the process, the transmission noise is decreased [Mal86]. Another popular technique, known as subband technique, processes separately components in different frequency bands. The reader will find a good tutorial on the subject in [Vai87].

In this chapter, we review preemphasis as a technique to reduce the subjective effects of noise. Then, a technique to transmit subbands sequentially using frequency modulation is presented. It is found that substantial improvements are possible when the rates of transmission of the subbands are adjusted according to their visual importance.

This technique also addresses another problem that appears in adaptive frequency modulation systems, namely how to transmit different analog samples streams in a given channel, with a variable fidelity criterion for each stream. Frequency mul-

tiplexing is a potential candidate, but is disregarded, among other reasons due to intermodulation between subcarriers.

5.1 Optimal Preemphasis and Deemphasis

In this section, optimal preemphasis is reviewed and improvements are evaluated in the context of FM transmission of pictures.

Cramer derived an optimal solution to the problem of pre- and deemphasis of analog signals in noisy channels [B.G66]. The solution is optimal in that the weighted mean square error due to transmission noise is minimized subject to the constraint that the standard deviation of the signal with or without preemphasis is the same. These results are directly applicable to frequency modulation systems where a limitation of the modulated signal bandwidth is imposed. The bandwidth B_z is not altered by preemphasis in a first order approximation.

In the one-dimensional case, the optimal preemphasis filter $H_{opt}(F)$ has the following magnitude response:

$$\mid H_{opt}(F)\mid^2 = K_1 \left(\frac{S_{e_n}(F)\mid W(F)\mid^2}{S_y(F)}\right)^{1/2} \tag{5.1}$$

where $S_y(F)$ is the signal spectrum, $S_{e_n}(F)$ is the noise spectrum, and $W(F)$ is the noise weighting function. K_1 is a constant so that the signal power is unaltered by preemphasis:

$$K_1 = \frac{\int_{-F_s}^{+F_s} S_y(F)dF}{\int_{-F_s}^{+F_s}(S_y(F)\mid W(F)\mid^2 S_{e_n}(F))^{1/2}dF} \tag{5.2}$$

The matching deemphasis filter $G_{opt}(F)$ is the inverse of $H_{opt}(F)$:

$$\mid G_{opt}(F)\mid^2 = \frac{1}{\mid H_{opt}(F)\mid^2} \tag{5.3}$$

Finally, the coding gain obtained by preemphasis is defined as the ratio of unprocessed weighted noise to processed weighted noise:

$$G_{preemphasis} = \frac{\int_{-F_s}^{+F_s} S_{e_n}(F)\mid W(F)\mid^2 dF \ \int_{-F_s}^{+F_s} S_y(F)dF}{[\int_{-F_s}^{+F_s}(S_{e_n}(F)\mid W(F)\mid^2 S_y(F))^{1/2}dF]^2} \tag{5.4}$$

It reflects the weighted SNR improvement when preemphasis is applied.

106

5.1.1 Example

The coding gain $G_{\text{preemphasis}}$ is computed for a continuous Gauss-Markov process $y(t)$ and displayed in figure 5.1. The channel noise is triangular in (a) and white in (b). (a) shows the improvement when preemphasis is applied in FM transmission. (b) shows an extra improvement, available when vertical preemphasis is also applied.
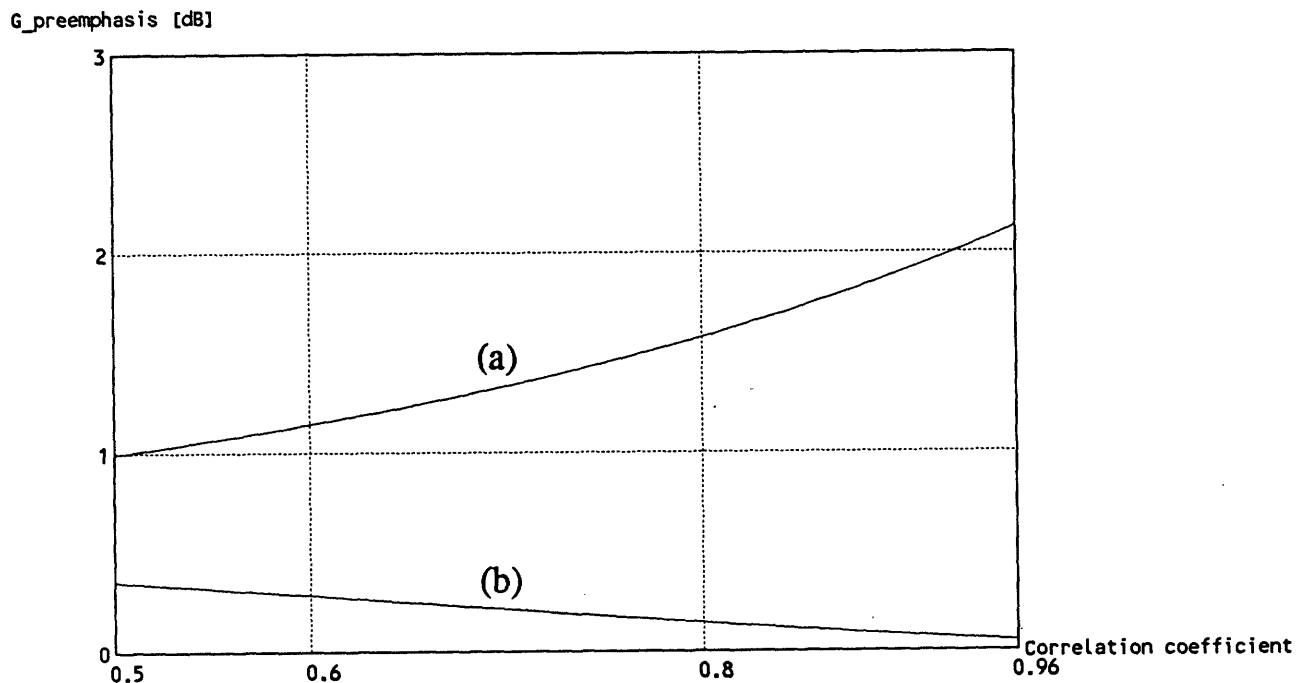
G_preemphasis [dB]

Figure 5.1: Coding gain $G_{\text{preemphasis}}$ for a Gauss-Markov process. (a) FM noise. (b) white noise.

In this example, the correlation coefficient ρ is defined as the autocorrelation function of $y(t)$ at $t = 1/2F_s$:

$$\rho = \exp(-F_0/2F_s) \tag{5.5}$$

where F_s is the maximum frequency of $y(t)$ to be transmitted. The use of the correlation coefficient permits comparison with other discrete coding strategies of processes with the same correlation.

107

The improvement $G_{\text{preemphasis}}$ is large when the mismatch between weighted noise and signal spectrum is large. In case of additive white noise, the low-pass signal spectrum is very similar to the weighted noise, and very little improvement is possible. For triangular noise, some spectral shaping is possible. However the improvement for a typical picture is smaller than 2 dB, showing the limitations of this technique[1].

5.2 Subband Decomposition

Subband coding is performed by splitting the signal $y(i, j, k)$ into L components $x_l(i, j, k)$, each of which corresponds to a given frequency band of the original signal. The components are coded independently. Either one-, two-, or three-dimensional frequency bands can be selected. A block diagram is shown in figure 5.2, where a three-dimensional decomposition is shown for the sake of generality.

The downsampling factor L is the product of the downsampling factors in each direction, for we shall confine ourselves to separable filters [Vet84].

$$L = L_i L_j L_k \tag{5.6}$$

perfect-reconstruction of the original is possible when no degradation occurs during transmission [Vet87].

As we shall be concerned with the noisy transmission of components, the relation between the error signal of one component and the error signal of the reconstructed sequence is described.

The total reconstruction error is called $e_r(i, j, k)$.

$$e_r(i, j, k) = y(i, j, k) - y'(i, j, k) \tag{5.7}$$

Let's define $e_l(i, j, k)$ the error due to transmission of component l and $e_{rl}(i, j, k)$ the error of the reconstructed sequence due to noisy transmission of component l alone.

$$e_l(i, j, k) = x_l(i, j, k) - x'_l(i, j, k) \tag{5.8}$$

---

[1]It should be noted here that the main reason for using preemphasis in NTSC satellite transmission is to decrease the effect of differential gain [Rho85].
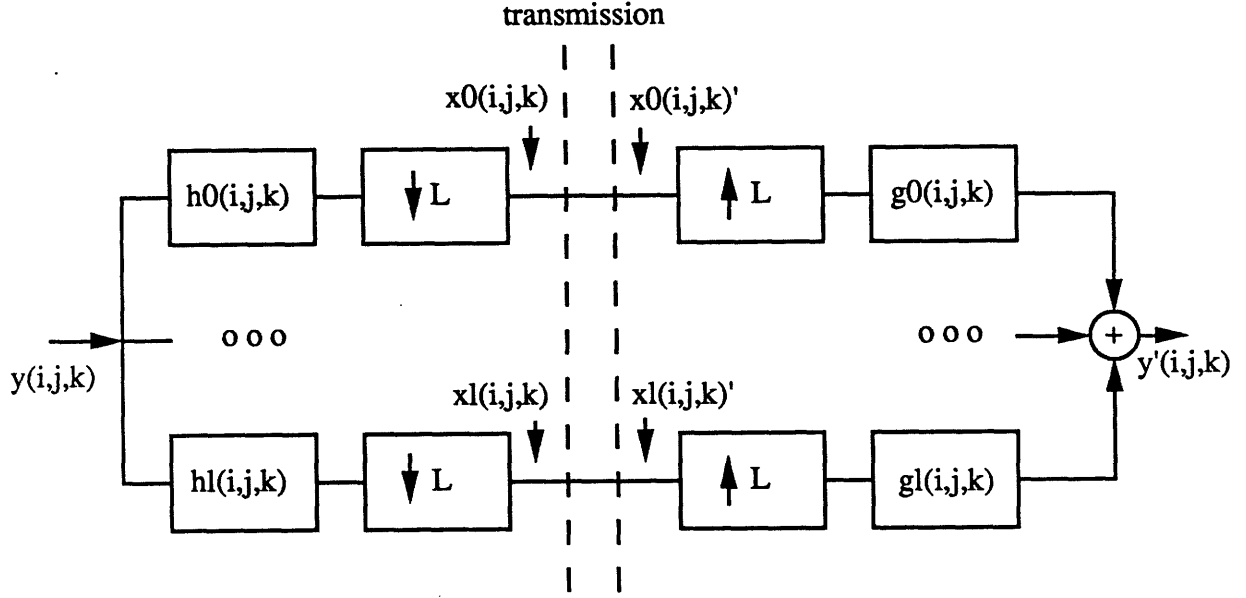
Figure 5.2: Decomposition of the sequence into components

In the frequency domain, we have:

$$S_{e_{rl}}(f_i, f_j, f_k) = S_{e_l}(f_i L_i, f_j L_j, f_k L_k) \mid G_l(f_i, f_j, f_k) \mid^2 \qquad (5.9)$$

Using the perfect-reconstruction assumption and the fact that the component error signals are not correlated, the total reconstruction error is in the mean square:

$$E(e_r^2(i, j, k)) = \sum_{l=0}^{L-1} E(e_{rl}^2(i, j, k)) \qquad (5.10)$$

where

$$E(e_{rl}^2(i, j, k)) = \int_0^1 \int_0^1 \int_0^1 S_{e_{rl}}(f_i, f_j, f_k) df_i df_j df_k \qquad (5.11)$$

The correlation between the component error signal is about zero when the frequency response of the synthesis filters have very little overlap, a condition typically met with sufficiently long filters.

