COMBINED SPEECH AND CHANNEL CODING FOR MOBILE RADIO APPLICATIONS

by
Hong Shi

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF APPLIED SCIENCE (ENGINEERING SCIENCE)

in the School
of
Engineering Science

SIMON FRASER UNIVERSITY
February 1993

APPROVAL

Name: Hong Shi

Degree: Master of Applied Science (Engineering Science)

Title of thesis: **Combined Speech and Channel Coding for Mobile Radio Applications**

Examining Committee: Dr. Andrew Rawicz, Chairman

Senior Supervisor
Dr. Paul Ho

Senior Supervisor
Dr. Vladimir Cuperman

Supervisor
Dr. Jacques Vaisey

Supervisor
Dr. Jim Cavers

Examiner
Dr. Shawn Stapleton

Date Approved: JAN. 28, 1993

PARTIAL COPYRIGHT LICENSE

I hereby grant to Simon Fraser University the right to lend my thesis, project or extended essay (the title of which is shown below) to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users. I further agree that permission for multiple copying of this work for scholarly purposes may be granted by me or the Dean of Graduate Studies. It is understood that copying or publication of this work for financial gain shall not be allowed without my written permission.

Title of Thesis/Project/Extended Essay

"Combined Speech and Channel Coding for Mobile Radio Applications"

Author:

(signature)

HONG SHI

(name)

March 1, 1993

(date)

To my dearest parents with lots of love

ABSTRACT

We present in this thesis a combined speech and channel coding scheme for digital mobile communications. The speech coding algorithm is based on Code-Excited Linear Prediction (CELP) and achieves good speech quality at a rate of 4 kb/s. For the channel code, both rate-compatible punctured convolutional (RCPC) codes and punctured Reed-Solomon (PRS) codes are considered. In the case of RCPC codes, soft decision decoding is considered in addition to the simpler hard decision decoding. The modulation format chosen in our study is $\pi/4$-DQPSK with differential detection, the format adopted for the North American Digital Cellular System. Unequal error protection is used based on the bit error sensitivities of the different speech parameters. The performance of the combined speech and channel coder is studied under different mobile channel conditions, such as fade rates, signal-to-noise ratios, and interleaving delays. The results indicate that with no interleaving delay and large channel signal-to-noise ratio, PRS codes provide marginally better protection in terms of the segmental signal-to-noise ratio (SSNR) of the reconstructed speech, while RCPC codes perform better at lower channel SNR. Informal listening tests were performed with the combined codecs using PRS codes and RCPC codes. The comparison shows that, for combined codecs designed using the procedure suggested in the thesis, the difference between the quality of the reconstructed speech in a clean channel and that in a Rayleigh fading channel is imperceptible at channel SNRs larger than 20 dB for both the RCPC and the PRS based codecs.

ACKNOWLEDGEMENTS

I would like to express my greatest gratitude to my senior supervisors, Dr. Vladimir Cuperman and Dr. Paul Ho, for their assistance, encouragement and guidance throughout the course of this research.

Special thanks to my colleague Peter Lupini for providing the source code for the speech coder used in this study.

I am also grateful to Dr. Geng Wu and my colleague Bhaskar Bhattacharya for their help during the last two years.

I would also like to thank my colleague Aamir Husain for his helpful suggestions to this thesis.

Finally, I would like to express my appreciation to all my friends who have shown their support and encouragement throughout my stay at Simon Fraser University.

CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS

ADPCM	Adaptive Differential Pulse Code Modulation
APCM	Adaptive Pulse Code Modulation
ATC	Adaptive Transform Coding
AVPC	Adaptive Vector Predictive Coding
AWGN	Additive White Gaussian Noise
BCH	Bose-Chandhari-Hocquenghem
BER	Bit Error Rate
BPF	Band-Pass Filter
BS-FEC	Bit-Selective Forward Error Correction
CELP	Code-Excited Linear Predictive
CELP/FEC	Combined CELP and Forward Error Correction system
CELP/PRS	Combined CELP and PRS codec
CELP/RCPC	Combined CELP and RCPC codec
DOD	Department of Defense
DPCM	Differential Pulse Code Modulation
DQPSK	Differential Quadrature Phase Shift Keying
DSP	Digital Signal Processing
FCC	Federal Communication Commission
FEC	Forward Error Correction
FM	Frequency Modulation
FSM	Finite-State Machine
GF	Galois Field
GSM	Groupe Speciale Mobile
IC	Integrated Circuit
LFSR	Linear Feedback Shift Register
LPC	Linear Predictive Coding
LSP	Line Spectrum Pairs
MDS	Maximum-Distance-Separable
MOS	Mean Opinion Score

MSB	Most Significant Bit
MSE	Mean-Squared Error
NBC	Natural Binary Code
PCM	Pulse Code Modulation
PRS	Punctured Reed-Solomon
QPSK	Quadrature Phase Shift Keying
RCPC	Rate-Compatible Punctured Convolutional
RS	Reed-Solomon
SBC	Subband Coding
SNR	Signal to Noise Ratio
SSNR	Segmental Signal to Noise Ratio
UEP	Unequal Error Protection
VLSI	Very Large Scale Integration
VQ	Vector Quantization
VQCELP	Vector Quantized Code Excited Linear Predictive
VSELP	Vector-Sum Excited Linear Prediction
VXC	Vector Excitation Coding
ZIR	Zero Input Response
ZSR	Zero State Response

CHAPTER 1

INTRODUCTION

Mobile radio communications is a broad concept of communication which includes radiophones, dispatching systems, radio paging systems, packet radios or radio-telephones (mobile phones). Mobile radio service can be seen as early as in 1921 [1]. We are focusing at mobile radio-telephone communication in this study.

The future of mobile radio-telephone communication depends on techniques of network planning and mobile radio equipment design that will enable efficient and economic use of the radio spectrum. The United States FCC (Federal Communication Commission) allocated a 40 MHz bandwidth in the 800-900 MHz frequency range for this purpose in 1974. To meet the steadily increasing customer demand for mobile radio service within the limitation of the available FCC frequency allocations, the solution is to develop a workable plan for reusing the assigned channels within each band of frequencies.

Cellular systems have experienced very fast growth in both North America and parts of Europe and Asia. In the near future, a tenfold increase in the number of users is expected according to commercial sources. The current cellular system in North

America uses analog voice transmission by means of FM modulation. The channel separation is 30 kHz, compared to 25 kHz in some of the analog systems in Europe. Since the United States FCC is currently not prepared to assign a larger frequency band for mobile radio applications, a standard for digital mobile radio system that provides much higher capacity is now under way in North America.

There are different standards proposed for digital mobile radio communications in North America, such as the Full Rate standard where the gross bit rate is designed at 16 kbps/user and the proposed Half Rate standard where the bit rate is fixed at 8 kbps/user. In our study, we are aiming at the Half Rate Digital Mobile Radio Standard for North America. In the Half Rate Standard [24], the gross bit rate is 8 kbps/user, of which about 6.4 kbps is assigned for source and channel coding, 1.5 kbps is available for the use of system overhead, and the remaining 0.1 kbps are left over for other uses, such as error detection.

Digital transmission of speech is becoming more prevalent in telecommunications because it provides numerous advantages such as: compatibility with data transmission, the use of modern transmission techniques, and the possibility of encryption. Recent advances in speech coding techniques have made digital transmission practical in mobile radio applications. For example, in 1991 the United States Department of Defense (DOD) developed a standard vocoder operating at the rate of 4.8 kbits/s for the purpose of facilitating inter-operability between radio telecommunication facilities and systems of the United States Federal Government. This standard is referred to as the Federal Standard 1016. Since 1984, NASA has been exploring the feasibility of speech coding at 4.8 kbits/s for their Mobile Satellite Experiment (MSAT-X) [3]. The Telecommunication Industry Association (TIA) adopted a speech codec called Vector-Sum Excited Linear Predictive Coding (VSELP) at a rate of 7.95 kbits/s for the North American Full Rate Digital Cellular Standard initialized recently.