Similarly, the total weighted reconstruction error $e_{rw}(i, j, k)$ is in the mean square:

$$E(e_{rw}^2(i, j, k)) = \sum_{l=0}^{L-1} \zeta_l \, E(e_{rl}^2(i, j, k)) \qquad (5.12)$$

where

$$\zeta_l = \frac{1}{E(e_{rl}^2(i, j, k))} \int_0^1 \int_0^1 \int_0^1 S_{e_{rl}}(f_i, f_j, f_k) \mid W(f_i, f_j, f_k) \mid^2 df_i df_j df_k \qquad (5.13)$$

109

The weighting factor ζ_l is a measure of the relative noise visibility in component l.

An important property of subband coding is that it allows an appropriate shaping of the transmission noise. By allowing more noise on components which have a small weighting factor ζ_l, we can get a good balance of noise in the different bands. Of course, this feature is used efficiently only for a transmission scheme with a noise-bandwidth exchange capability, such as FM or digital coding.

5.2.1 Frequency Modulation of the Subbands

In this section, a sequential transmission of the subbands using frequency modulation is described. The frequency deviation for each component is computed and the coding gain over regular frequency modulation is evaluated. The deviation D is shown to be larger for components with low power.

We shall use again a stochastic description of the components. The component $x_0(i, j, k)$ has mean m_0, whereas all others have a mean equal to zero. The block model of chapter 3 is used again as it allows us to keep track of the short-time bandwidth of the modulated signal. In this model, the signal in each small block is modeled by a multivariate Gaussian process, whose standard deviation obeys a probabilistic law.

Let's define σ_{ml} the maximum standard deviation in a block of component l, such that a block with smaller standard deviation appears with probability 0.99. The maximum standard deviation σ_{ml} is an important parameter of the process as the peak short-time bandwidth is linearly related to it, as a first approximation. Let's also call x_0 a reference value. D_l is the frequency deviation of band l, when the input is the reference value x_0.

The channel used for transmission of the frequency modulated signal $z(t)$ is bandwidth limited. We impose the short-time bandwidth b_z of $z(t)$ to be limited to B.

Hence the frequency deviation of band l is:

$$D_l = \frac{x_0}{4\sigma_{ml}} B \tag{5.14}$$

For convenience, let's call σ_{nl}^2 the mean square of the component error.

$$\sigma_{nl}^2 \equiv E(e_l^2(i, j, k)) \tag{5.15}$$

110

Using equation (3.16) we define SNR_0, the signal-to-noise ratio of any component:

$$SNR_0 = \frac{\sigma_{ml}^2}{\sigma_{nl}^2} = \frac{6}{16\Delta_s^2\Delta_n^2} k^2 CNR_0 \qquad (5.16)$$

Hence, the signal-to-noise ratio is the same for all components.

The mean square noise of component l is:

$$\sigma_{nl}^2 = \frac{\sigma_{ml}^2}{SNR_0} \qquad (5.17)$$

and the total weighted reconstruction error is:

$$E(\epsilon_{rw}^2(i,j,k)) = \frac{1}{SNR_0} \sum_{l=0}^{L-1} \zeta_l \sigma_{ml}^2 \qquad (5.18)$$

A sequential transmission of the subbands leads to the following coding gain over regular FM with same peak bandwidth:

$$G_{\text{subband}} = \frac{\zeta_{w_{tot}} \sigma_y^2}{\sum_{l=0}^{L-1} \zeta_l \sigma_{ml}^2}, \qquad (5.19)$$

where $\zeta_{w_{tot}}$ is the total weighting factor when no subband decomposition is performed, and σ_y^2 is the variance of the signal $y(i,j,k)$.

5.2.2   Example

In order to demonstrate the performance of the subband decomposition in a FM channel, the coding gain G_{subband} is evaluated for a two dimensional separable Gauss-Markov field.

We shall assume the analysis and synthesis filters to be ideal bandpass, in particular when computing the weighting factors and component variance. In practice, there is a non-zero transition band between stop band and pass band. With long filters, it is sufficiently small so as not to change the results significantly.

The standard deviation in each band is evaluated numerically assuming an ideal analysis filter:

$$\sigma_{ml}^2 = \int\int_{(f_i,\, f_j \text{ in band } l)} \frac{\sigma_y^2(1-\rho_i^2)(1-\rho_j^2)}{(1+\rho_i^2-2\rho_i\cos(2\pi f_i))(1+\rho_j^2-2\rho_j\cos(2\pi f_j))} df_i df_j$$

$$(5.20)$$

111

Similarly, ζ_l is evaluated using (5.13).

Figure 5.3 shows the coding gain when horizontal and vertical decomposition are used with $L_i = L_j = 5$. The coding gain expressed in dB is negative for positive correlation coefficient. For a low-pass spectrum, the performance of the system is actually decreased when using subband decomposition. For a high-pass spectrum, i.e. for negative correlation coefficient, this gain is positive. In this case, the component error for the low bands is decreased as energy in these bands is smaller. Subjectively, this improvement is much more effective and yields an overall improvement of subjective noise performance.
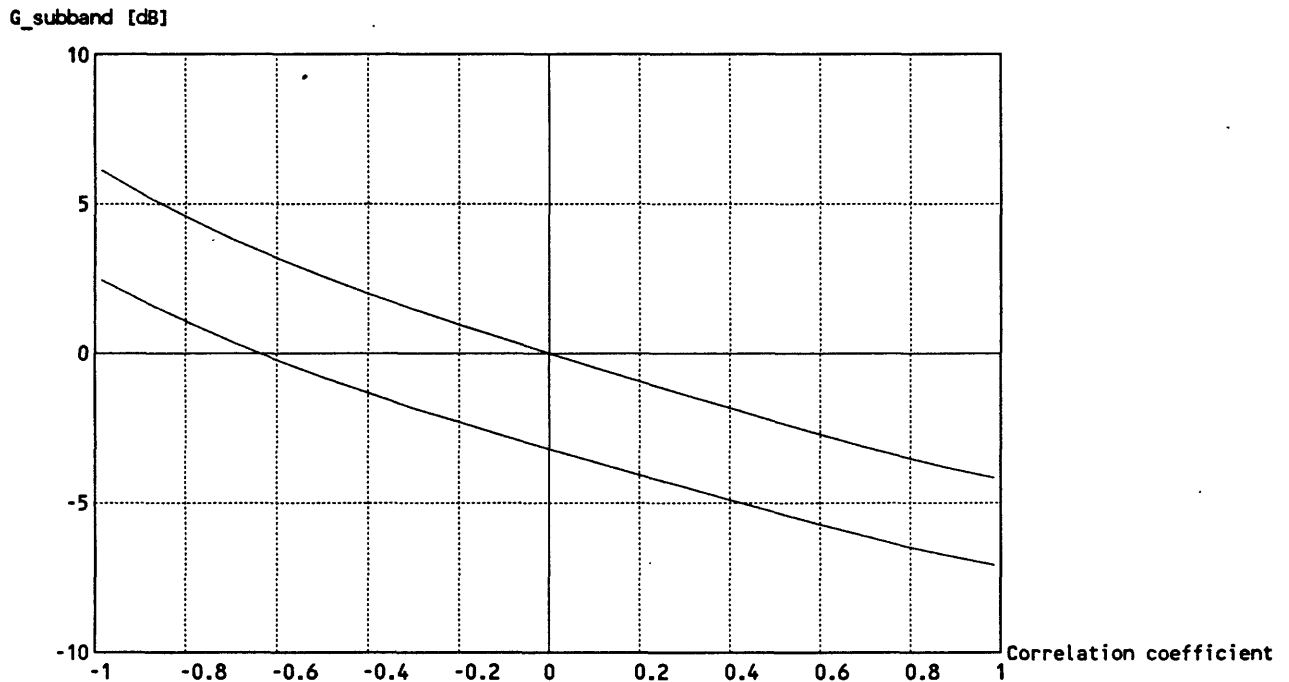
G_subband [dB]

Figure 5.3: Coding gain $G_{subband}$ for a 2-D separable Gauss-Markov field. (Bottom) Horizontal decomposition. (Top) Vertical decomposition. A five-band decomposition is performed.

5.3 Multirate Frequency Modulation

In the previous section, the performance of subband decomposition in FM transmission was shown to be effective only for a high-pass spectrum. Of course, television pictures are of low-pass nature and don't benefit from this decomposition. In this section, a sequential FM transmission of the components is demonstrated where the rate of transmission of each component is variable, yielding improved performance. The system is shown in figure 5.4. As explained in section 3.2, the components are still continuous in amplitude.
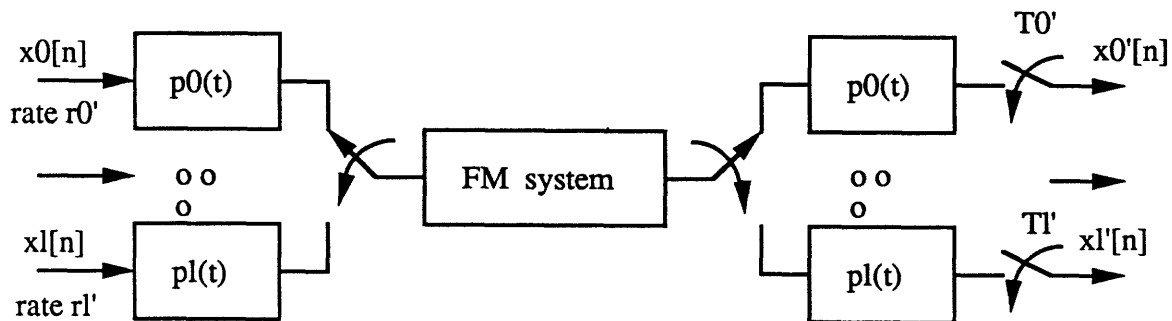
Figure 5.4: Sequential transmission of the components with variable rate.

Optimal rate allocation is achieved when the mean square value of the weighted reconstruction error is minimized. The problem of multirate frequency modulation can be stated in the following way:

- Derive r'_i, the rate of transmission of component i, and D_i, the frequency deviation for component i, so as to minimize $E(e^2_{rw}(i, j, k))$ under the constraint $b_z < B$.

This problem is very similar to the bit allocation problem in transform coding [JN84]. By varying the rate of transmission of each component, the noise character-

113

istics of the reconstructed picture can be tuned to match the human sensitivity to noise.

5.3.1 Derivation of the Optimal Rates

Let's define the average rate of the source:

$$r_{av} = 2F_s \qquad (5.21)$$

r_{av} is the rate necessary to transmit the discrete sequence when no component separation is performed. It is also the rate of the components when transmitted at equal rate in a time-multiplex scheme.

The average bandwidth expansion factor k_{av} is :

$$k_{av} = \frac{B}{r_{av}/2} \qquad (5.22)$$

The rate of each component l is:

$$r_l = \frac{r_{av}}{L} \; ; l = 0, .., L - 1 \qquad (5.23)$$

Let's also define the time-expansion factor of component l:

$$\gamma_l = \frac{r_{av}}{r'_l} \; ; l = 0, .., L - 1 \qquad (5.24)$$

where r'_l is the rate after time-expansion. This factor expresses the relative slow-down factor used to transmit this component.

After time-expansion γ_l, the noise mean square value is:

$$\sigma^2_{nl} = \frac{1}{\text{SNR}_0} \frac{\sigma^2_{ml}}{\gamma_l^3} \qquad (5.25)$$

where SNR_0 was defined in (5.16). It can be seen that the noise is decreased with the cube of the time expansion factor. This relation comes from equation (2.41), where the ratio of the SNR to CNR_0 was found to be proportional with the square of the bandwidth expansion factor, which is directly related with γ_l. Additionally, CNR_0 is increased by another term γ_l since it is defined for the noise integrated over the basebandwidth, and that the basebandwidth is also decreased by γ_l.

114

Additionally, the set of time-expansion factors have to satisfy the following constraint in order to make possible time-multiplexing :

$$\sum_{l=0}^{L-1} \frac{r_l}{r_{av}} \gamma_l \leq 1 \tag{5.26}$$

which is equivalent to:

$$\sum_{l=0}^{L-1} \gamma_l \leq L \tag{5.27}$$

The total weighted reconstruction error is:

$$E(e_{rw}^2(i,j,k)) = \frac{1}{\text{SNR}_0} \sum_{l=0}^{L-1} \frac{\zeta_l \sigma_{ml}^2}{\gamma_l^3} \tag{5.28}$$

where ζ_l is the weighting factor for component l, as defined in (5.13).

Using a Lagrange multiplier, the minimum of (5.28) under the constraint (5.27) is obtained with:

$$\gamma_l = \frac{(\zeta_l \sigma_{ml}^2)^{1/4}}{\frac{1}{L} \sum_{l=0}^{L-1} (\zeta_l \sigma_{ml}^2)^{1/4}} \tag{5.29}$$

Equation (5.28) reduces to:

$$E(e_{rw}^2(i,j,k)) = \frac{1}{\text{SNR}_0} (\frac{1}{L} \sum_{l=0}^{L-1} (\zeta_l \sigma_{ml}^2 L)^{1/4})^4 \tag{5.30}$$

Hence, the coding gain of multirate FM over regular FM is:

$$G_{\text{multirate}} = \frac{\zeta_{\mathbf{w}_{tot}} \sigma_y^2}{(\frac{1}{L} \sum_{l=0}^{L-1} (\zeta_l \sigma_{ml}^2 L)^{1/4})^4} \tag{5.31}$$

The weighting factor $\zeta_{\mathbf{w}_{tot}}$ was defined in (2.52).

5.3.2   Example

The coding gain of multi-rate frequency modulation is computed for a separable Gauss-Markov field. The performance for vertical and horizontal decomposition is shown in figure 5.5 with $L_i = L_j = 5$. A substantial improvement of performance can be observed when compared to figure 5.3. In particular, vertical decomposition is very effective, with a gain ranging from about 1 to 10 dB.

Horizontal decomposition is less effective. For a uniform signal spectrum, in which case the correlation coefficient is zero, the coding gain expressed in dB is negative, a

115

result not surprising when it is remembered that frequency modulation is horizontally very well matched to the human observer sensitivity. Variable rate allocation can however improve this when either a high-pass or a low-pass spectrum is transmitted.

Two-dimensional decomposition yields a gain which is, in dB, the sum of the horizontal and vertical decomposition gain. For large correlation coefficients, it can be advantageous over vertical decomposition alone.
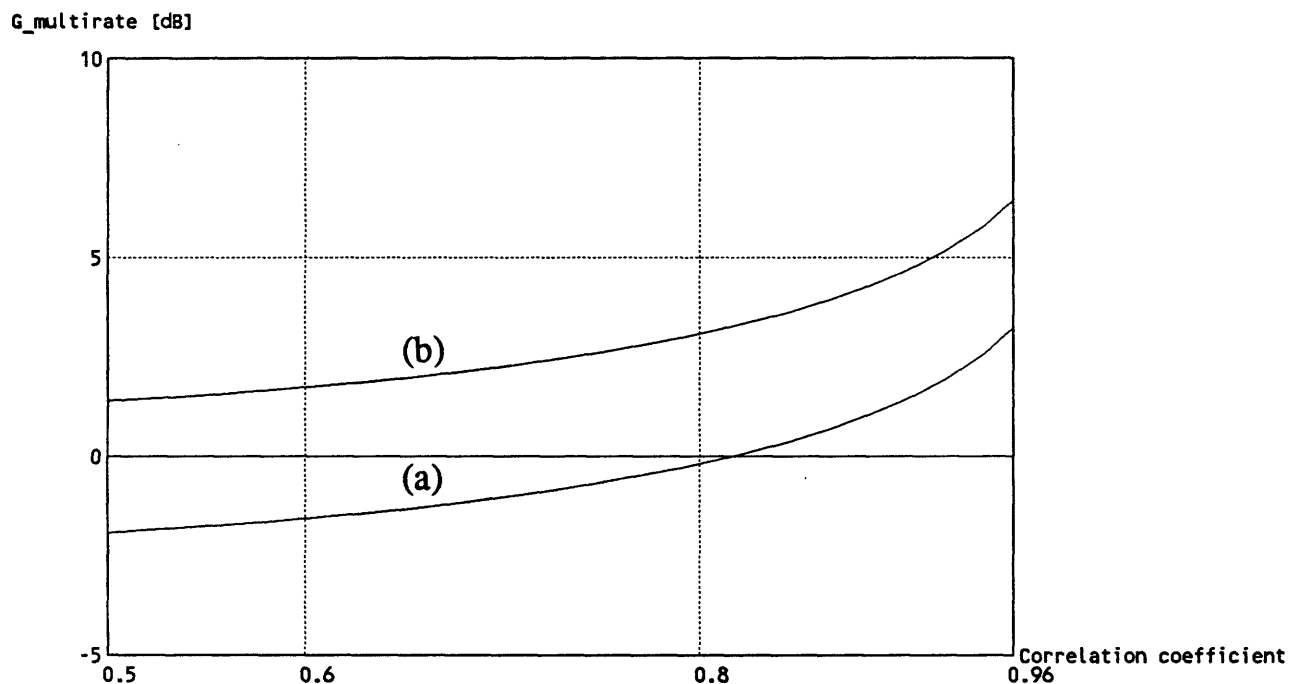
G_multirate [dB]

Figure 5.5: Coding gain $G_{multirate}$ for a 2-D separable Gauss-Markov field. (a) Horizontal decomposition. (b) Vertical decomposition. A five-band decomposition is performed.

5.3.3 Another Example

The prior example covered the case of a stationary model of pictures. Here, a vertical subband decomposition of "Girl" is performed, and the performance of multirate frequency modulation is evaluated.

Band l	σ_{ml}	ζ_l/ζ_{wi}	γ_l
$y[i]$	121	0.24	1
0	119	0.630	2.3
1	37	0.278	1.0
2	24	0.148	0.7
3	16	0.097	0.5
4	12	0.056	0.4

Table 5.1: Multirate system for "Girl", vertical decomposition.

In this example, analysis and synthesis are performed using five-band filters designed by Vaidyanathan et al. [DV88]. The five components are shown in figure 5.6, where it is apparent that higher bands have a smaller variance. The nonstationary nature of each component is very visible.

Based on the block model, a probabilistic description of the standard deviation in a block is derived. Finally, the time-expansion factors are computed. The system parameters are shown on table 5.1. The horizontal weighting factor ζ_{wi} was defined in (2.52).

Figure 5.7 shows a part of the reconstructed picture after transmission with $B/F_s = 1.5$ and $CNR_0 = 16$ dB. It should be compared with figure 5.8 where regular FM was used with the same parameters. The noise structure in figure 5.7 is more pleasing, like the high frequency grains of photographic prints. It is also less objectionable when viewed from the regular viewing distance. The noise improvement is due both to a better usage of a priori information on the source -the spectrum is not uniform- and a better match to the human visual response.

The coding gain $G_{\mathrm{multirate}}$ is in this case 3.4 dB. This value is smaller than that predicted for a Gauss-Markov process, since the peak standard deviation in a block is taken rather than the standard deviation.
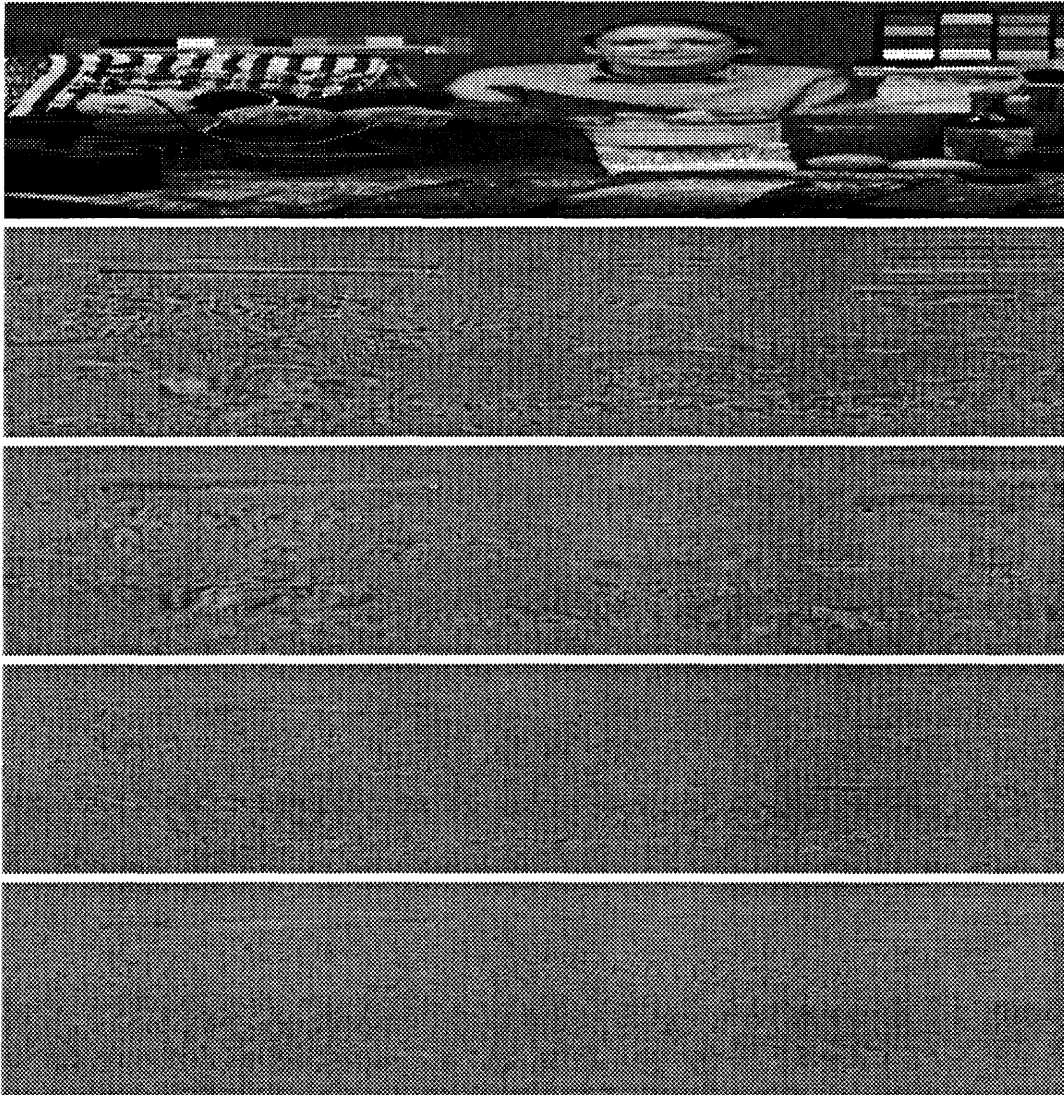
Figure 5.6: Vertical decomposition of the picture "Girl"