In mobile radio applications, speech coders are required to have low bit rates and

provide high quality speech (near toll quality). Many efficient speech coders have been developed for providing high quality speech at rates between 4.8 kbits/s and 16 kbits/s [3, 4, 5, 6, 7]. There are even speech coders which provide speech with good intelligibility at rates as low as 2.4 kbits/s [8, 9]. However, one major point to consider is that when the parameters of the low rate speech coder is transmitted on a physical mobile radio channel characterized by severe disturbance due to adjacent channel interference and multi-path fading, the quality of the reconstructed speech degrades dramatically. Thus, the speech coders used in mobile communications are required to have low bit rates, high quality and robustness to channel errors which may be either random or bursty. With this motivation, we study combined speech and channel coding with possible application to the North American Half Rate Digital Cellular Standard.

Recent research advances in speech coding have shown that Code Excited Linear Predictive (CELP) coding is a very promising technique for transmission of high quality speech at low bit rates [3, 4, 5, 10]. This results in the great interest of applying CELP coding scheme to mobile radio communications. CELP coding is one of the analysis-by-synthesis schemes which utilize the coding information of the spectrum envelope and source information. It provides a low rate output but is vulnerable to noise and transmission errors as compared to waveform coding methods which makes use of the speech waveform and produces a higher rate signal output which is more robust to transmission errors.

In a CELP coder, the speech signal is represented by a set of parameters which include LPC coefficients (short-term filter coefficients), pitch period, pitch gain, excitation codeword, and excitation gain. These parameters are quantized, coded, and transmitted over a physical channel, such as a telephone line, satellite link, or mobile radio channel. In the absence of channel errors, the CELP coder can reproduce speech with good intelligibility at bit rates as low as 2.4 kbits/s [11], however, when there is channel disturbance, the reconstructed speech quality degrades significantly.

Thus, error protection must be applied to attain acceptable speech quality over these channels.

There are several ways of improving the robustness of the CELP coded transmission, some of these methods are efficient index assignment [12, 13, 14], parameter smoothing with error detection [15, 16], and Forward Error Correction (FEC) [17, 18, 19, 20, 21]. Efficient index assignment, or Gray ordering, usually assumes that only one bit in a binary representation of a parameter is disturbed by a channel error, which may not be the case in a harsh channel with a high bit error probability and/or bursty errors. The parameter smoothing method is based on the error detection of a binary code representing a parameter. If an error is detected, the current parameter value will either be replaced by the previous one or interpolated with the previous values. This technique works in the case when only one frame is corrupted. In a harsh channel with a high bit error rate, there can be errors in the same parameter in several consecutive frames. In these cases this scheme may fail and cause very audible glitches, squeaks, or blasts because the parameter smoothing is done with incorrect references. FEC is considered as a more powerful technique for improving the communication performance with the trade-off of increasing the band-width by introducing redundancy to the transmitted information sequence. According to the channel coding theorem by Shannon [22, 23], the transmitted information can be recovered at the receiver to any degree of accuracy desired as long as the channel capacity C is greater than the information rate R.

Since our study is on mobile radio applications, the communication channel is modeled as a Rayleigh fading channel where the channel errors are bursty in nature [25]. In this work, we consider Reed-Solomon (RS) codes and convolutional codes with both hard and soft decision decoding for error protection. RS codes are chosen for their capability of correcting multiple errors [26, 27]. Convolutional codes with interleaving can also provide good protection against error bursts. We first carefully examined the bit error sensitivity of the CELP coder's output. Since there exists a

large dynamic range in the bit error sensitivity among the different CELP parameters, unequal error protection can be applied. According to the requirement of unequal protection, rate-compatible punctured Reed-Solomon (PRS) codes have been used [28, 29, 30]. Experiments of combining the speech codec and the rate-compatible Punctured Convolutional (RCPC) codec [52, 53] have also been conducted.