Figure 5.7: Multirate FM coding of a segment (256 x 256) of "Girl".

Figure 5.8: Regular FM coding of a segment (256 x 256) of "Girl".

5.3.4 Derivation of the Optimal Rate with an Additional Constraint

As shown in chapter 4, one of the effects of the transmission channel on the demodulated signal is linear filtering by an equivalent low-pass filter. On the other hand, transmission free of intersymbol interference -a necessary condition for perfect reconstruction- can take place only if the bandwidth of $H^+(F)$ is larger than or equal to half the transmission rate. It appears that the rate of a component after time-expansion must be upper bounded by $2B^+$.

$$r'_l = \frac{r_{av}}{\gamma_l} \leq 2B^+ = \frac{r_{av}}{\gamma_{min}} \tag{5.32}$$

or, equivalently:

$$\gamma_l > \gamma_{min} = \frac{r_{av}}{2B^+} \tag{5.33}$$

A possible procedure to handle this additional constraint follows. It defines an approximate solution γ'_l that satisfies the constraint 5.33.

- solve γ_l; $l = 0, ..., L - 1$ using (5.29)

- replace γ_j by $\gamma'_j = \gamma_{min}$ for all j such that $\gamma_j < \gamma_{min}$

- find L' = number of subbands with $\gamma_l \geq \gamma_{min}$

- solve

$$\gamma'_l = \frac{(\zeta_l \sigma^2_{ml})^{1/4}}{\frac{1}{L'} \sum_{l=0}^{L'-1} (\zeta_l \sigma^2_{ml})^{1/4}} \tag{5.34}$$

for each l such that $\gamma_l > \gamma_{min}$

This procedure limits the time-expansion that is applied on components with very low fidelity requirements. Such components don't need a large bandwidth expansion factor to decrease the channel noise, but require a bandwidth at least equal to half the transmission rate in order to ensure no intersymbol interference.

121

5.4 Adaptive Multirate FM

The subbands described in the previous section present space-varying statistics. Adaptation of the modulator to the signal activity in each band is presented in this section.

Here the subband decomposition is performed on the Highs only. The Lows are transmitted by another mean, for example by frequency modulation with a large time-expansion factor, so as to transmit them with little degradation.

In order to guarantee that the short-time bandwidth doesn't exceed a certain value B, the frequency deviation D_l is set according to the maximal standard deviation σ_{ml} in all blocks. The deviation in block (m, n) can be raised by the adaptation factor f_a without increasing the peak bandwidth, or the bandlimiting distortion.

$$D(m,n) = D_0 f_a(m,n), \tag{5.35}$$

where $f_a(m,n)$ is found using some algorithm described in chapter 3 or 4.

In adaptive multirate FM, the rate of transmission of each component is adjusted according to its psychovisual importance and adaptive modulation is applied to the L subbands of the Highs.

The total weighted reconstruction error $e_{rw}(i,j,k)$ after transmission is in the mean square:

$$E(e_{rw}^2(i,j,k)) = \sum_{l=0}^{L} \gamma_l E(e_{rl}^2(i,j,k)) \tag{5.36}$$

where γ_l is the weighting factor measuring both the frequency weighting and the improvement due to adaptive modulation.

Unlike in chapter 3, it is hard to evaluate the weighting factor γ_l. When subband decomposition is performed, there is no simple way to combine the masking factor with the noise improvement obtained with adaptive modulation. However, the average improvement due to adaptive modulation G_{adap-m} can be arbitrarily set, for example, to 10, a conservative value. This worst-case approach provides us with a design guideline that can be useful.

Hence, the weighting factor for the Highs components is approximated by:

$$\gamma_l = \frac{1}{G_{\text{adap-m}} \, E(e_{rw}^2(i,j,k))} \int_0^1 \int_0^1 \int_0^1 S_{e_{rl}}(f_i, f_j, f_k) \mid W(f_i, f_j, f_k) \mid^2 df_i df_j df_k$$

(5.37)

with

$$G_{\text{adap-m}} = 10$$

(5.38)

5.4.1 Subband Grouping for Adaptive Modulation

Adaptive frequency modulation requires the transmission of the factors $f_a(m,n)$, or, equivalently, the estimated block standard deviation, or the peak absolute value in a block. As a result, the subjective effect of noise is greatly reduced. However, this effect relies heavily on the masking effect and the block size cannot be increased much above 4 x 4, for the noise would spread too far from the region where masking is effective.

Adaptive modulation of the subbands is also possible. However, because the subbands are downsampled, the size of the adaptation block has to be reduced accordingly, typically to a size of 1 x 1 or 2 x 2, in order not to exceed much an effective size of 4 x 4 after upsampling. Hence, if all subbands are adaptively modulated, the side information is increased considerably, more precisely by a factor L. A suboptimal approach is proposed here to reduce this side information.

Typically only a few subbands display significant activity in any small area of the picture. For example, this region can be an edge in the original picture and depending on the edge angle, certain frequency bands will present locally a large activity. Another example is a texture with periodic components. By grouping the subbands and assigning a single factor for the group, little loss of performance occurs. The number of groups G will determine the amount of side information. Here only a small number of groups -typically one to four- is considered in order to minimize the side information.

Let's define $f_l(m,n)$ the adaptation factor in block (m,n) computed for the component l.

123

The shared adaptation factor $f'_g(m, n)$ in block (m, n) for all bands belonging to a same group g is used to control the frequency modulator:

$$f'_g(m, n) = \text{Min}(f_l(m, n))_{(l \text{ in group } g)} \tag{5.39}$$

For some subbands, this factor is suboptimal, as larger factors could be used.

For a two-dimensional decomposition, experience shows that there is no need to use more than four different groups of factors. For example, a division of 16 subbands into four groups is proposed in figure 5.9. As edges are the elements in the picture most likely to produce a large local variance, this classification amounts to separating horizontal edges (group 0), vertical edges (group 1), diagonal edges (group 2), and the low frequencies (group 3), and assigning to each group a different set of adaptation factors.
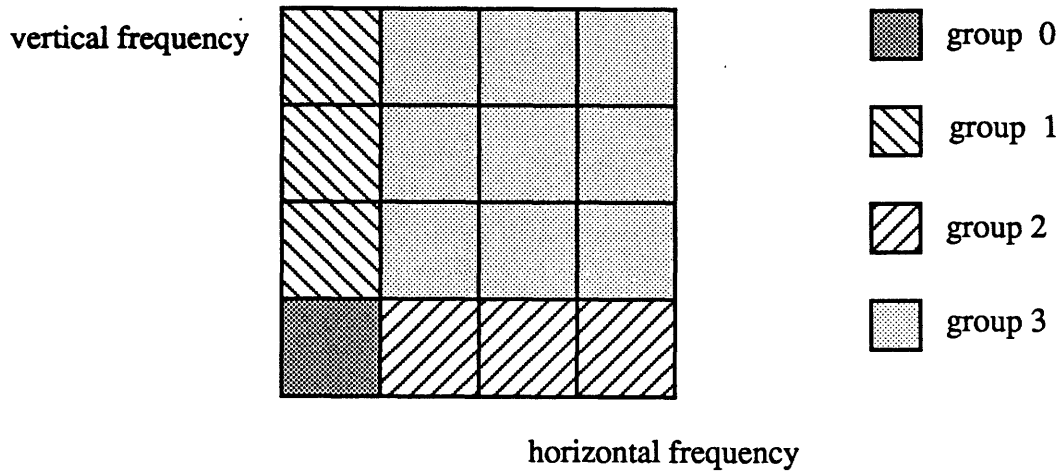
Figure 5.9: Groups of subbands with common frequency deviation

5.4.2 Derivation of the Optimal Rates

Adaptive modulation requires the transmission of side information. In this section, the derivation of the optimal rates is performed for the case where the Lows transmis-

sion is included in the design. A similar procedure can be used when the adaptation factors, or the chrominance factors are also included. It is illustrated in chapter 6.

The Lows is a critical component as no adaptive modulation can be applied, which results in no adaptation coding gain. This has to be compensated by a large expansion factor that will significantly raise its SNR, but will also use a large amount of the channel capacity. The Lows are averaged over large blocks so as to keep the data rate very low. Typical block sizes range from 4 x 4 to 8 x 8, and even 16 x 16, with a somewhat reduced performance. A value of 8 x 8 was found to be a good trade-off.

The average rate of the source is:

$$r_{av} = 2F_s(1 + \frac{1}{M_l N_l}) \tag{5.40}$$

where M_l, N_l are the downsampling factors of the Lows.

The $L + 1$ components –one component for the Lows and L subbands for the Highs– are time-expanded with factor γ_l.

$$\gamma_l = \frac{r_{av}}{r_l'} \; ; l = 0, .., L \tag{5.41}$$

where r_l' is the rate after time-expansion.

The time-expansion factors have to satisfy the constraint:

$$\sum_{l=0}^{L} \frac{r_l}{r_{av}} \gamma_l \leq 1 \tag{5.42}$$

The total weighted reconstruction error is the function to be minimized:

$$E(e_{rw}^2(i,j,k)) = \frac{1}{\text{SNR}_0} \sum_{l=0}^{L} \frac{\zeta_l \sigma_{ml}^2}{\gamma_l^3} \tag{5.43}$$

Its minimum is obtained with

$$\gamma_l = \frac{(\frac{r_{av}}{r_l} \zeta_l \sigma_{ml}^2)^{1/4}}{\sum_{l=0}^{L} \frac{r_{av}}{r_l}(\frac{r_{av}}{r_l} \zeta_l \sigma_{ml}^2)^{1/4}} \tag{5.44}$$

In this case, the reconstruction error is in the mean square:

$$E(e_{rw}^2(i,j,k)) = \frac{1}{\text{SNR}_0} (\sum_{l=0}^{L} (\frac{r_l}{r_{av}})(\frac{r_{av}}{r_l} \zeta_l \sigma_{ml}^2)^{1/4})^4 \tag{5.45}$$

5.5 Summary

This chapter addresses the problem of balancing the transmission impairment of frequency modulation between subbands, when a subband decomposition is performed, or between components, when the visual information consists of multiple components such as Highs/Lows, or luminance/chrominance.

It is found that by adjusting the transmission rate of the components, their SNR can be changed. The optimal rates are found such that they minimize a subjective measure of the transmission impairment.

In a subband decomposition application, a vertical decomposition was found to be very effective. The favorable horizontal noise structure is maintained, while the vertical noise spectrum is better matched to the human visual system. When adaptive modulation of the subbands is performed, the side information can be transmitted at a smaller rate, while the rate of the subbands must be small enough so that they can be accomodated in the transmission channel.

Chapter 6

Applications

In this chapter, two applications of adaptive frequency modulation and multirate transmission are presented. The first application is a dual-in-one-transponder system. In this system, two NTSC signals are fit into one 27 MHz transponder. The system can be used for broadcasting or for distribution between terrestrial and cable transmitters. The second application involves direct broadcasting by satellite of high definition television in a 36 MHz transponder.

6.1 A Dual-in-One-Transponder System

There are many situations where more channels are required than can be provided for. For example, News Gathering is often required from remote locations in which case a satellite link is required for rapid transmission. A dual-in-one-transponder system will double the number of links available for this purpose. Another application is video distribution by satellite of cable programs. In this case, operating cost can be reduced as more channels are available from a given satellite. The dual-in-one system presented next is based on the MAC system currently under development in Europe. A brief description of MAC is included followed by the specifications of an adaptive MAC system.

Frequency multiplexing, when used in FM transmission of composite signals, has long been described as inadequate. The color information lies in the frequency range

where triangular noise is largest, and is badly affected by the transmission. Additionally, the presence of a subcarrier increases the effects of the non linearity that take place in the traveling-wave-tube amplifier. As a response to these problems, many researchers have promoted a time-multiplex approach of the components, one example of which is the MAC system. MAC stands for Multiplex Analog Components, and a more complete description can be found in [Gar88]. The components arrangement is illustrated in figure 6.1. The two chrominance components U and V are transmitted alternatively on a line-by-line basis, but are shown here in a field-sequential manner for more clarity. The luminance signal is time-compressed by 3:2, while chrominance is time-compressed by 3:1. This increases the MAC basebandwidth from 4.2 MHz required for the NTSC composite signal up to 6.3 MHz. In order to reduce the chrominance noise, U and V are low-pass filtered to typically one quarter of the luminance bandwidth.
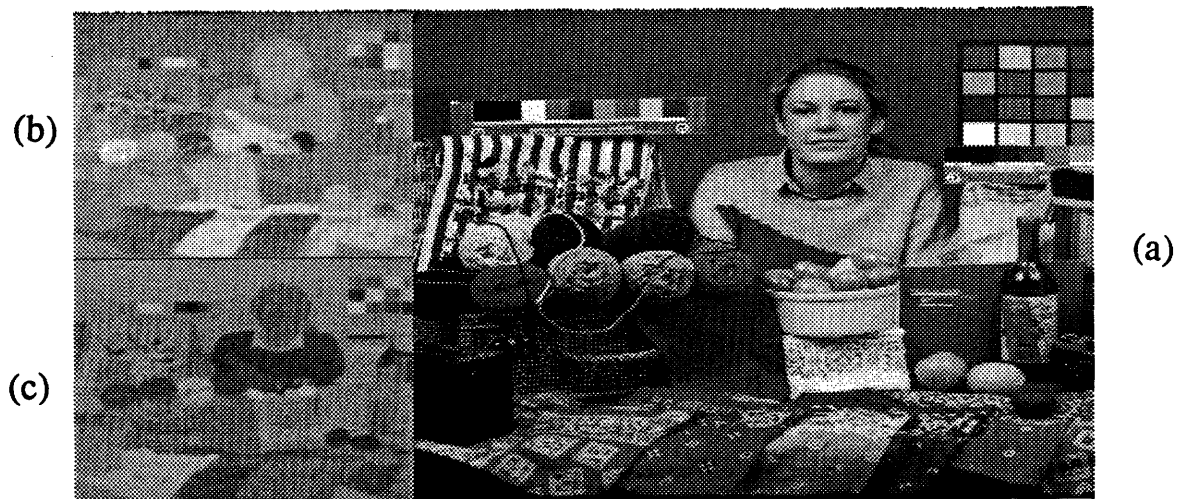

(b)

(a)

(c)

Figure 6.1: Undecoded field of a MAC system. (a) luminance. (b) U component. (c) V component.

MAC typically requires a bandwidth of 24 to 27 MHz in order to improve the low carrier-to-noise ratio offered by satellites. This bandwidth expansion from 6.3 MHz

to 27 MHz is not available with dual-in-one-transponder systems. If the two signals that share the transponder are time multiplexed, the basebandwidth is doubled and the bandwidth expansion factor is halved. This makes the video signal more prone to noise. If time multiplex is applied, adaptive frequency modulation and multirate transmission can compensate for the loss of performance, as described below. In the following, the adaptive MAC system is demonstrated for a 13 MHz transponder, which is equivalent to time multiplexing two signals in a 27 MHz transponder.

6.1.1   Baseband Specifications of Adaptive MAC

In this section, a new system, called adaptive MAC, is proposed. As shown in chapter 3, adaptive modulation requires the transmission of side information, namely the Lows component and the adaptation factors. In order to make room for the side information, the chrominance information is subsampled temporally by a factor of two. Hence the refresh rate of the chrominance is 30 Hz. The subjective impairment is not significant, as human sensitivity to moving color details is reduced.

The vertical and horizontal resolutions of the chrominance are also reduced through downsampling in a way similar to what is done in the MAC systems. First, as in the original MAC system, every other line is discarded. Second, horizontal subsampling is performed. The chrominance component in MAC systems is horizontally low-pass filtered, and then time-compressed. We can equivalently prefilter and downsample horizontally, and then time-expand. This approach is preferred here as it fits better in the multirate approach developed in chapter 5. The downsampling factors for the different components are shown in table 6.1, together with the block size of the adaptation factor.

6.1.2   Transmission of the Adaptation Factors

Transmission of the adaptation factors is considered in this section using frequency modulation.

In the effective peak frequency deviation algorithm (EPFD), the adaptation infor-

Component	Horizontal	Vertical	Temporal
Highs	1	1	1
Lows	8	4	1
V	4	2	2
U	4	2	2
Adaptation	4	4	1

Table 6.1: Downsampling factors of the analog components and block size of the adaptation factors in a field of adaptive MAC

mation is computed from the maximum absolute value of $y(i,j)$ in each block. The adaptation information can be transmitted either in form of the factors $f_a(m,n)$, or in form of the maximum value that was used to compute it. The latter solution is chosen, for reasons that will become clearer in the following. Let's call Max, the maximum absolute value of $y(i,j)$ in a block.

$$\text{Max} = \text{Max}(|y(i,j)|)_{\text{in block m,n}} \qquad (6.1)$$

After transmission, the value Max $+ \Delta$Max is used to reconstruct the Highs, where ΔMax is a Gaussian random variable that accounts for the transmission noise. If the interpolation of the factors between blocks is ignored, we have in block (m,n):

$$y'(i,j) = y(i,j) + \Delta\text{Max} \frac{y(i,j)}{\text{Max}} \qquad (6.2)$$

where only transmission noise on the side information is considered.

Let's call $e_f(i,j)$ the reconstruction error due to the noisy transmission of Max.

$$e_f(i,j) = \Delta\text{Max} \frac{y(i,j)}{\text{Max}} \qquad (6.3)$$

The error is in the mean square:

$$E(e_f^2(i,j)) = E(\Delta\text{Max}^2) \, E(\frac{y^2(i,j)}{\text{Max}^2}) \qquad (6.4)$$

where we used the fact that ΔMax is statistically independent of $y(i,j)$ and Max, as is the case when the bandlimiting distortion is negligible.

We shall again use the assumption that $y(i,j)$ is a stationary first-order Gauss-Markov process when it is only considered in a block of small size. The process is doubly stochastic, as the standard deviation in a block is itself a random variable. The pdf of the standard deviation describes the amount of "activity" that is expected in any block. An example of such a probability density function can be found in figure 3.11, for the picture "Girl".

The expected value of the error square in a block with standard deviation σ is:

$$E(e_f^2(i,j)|\sigma) = E(\Delta \text{Max}^2)E(\frac{y^2(i,j)}{\text{Max}^2}|\sigma) \qquad (6.5)$$

The second term, obviously smaller than one, could be evaluated using a Monte Carlo simulation. What is important here is that the quantity is scale-invariant and thus doesn't depend on σ. However, it varies a little from block to block, since the correlation coefficient is not constant in all blocks.

$$E(\frac{y^2(i,j)}{\text{Max}^2}|\sigma) = E(\frac{y^2(i,j)}{\text{Max}^2}) \qquad (6.6)$$

It appears that the reconstruction error in a block due to a noisy transmission of Max is, in the mean square, independent of the block standard deviation. In other words, it doesn't depend on the signal activity in the block. This is not the case when the adaptation factor for each block is transmitted, and explains our choice of transmitting Max rather than $f_a(m,n)$. The scale-invariance property holds regardless of the pdf of $y(i,j)$. A Gaussian model is used here for simplicity.

Ideally, we wish to decrease noise in blank areas, since masking has no effect there. We show here that $e_f(i,j)$ can be made small in blank areas, due to a nonlinearity introduced in the computation of the adaptation factor. This makes transmission of the side information very robust to transmission noise in terms of image quality. An example is included to demonstrate this.

As already mentioned in chapter 4, the adaptation factor $f_a(m,n)$ doesn't need to be increased up to very large values. In blank areas, where Max is small, $f_a(m,n)$ doesn't need to be larger than 32, in which case the noise reduction is already about 30 dB, a very large improvement. Hence, with no loss of performance, we have:

$$f_a(m,n) = \text{Min}(\alpha\frac{y_0}{\text{Max}}, 32) \qquad (6.7)$$

CNR	B_{IF}/r	SNR of Max	picture SNR -masked noise-
12 dB	1	16 dB	37.5 dB

Table 6.2: FM transmission of side information

where, as in section 4.3.1, y_0 is the peak value of $y(i,j)$ in all blocks.