The CELP speech coding scheme was employed in this study. We used a 4 kbps CELP coder which is a variant of the DOD 4.8 kbps standard [31]. Channel bandwidth is a scarce resource in mobile radio communication. Our goal is thus to find a protection scheme which provides as much improvement as possible while not significantly increasing the transmission rate. In this study, we employ a scheme of joint speech and channel coding where channel protection is optimally applied to the speech elements according to the different sensitivities of the speech elements to channel errors.

The objective performance measurement in this thesis is the Segmental Signal to Noise Ratio (SSNR) of the reconstructed speech compared to the original speech. Combined source and channel coding has been previously studied for simple waveform coders such as 32 kbits/s DPCM [17] or 16 kbits/s subband coder [19]. However, in these systems, the channel code rate allocation is not designed optimally. There are also work done earlier in combining speech and channel coding in an optimum fashion [32, 33]. In this thesis, an exhaustive search method for both the combined CELP/PRS codec and the combined CELP/RCPC codec is used to find the optimal channel code rate allocation under certain channel conditions according to the SSNR of the reconstructed speech.

In this thesis, simulation has been done under different sets of channel and transmission conditions. Simulation results for combined CELP/PRS codec and combined CELP/RCPC codec with both hard and soft decision decoding show that error protection provides significant improvement in the reconstructed speech quality. We also conducted studies from the perspectives of the Doppler frequency and the interleaving

delay. Simulation results for these purposes are also included.

In Chapter 2 we give a overview of the system configuration. Chapter 3 will focus on the forward error correction schemes. The combined speech and channel coding system we studied in this thesis will be introduced in Chapter 4, and the experimental results will be presented in Chapter 5. We will draw the conclusion and lay out the future work in the last chapter.

CHAPTER 2

System Overview

The mathematical foundation for digital communications was established by Claude Shannon in two papers published in 1948 [22, 23]. In this pioneering work he formulated the basic problem of reliable transmission of information in statistical terms, using probabilistic models for information sources and communication channels. With unique intuition, Shannon perceived that the goals of approaching error-free digital communication over noisy channels and of maximally efficient conversion of analog signals to digital form were dual facets of the same problem.

Source coding is a process which represents the source output by a sequence of binary digits. The goal of channel coding is to reliably transmit this binary sequence from the source encoder to the source decoder at the receiving end through a noisy channel. This thesis focuses on combined source and channel coding as implied by the title. This chapter begins with a description of the basic block diagram of a combined source and channel coding system in Section 2.1. In Section 2.2 we give an overview of the source coding techniques. The communication channel in our study is briefly discussed in Section 2.3. In Section 2.4, the channel coding techniques are outlined. Finally, an overview of the combined source and channel coding methods are given in

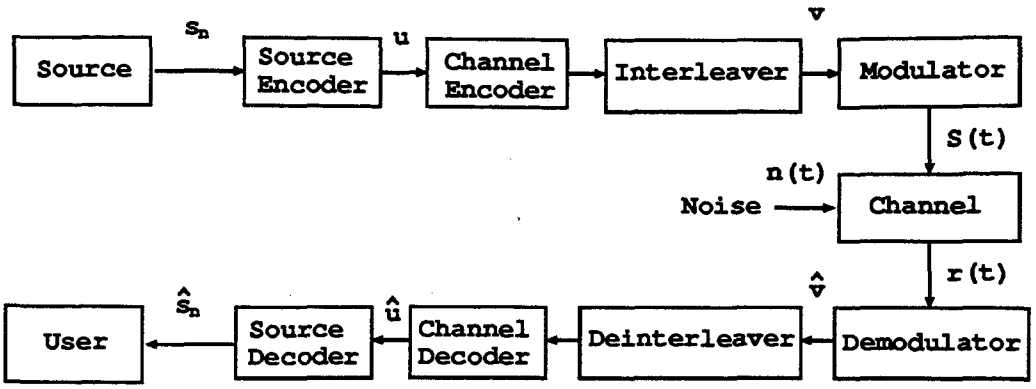


Figure 2.1: Model of a digital communication system

Section 2.5.