It appears that if Max is small enough, such as in blank areas, the effect of transmission noise is null. It is due to the fact that $f_a(m,n)$ is clipped to 32. Namely, if $f'_a(m,n)$ is the adaptation factor obtained from $\mathrm{Max}' = \mathrm{Max} + \Delta\mathrm{Max}$, we have:

$$f'_a(m,n) = f_a(m,n) = 32 \leftrightarrow e_f(i,j) = 0 \tag{6.8}$$

when the following two conditions are satisfied:

$$\mathrm{Max} < \frac{\alpha\, y_0}{32} \tag{6.9}$$

$$\mathrm{Max} + \Delta\mathrm{Max} < \frac{\alpha\, y_0}{32} \tag{6.10}$$

Also, $E(e_f^2(i,j))$ is decreased when only condition (6.10) is satisfied, since $f'_a(m,n)$ is clipped to 32.

As a result, the effect of transmission noise on the side information is negligible in blank areas if the peak value in each block is used as side information, and if clipping of the adaptation factor is performed. This is demonstrated in the following example, where frequency modulation is used to transmit the side information.

In this example, the FM noise improvement is kept to a minimum by using a transmission rate equal to the channel bandwidth. This amounts to very-narrowband FM. The channel used in the simulation is an ideal low pass filter. The noise in the channel represents a critical case, with a CNR of 12 dB. The effect of the noisy transmission of the side information is shown in figure 6.2, together with the reconstruction error $e_f(i,j)$. Noise is not perceptible, although the transmission of the side information was very noisy. Its measurement is reported in table 6.2.
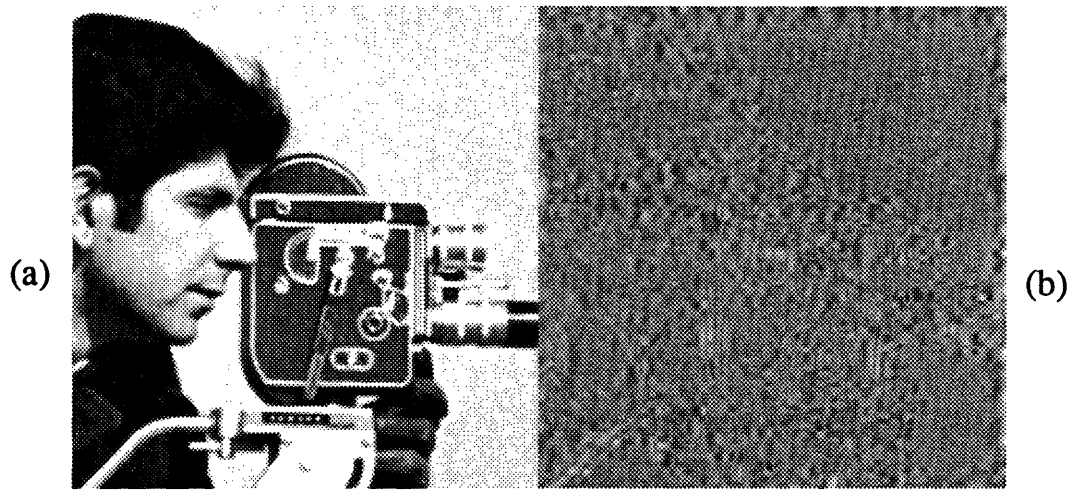
Figure 6.2: Effect of noisy transmission of side information. (a) reconstructed picture. (b) reconstruction error scaled up by five.

6.1.3 Multirate Specifications of Adaptive MAC

In this section, the rate of transmission of each component is varied in order to get a good balance of noise between components. The time expansion factors derived should have simple ratios with each other, in order to simplify the coder and decoder. In this case, the sampling clocks for the components are easily derived from a single master clock.

The following factors are relevant to determine the expansion factors. First, larger time-expansion factors γ_l yield better SNR, as the SNR increases with the cube of γ_l. Also, the subsampling performed on some components affects its weighting factor. When interpolation of the component is performed up to full size, the noise frequency distribution is compressed towards lower frequencies, where the weighting function is larger. The larger the downsampling, the more visible the noise, as the size of a noisy pixel corresponds to a larger area. The derivation of the weighting factor in this case is very similar to that of section 5.2. Finally, adaptive modulation improves the masked SNR by about 16 dB on the Highs component.

133

Component	l	r_l	$\gamma_l(2F_s/r_{av})$	$10\log(\gamma_l/\gamma_0)^3$	$-10\log(\zeta_{wi})$	$10\log(G_{adap})$
Highs	0	$2F_s$	2:3	0 dB	10.4 dB	16.4 dB
Lows	1	$2F_s/32$	8:3	18 dB	2.2 dB	-
U,V	2	$2F_s/16$	4:3	9 dB	4 dB	-
Max	3	$2F_s/16$	4:3	9 dB	-	-

Table 6.3: Multirate specification of adaptive MAC system

A good compromise is found by allocating a large expansion factor to the Lows, and a medium factor to the chrominance components and to the Max component. The multirate specification of the system is shown in table 6.3, together with the SNR improvement due to the time expansion, the horizontal weighting, and the adaptive frequency modulation.

The average transmission rate of the source is:

$$r_{av} = 2F_s(1 + \frac{1}{32} + \frac{2}{16} + \frac{1}{16}) = 2F_s \frac{39}{32} \tag{6.11}$$

It describes the sample rate of the source, when all components have equal rate. For NTSC signals, we have $F_s = 4.2$ MHz, and:

$$r_{av} = 10.2 \text{ Msamples/sec} \tag{6.12}$$

The basebandwidth of the encoded signal is 6.3 MHz. The component arrangement is very similar to that of the regular MAC system, allowing for easy transcoding between regular MAC and adaptive MAC. As can be seen in figure 6.3, one chrominance component is not transmitted. Instead, the side information is transmitted, consisting of the time-expanded Lows, and the time-expanded Max component. The chrominance component from the previous field is used in place of that not transmitted. In other words, U and V are transmitted alternatively on a field-by-field basis. An enlargement of a segment of the reconstructed luminance frame is shown in figure 6.4 after transmission through a 13 MHz channel with a CNR of 12 dB. The SNR measurement is shown in table 6.4. Similarly, in a simulation involving a sequence of

(b)

(c)

(d)

(a)

Figure 6.3: Undecoded field of an adaptive MAC system. (a) Highs. (b) U/V component (alternatively). (c) Lows. (d) Max.

CNR	B_{IF}	picture SNR -masked noise-
12 dB	13 MHz	24 dB

Table 6.4: Measurement of the decoded SNR in a frame of the adaptive MAC system

30 fields transmitted in the same channel, the lower frame rate of the chrominance didn't present any visible impairment. From a regular viewing distance, neither the noise nor the chrominance loss of resolution is perceptible.

6.2 Satellite Broadcasting of High Definition Television

In the next decade, high definition television will be introduced in the North American continent. Many systems are competing to become the national standard, such as those from Zenith, David Sarnoff Laboratories, and MIT, among others. The different
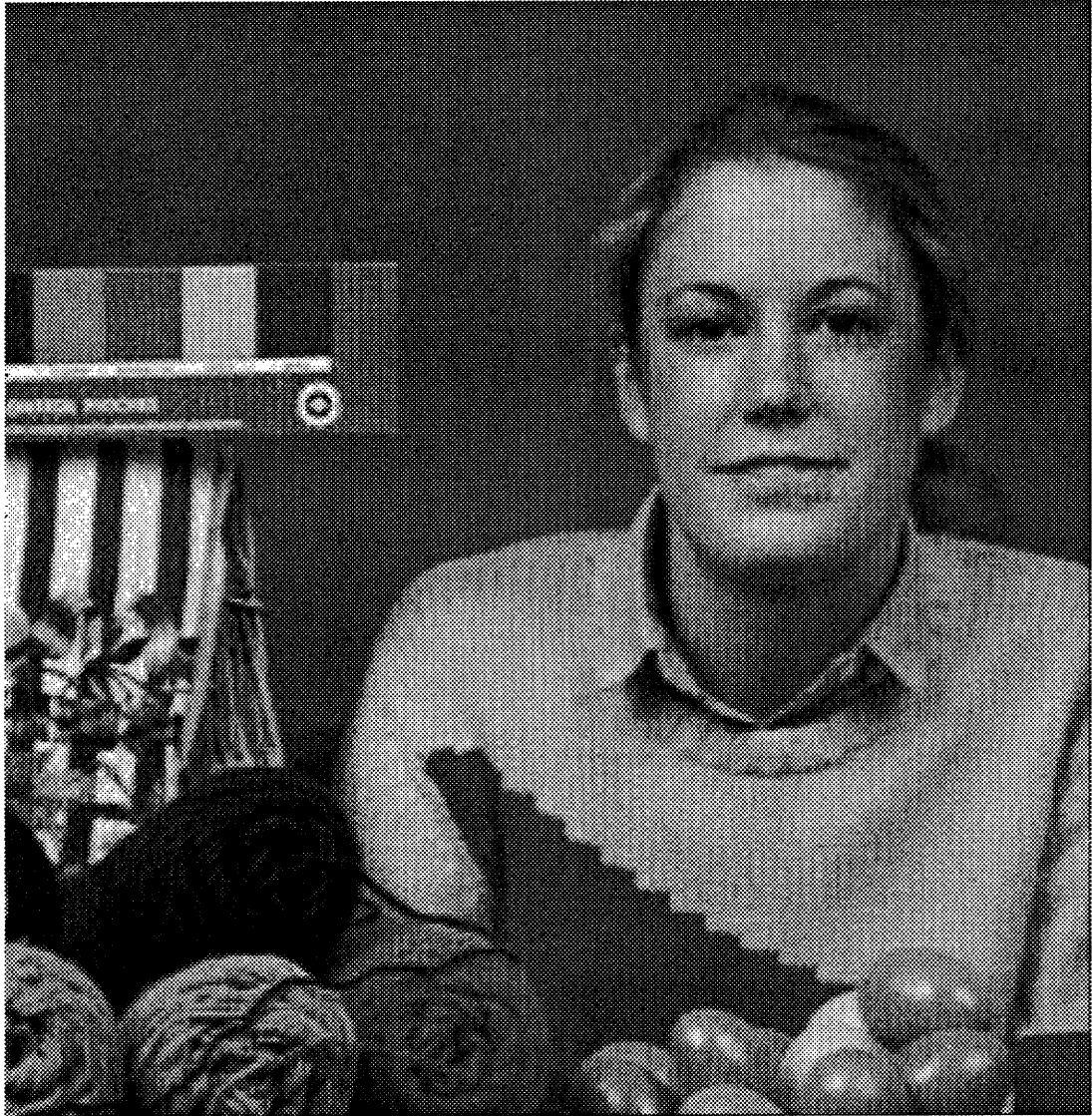
Figure 6.4: Segment (256 x 256) of decoded frame of the adaptive MAC system after transmission through a 13 MHz channel with a CNR of 12 dB

media are likely to coexist, and it is not clear which one will dominate the market, if any. These media are over-the-air broadcasting, optical fibers, cable, and finally satellite broadcasting. Another uncertainty factor arises as what will be the main production standard. Film is the most widespread source material nowadays, but the high definition systems proposed in Japan, Europe, and recently in the U.S. are likely to become more important. In this perspective, easy interfacing between different production standards and media is an important feature that should be included in any new transmission scheme. Along these lines, a friendly-family of transmission standards has been proposed, where the three dimensional video signal is decomposed into subbands [Sch89]. The "packaging", i.e. the selection of the number of subbands can be tailored to each transmission system. In this section, a system is proposed based on the friendly-family concept. It illustrates the "component" nature of the family, and the flexibility available with frequency modulation, when adaptive modulation and multirate transmission are used.

It is clear that any system design involves a trade-off between spatial resolution, temporal resolution, and transmission noise. In the system below, high emphasis is given to motion rendition. It is desired to maintain the sharpness of moving edges, hence providing the spatial resolution that the human visual system is capable of when tracking a moving object.

6.2.1   Baseband Specifications of HDTV System

The friendly family of transmission standards [Sch89] involves a division into components, each of which comprises 144 lines at 256 lines per line, and at a rate of 60 Hz. A component is a subband obtained from a very high resolution original, and the larger the number of components, the better the resolution. The system presented here involves only subbands obtained from a spatial decomposition, in order to keep the cost of the decoder down. More elaborate systems also involve temporal decomposition into subbands at 12 Hz (5 bands decomposition).

In this system, each frame consists of 4 x 4 components, which brings the resolution to 576 lines with 1024 pels per line at 60 frames per second. The static resolution is

about that of other HDTV proposals, when the Kell factor is taken into account and when sampling at twice the luminance bandwidth is performed. However, two major impairments of other interlaced HDTV systems are eliminated [Kra87]:

- interline flicker

- vertical modulation of moving detail

Out of the 16 components, only 10 are actually transmitted with no loss of quality. The fixed selection of 10 out of 16 components is based on the reduced diagonal resolution of the human visual system. The components involving high diagonal frequencies are discarded. The fixed selection scheme contrasts with adaptive selection schemes where some components are selected depending on their local energy [JN84]. The transmitted components are shown in figure 6.5, together with the two chrominance components I and Q. The color space used here is the same as that used in the NTSC standard. Additionally, the Lows component is subtracted from the signal in order to make possible adaptive frequency modulation.

Adaptive modulation is applied to the Highs components. In order to decrease the adaptation information, the components are grouped together in four groups that share a common set of adaptation factors, as shown in figure 5.9. The block size for the adaptation information is 1 x 1, i.e. one factor for each pel, except for group 3, where blocks of size 2 x 2 are used. The energy of components in group 3 is smaller, and so is the transmission noise, making the group less sensitive to noise. Obviously the group 0, consisting of only one component, is the most sensitive since noise added on this component is of lowest spatial frequency, and since energy in this band is larger yielding in turn larger transmission noise. To further decrease the adaptation information, two successive frames are processed with a single set of factors. The smallest factor for each block in the two frames is used in the modulator, and transmitted to the receiver. The loss of performance when objects are moving was found to be negligible.

Another component not displayed in figure 6.5 is the audio component. A rate of 88.2 ksamples/sec is necessary for one high quality stereo audio channel.

Component	Horizontal	Vertical	Temporal
Highs subbands	4	4	1
Lows	8	8	1
I	4	4	2
Q	8	8	2
Adaptation	1 (2)	1 (2)	2

Table 6.5: Downsampling factors of the analog components and block size of the adaptation factors in a frame of HDTV system

The baseband specifications of the components are shown in table 6.5.

Finally, the rate of the HDTV source is computed:

$$r_{av} = 28.84 \text{ Msamples/sec} \tag{6.13}$$

Such a source could be fit in a transponder as narrow as 29 MHz when very narrow band FM is used, with a very small roll off. However, it was found that 36 Mhz was required in order to allow for some components to be time-expanded.

6.2.2 Transmission of the Adaptation Information

For the same reasons as in section 6.1.2, the maximum absolute value in a block is transmitted. We shall call Max_g the adaptation component for group g. However, unlike the adaptive MAC system, we allow for the adaptation factor to get very large, otherwise some loss of performance is observed after the components are put together. It follows that noise due to the transmission of Max is no longer negligible in blank areas. Since noise in blank areas is most undesirable, static companding is used [SB81]. It consists of transmitting $f(Max)$ rather than Max, where $f(x)$ is a nonlinear function. After reception, the inverse $g(x)$ of the nonlinear function is used to get back the original value of Max.

In our case:

$$f(x) = x_0 (\frac{x}{x_0})^{1/2} \tag{6.14}$$

139
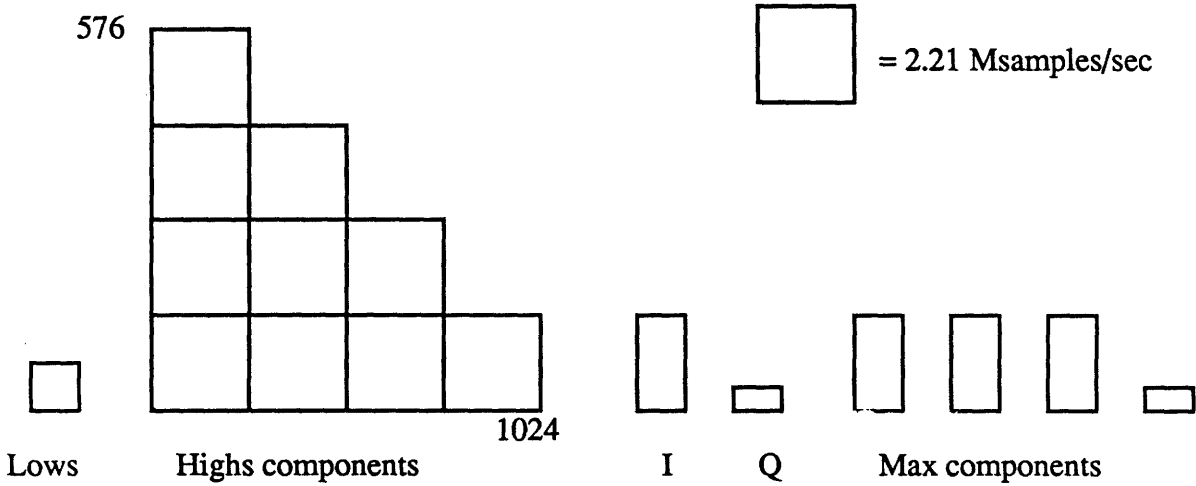
Figure 6.5: Components of HDTV system

$$g(x) = x_0(\frac{x}{x_0})^2 \tag{6.15}$$

where x_0 is the largest possible value of Max, in order to normalize the argument between 0 and 1.

Let's call $e_{fc}(i,j)$ the reconstruction error due to a noisy transmission of the companded value of Max. The error is in the mean square a function of Max. It can be shown that in a first approximation the mean square error is:

$$E(e_{fc}^2(i,j)) = E(\frac{4\Delta \text{Max}^2}{\text{Max } x_0}y^2(i,j)) \tag{6.16}$$

Hence, for a given value of Max, the ratio of companded to uncompanded mean square error is:

$$\frac{E(e_{fc}^2|\text{Max})}{E(e_f^2|\text{Max})} = \frac{4\text{Max}}{x_0} \tag{6.17}$$

It appears that when Max is very large -on the order of x_0-, the error mean square is increased by a factor of four. However, the visual impairment in such a block is not significant, since the masking effect is large. On the other hand, the error is reduced largely in blocks where the Max value is small. In these blocks, masking has little effect, and the subjective improvement is very large.

140

6.2.3 Multirate Specifications of HDTV System

Again, the rate of transmission of each component is adjusted so as to get the best perceptual quality. The different rates should be easily obtained from a single master clock.

The Highs components are less sensitive to noise, since they are adaptively modulated. High-frequency subbands are very robust to noise because they generate only high-frequency noise, and also because the noise variance is in direct proportion to their energy, the latter being smaller than in low frequency subbands. Interestingly, the subband that represents both lowest horizontal and vertical frequencies is quite sensitive to transmission noise, although adaptive modulation is applied. Adaptive modulation is possible on this subband, since the Lows component has been removed. The very low frequency nature of the noise contribution from this component makes it necessary to increase its time expansion factor with respect to the other subbands. The other subbands only require the smallest possible expansion factor γ_{min}. When the Highs components are time-expanded with γ_{min}, the peak transmission rate results. In our case case, $\gamma_{min} = 0.838$, and the peak transmission rate is 34.4 Msamples/sec. This peak transmission rate can be accommodated in a 36 MHz transponder using very narrow band FM with a small roll off factor. The time-expansion factors for each component are shown in table 6.6.

An enlargement of a segment of the full HDTV picture after transmission through a 36 MHz channel with a CNR of 15 dB is shown in figure 6.6. The SNR measurement is shown in table 6.7. The SNR in this case is smaller than that of the adaptive MAC system, because some subbands are not transmitted and because a small amount of clipping has been applied on those transmitted. These effects are subjectively minor. However, at a closer distance, the noise can be more easily distinguished than in the adaptive MAC case. This effect is due to the subband decomposition. When only some subbands are noisy, as in the neighborhood of edges, there is a structure in the noise due to the synthesis filters. This structure makes the noise more visible, and is one drawback of subband decomposition in adaptive frequency modulation. Additionally, a slight amount of ringing can be seen on sharp diagonal edges due to

Component	l	r_l	$\gamma_l/\gamma_{\min}$	$10\log(\gamma_l/\gamma_{\min})^3$
H subband	0	$2F_s/16$	3:2	5.3 dB
H subbands	1-9	$2F_s/16$	1:1	0 dB
L	10	$2F_s/64$	3:1	14.3 dB
I	11	$2F_s/32$	3:2	5.3 dB
Q	12	$2F_s/128$	3:2	5.3 dB
Max_0	13	$2F_s/32$	2:1	9 dB
Max_1	14	$2F_s/32$	3:2	5.3 dB
Max_2	15	$2F_s/32$	3:2	5.3 dB
Max_3	16	$2F_s/128$	1:1	0 dB

Table 6.6: Multirate specification of HDTV system

CNR	B_{IF}	picture SNR -masked noise-
15 dB	36 MHz	20.4 dB

Table 6.7: Measurement of the decoded SNR in a frame of the HDTV system

Figure 6.6: Segment (256 x 256) of a decoded frame of the HDTV system after transmission through a 36 MHz channel with a CNR of 15 dB

144

the missing diagonal bands.