2.1 Model of a Digital Communication System

The basic elements of a digital communication system are illustrated by the general block diagram shown in Fig. 2.1. The function of this system is to transmit the information from the source to a destination user as accurately as possible.

The information source generates messages which are to be transmitted to the receiver. In general, the characteristics of the messages can be classified into two categories: analog information and discrete information. In a digital communication system, the analog information is first converted to discrete form via the process of sampling and quantization.

The source encoder first converts the source output, denoted by s_n, into a sequence of binary digits **u**. We would like to find an efficient representation that results in little or no redundancy, that is, we would like to represent the source output by as few binary digits (bits) as possible. The process of efficiently converting the output of a source into a sequence of bits is called *source coding*. We shall describe source

coding techniques in Section 2.2.

The sequence of binary digits from the source encoder is to be transmitted through a channel to the receiver. In order to combat noise and interference and, thus to increase the reliability of the data transmitted through the channel, it is often necessary to introduce in a controlled manner some redundancy to the sequence $\mathbf{u}$. This is known as *channel coding*. The channel encoder transforms the sequence $\mathbf{u}$ into another sequence $\mathbf{v}$ by adding redundancy to $\mathbf{u}$. The codeword $\mathbf{v}$ can be either a binary sequence or an M-ary sequence in different applications. In this thesis we only deal with binary sequences.

Real channels are basically waveform channels and hence, they can not be used to directly transmit the sequence of binary digits $\mathbf{v}$. A device known as a digital modulator is required that converts the digital information into waveforms that are compatible with the characteristics of the channel. This waveform $s(t)$ enters the channel and is corrupted by the channel noise and interference. At the receiving end, the received waveform from the channel is represented by $r(t)$. First, the digital demodulator transforms each received channel waveform into a binary stream $\hat{\mathbf{v}}$, which is in correspondence with the channel encoded sequence $\mathbf{v}$. The channel decoder processes the received sequence $\hat{\mathbf{v}}$ to generate the estimated information sequence $\hat{\mathbf{u}}$. The channel decoder uses the codeword redundancy to correct the errors in the received sequence $\hat{\mathbf{u}}$ caused by the channel noise. Ideally, the decoded sequence $\hat{\mathbf{u}}$ will be the same as the information sequence $\mathbf{u}$ if the error caused by the channel noise is correctable.

The source decoder reconstructs the source output from the estimated information sequence $\hat{\mathbf{u}}$ and delivers it to the destination. Due to channel decoding errors and possible distortion in the source encoder, the reconstructed source signal $\hat{s}_n$ will not be exactly the same as the original source signal s_n.

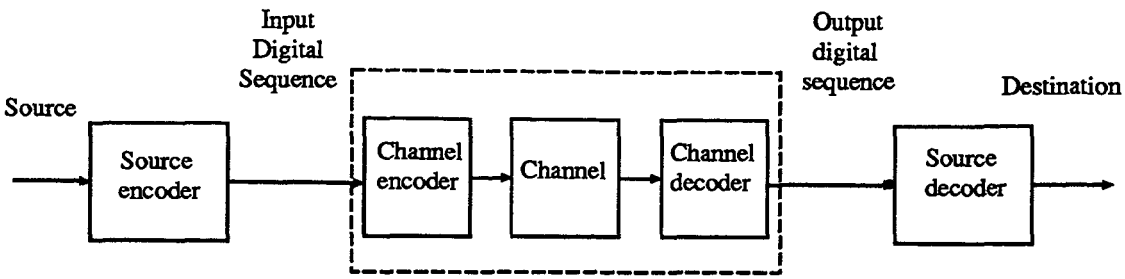


Figure 2.2: Basic model of a digital communication system

It should be mentioned that in this study, the source codec is a CELP (Code-Excited Linear Predictive) codec and that the channel codec is either a Rate-Compatible Punctured Reed-Solomon Codec or a Rate-Compatible Punctured Convolutional Codec. These codecs will be discussed in detail later in chapter 2 and 3 respectively.