In a simulation involving a sequence of 30 frames transmitted in a 36 MHz channel, the effect of noise for a CNR of 12 dB is barely perceptible at regular viewing distance (four times the picture height). At a closer distance, a flicker can be noticed in moving texture. It is attributed to the adaptation factors block size, which is two frames wide, and degrades the performance of adaptive modulation for moving objects.

6.3   Summary

Two applications of adaptive frequency modulation have been described. They illustrate the potential of the technique for real-world systems.

In a first system, called adaptive Mac, it is shown that the side information necessary to perform adaptive modulation can be transmitted in place of one chrominance component. In this case, the frame rate of the chrominance information is halved with no loss of quality. The coding of the adaptation information is then shown to be very robust to noise, and is well suited to frequency modulation.

In another system, direct broadcasting by satellite is demonstrated for high definition television. By transmitting only 10 subbands out of a 4 x 4 subband decomposition, a lower transmission rate is achieved with no loss of quality. Adaptive modulation of the subbands allows for robust transmission in a noisy channel. Here again, the adaptation information is well suited to frequency modulation, but static companding proved necessary to decrease noise in blank areas.

In the two systems, the SNR of the different components is adjusted by varying the time-expansion factor. In this way, the Lows and the chrominance components are transmitted with minimal degradation, even in noisy channel.

Additionally, a complete simulation of the two systems has been conducted for a color sequence of 30 frames. The resulting quality demonstrates that the systems specifications are well suited to high-quality television in a narrow bandwidth FM channel.

Chapter 7

Conclusions

7.1 Contributions

This thesis is concerned primarily with the problem of improving the performance of frequency modulation for television signals. Two findings motivate this research. The first finding involves the relative inefficient usage of the bandwidth in existing systems, in order to handle some worst-case signals. The second finding is that a system can always be improved when the subjective impairments in different components, frequency bands, or regions of the picture is not well balanced. Such is typically the case of FM transmission of NTSC composite signals, in which case the two chrominance components are much more noisy than the luminance component. Similarly, blank regions of the picture are badly affected by transmission noise, more than busy regions.

Along these lines, the contribution of this thesis is twofold. Namely, subjective improvements are possible by:

- adjusting the modulation characteristics in order to improve the usage of the channel bandwidth (adaptive modulation).

- adjusting the rate of transmission of the different components in order to balance the subjective impairments (multirate transmission).

As a mean to perform these tasks, the different elements involved in the transmis-

sion process are modeled. First, the source is modeled by a Gauss Markov process, whose second order statistics are allowed to vary from block to block. Second, the communication link is seen globally as a linear system, at the output of which two error terms are added , one being independent noise, and the other being bandlimiting distortion. Finally, the sink, i.e. the human viewer, is shown to object to a transmission error in a way related to its frequency content (weighting), and its position with respect to masking stimuli.

An interesting result of this approach it to allow for connection between the different models. For example, the standard deviation in a block is related to the short-time modulated bandwidth. The bandlimiting distortion is then shown to depend on the short-time bandwidth. Finally, the average masking in a block is also related to the standard deviation in this block. These models, although simplistic, give us some insight, and also provide us with tools to analyze the system performance.

In a second step of the research, investigation of the transmission limitations suggests different algorithms to control the modulator. They are the effective peak frequency deviation algorithm (EPFD), the iterative algorithm, and the frame iterative algorithm. These algorithms are shown to improve performance when compared to the simple-minded algorithm based on signal modeling.

Finally, the performance of adaptive modulation and multirate transmission are demonstrated in two real-world systems. A full simulation of the system for a color sequence demonstrates the quality available from a noisy narrow band channel when improved frequency modulation is used.

7.2   Suggestions for Further Work

The HDTV system proposed in chapter 6 relies on a spatial subband decomposition in order to decrease the transmission rate of the source. Designing a HDTV system based on three-dimensional subband decomposition would lower this rate further, making possible DBS of HDTV in a 27 MHz channel. Such a system could use a fixed selection of components, at the expense of a reduced sharpness of moving

146

objects, or adaptive selection, with improved resolution of moving objects.

The design should involve a trade-off of noise between the subbands, even if they are adaptively modulated. Noise on lower temporal bands is likely to be noticeable, unless extra care is taken, such as increasing their time expansion factor. This may be necessary as noise in FM channels can be very large.

Bibliography

[AM83]   K. Arbenz and J.C. Martin. *Transmission de l'information*. Masson, 1983.

[Ber71]   T. Berger. *Rate Distortion Theory A Mathematical Basis for Data Compression*. Prentice-Hall, 1971.

[B.G66]   B.G.Cramer. Optimum linear filtering of analog signals innoisy channels. *IEEE Transactions on Audio and Electroacoustics*, Au-14(1):3–15, March 1966.

[BR68]   E. Bedrosian and S.O. Rice. Distortion and crosstalk of linearly filtered, angle-modulated signals. *Proceedings of the IEEE*, 56(1):2–13, January 1968.

[dC84]   Frederic de Coulon. *Theorie et Traitement des signaux*. Editions Georgi, 1984.

[DV88]   Z. Doganata and P.P. Vaidyanathan. General synthesis procedures for fir lossless transfer matrices, for perfect-reconstruction multirate filter bank applications. *IEEE Trans. Acoust. Speech Signal Process.*, ASSP-36(10):1561–1574, Oct. 1988.

[Gag87]   H.P. Gaggioni. The evolution of video technologies. *IEEE Communications Magazine*, 25(11):20–36, November 1987.

[Gal68]   R.G. Gallager. *Information Theory and Reliable Communication*. Wiley, 1968.

[Gar88]   P.N. Gardiner. The UK D-MAC/packet standard for DBS. *IEEE Transactions on Consumer Electronics*, 34(1):128–135, February 1988.

[Gir88]   B. Girod. *Ein Modell der Menschlichen Visuellen Wahrnehmung zur Irrelevanzenreduktion von Fernsehluminanzsignalen*. PhD thesis, Institut fuer Theoretische Nachrichtentechnik und Informationsverarbeitung, Universitaet Hannover, 1988.

[HC76]   B.R. Hunt and T.M. Cannon. Nonstationary assumptions for Gaussian models of images. *IEEE Transactions on Systems, Man, and Cybernetics*, pages 876–881, December 1976.

[JI65]   J.M.Wozencraft and I.M.Jacobs. *Principle of Communication Engineering*. John Wiley and Sons, 1965.

[JN84]   N.S. Jayant and P. Noll. *Digital Coding of Waveforms*. Prentice-Hall, 1984.

[JT37]   J.R.Carson and T.C.Fry. Variable frequency electric circuit theory. *Bell Syst. Tech. J.*, 16:513–540, October 1937.

[Kra87]   E.A. Krause. *Motion Estimation for Frame-Rate Conversion*. PhD thesis, MIT Dept. of Electr. Eng., June 1987. CIPG Memo ATRP-T-70.

[Mal86]   H.S. Malvar. *Optimal Pre- and Post-Filtering in Noisy Sampled-Data Systems*. PhD thesis, Massachusetts Institute of Technology, August 1986.

[MMHY87]   H. Murakami, S. Matsumoto, Y. Hatori, and H. Yamamoto. 15/30 mbit/s universal digital TV codec using a median adaptive predictive coding method. *IEEE Trans. Commun.*, 35(6):637–645, June 1987.

[NP77]   A.N. Netravali and B. Prasada. Adaptive quantization of picture signals using spatial masking. *Proc. IEEE*, 65(4):536–548, Apr. 1977.

[OS75]   A.V. Oppenheim and R.W. Schafer. *Digital Signal Processing*. Prentice-Hall, 1975.

[PA86]   L.C. Palmer and A.Shenoy. Simulation of tv transmission over the communications satellite channel. *IEEE Transactions on Communications*, COM-34(2):188–199, February 1986.

[Pan65]   P.F. Panter. *Modulation, Noise, and Spectral Analysis*. McGraw-Hill, 1965.

[Pap83]   A. Papoulis. Random modulation: A review. *IEEE Transactions on Accoustic, Speech, and Signal Processing*, ASSP-31(1):96–105, February 1983.

[Pra78]   W.K. Pratt. *Digital Image Processing*. Wiley-Interscience, 1978.

[Pro83]   John G. Proakis. *Digital Communications*. McGraw-Hill, 1983.

[Rho85]   C.W. Rhodes. A tutorial on improved systems for color television transmission by satellite. *IEEE Transactions on Broadcasting*, BC-31:1–9, 1985.

[Rut68]   C.L. Ruthroff. Computation of fm distortion in linear network for bandlimited periodic signals. *The Bell System Technical Journal*, 47(6):1043–1063, July 1968.

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

[Sch89]   W.F. Schreiber. A friendly family of transmission standards for all media and all frame rates. Cipg memo atrp-t-97, Feb. 1989.

[SP88]   W.F. Schreiber and J. Piot. Video transmission by adaptive frequency modulation. *IEEE Communications Magazine*, November 1988.

[ST85]   W.F. Schreiber and D.E. Troxel. Transformation between continuous and discrete representations of images: A perceptual approach. *IEEE Trans. Pattern Anal. Mach. Intell.*, 7(2):178–186, Mar. 1985.

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

[Tom86]   A.S. Tom. Prediction of FIR pre and post-filter performance based upon a visual model. Master's thesis, MIT Dept. of Electr. Eng., June 1986.

[Tre68]   H.L. Van Trees. *Detection, Estimation, and Linear Modulation Theory -Part 2*. John Wiley and Sons Inc., 1968.

[Vai87]   P.P. Vaidyanathan. Quadrature mirror filter banks, M-band extensions and perfect-reconstruction techniques. *IEEE Acoust. Speech Signal Process. Mag.*, 4(3):4–20, July 1987.

[Vet84]   M. Vetterli. Multi-dimensional sub-band coding: Some theory and algorithms. *Signal Processing*, 6(2):97–112, Apr. 1984.

[Vet87]   M. Vetterli. A theory of multirate filter banks. *IEEE Trans. Acoust. Speech Signal Process.*, ASSP-35(3):356–372, Mar. 1987.

[Wea71]   L.E. Weaver. *Television Measurement Techniques*. Peter Peregrinus Ltd., 1971.

[WO86]   J.W. Woods and S.D. O'Neil. Subband coding of images. *IEEE Trans. Acoust. Speech Signal Process.*, ASSP-34(5):1278–1288, Oct. 1986.

Biography

Julien Piot was born in Bangalore (India) in 1960. He received the Dipl. Ing. El. degree from the Swiss Federal Institute of Technology, Lausanne, Switzerland, in 1983.

From 1983 to 1986 he worked in the Digital Audio Lab of Willy Studer AG., Regensdorf, Switzerland, where he developed a new type of equalizer for digital magnetic recording.

From 1986 to 1989 he was a Research Assistant in the Advanced Television Research Program at the Massachusetts Institute of Technology, Cambridge.

His research interests include communications, television picture processing, digital signal processing, and speech and audio processing.

Mr. Piot holds several patents in the field of digital magnetic recording.