2.2 Source Coding

The digital communication system shown in Fig. 2.1 can be simplified to the one shown in Fig. 2.2. Ignore for the moment the channel with its encoder and decoder (within the dashed contour in Fig. 2.2) and replace it by a direct connection called a "noiseless channel", we will get the block diagram of a source coding system. In this section, we will talk about one of the two facets of the dual problem in digital communications, *source coding* or how to efficiently convert an analog signal to digital form.

In speech coding, the source signal is analog with continuity in both time and amplitude. The frequency range of the speech signal is assumed to be between 200 to 3400 Hz. Before applying any source coding technique, the analog speech signal has to be sampled. The Nyquist sampling theorem [34] provides a link between continuous-time signals and discrete-time signals. If the analog source signal is band-limited within the frequency range of 0 - W Hz, then the sampling rate f_s has to satisfy f_s

$\geq$ 2W in order not to result in a loss of information. Sampling of the analog speech signal produces a discrete-time signal s_n with continuous amplitude.

2.2.1 Data Compression

The source signal s_n with continuous amplitude can not be represented exactly by a digital sequence because the source output sequence takes on values from an infinite set, and thus can not be mapped one-to-one onto a discrete set, i.e., a digital alphabet. This is because the entropy H(s) of the continuous amplitude signal s_n is infinite. According to the converse to the channel coding theorem [35], if the entropy per second, $H(s)/T_s$, of the source is greater than the channel capacity per second, C/T_c, then there exists a constant $\alpha > 0$ such that $P_e \geq \alpha$ for all sequence lengths (P_e is the error probability). This theorem shows that it is impossible for a communication system to operate with arbitrarily small average error probability when the information rate R of the source is greater than channel capacity C.

In practice, there is no such physical channel that has a infinite capacity. The best that can be done in mapping the source into a digital sequence is to tolerate some distortion at the destination after the source decoder operation. If we introduce a digitized sequence $\mathbf{u}$ with $H(\mathbf{u}) \leq C$ so that it approximates the continuous amplitude speech signal s_n, then according to the channel coding theorem, the approximation $\mathbf{u}$ can be reconstituted to any degree of accuracy desired. This is the concept of *rate distortion theory*. The operation of transforming the speech signal s_n into its approximation sequence $\mathbf{u}$ is referred to as data compression or source coding. Rate distortion theory is the basis of data compression and it establishes the theoretically minimum *rate* required to represent a source such that the representative satisfies a given fidelity criterion, one within the allowed *distortion* [35].

A block diagram of the source coding system is shown in Fig. 2.3. The source

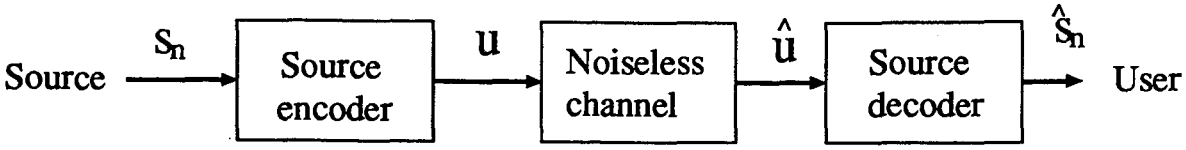


Figure 2.3: A simple diagram of a source coding system

encoder converts the speech signal s_n into a binary sequence $\mathbf{u}$ and the sequence $\mathbf{u}$ will be transmitted to the receiver. Here, we assume a noiseless channel. Consequently the received binary sequence $\hat{\mathbf{u}}$ will be identical to the transmitted information sequence $\mathbf{u}$. The source decoder then constructs the speech signal $\hat{s}_n$ from the binary sequence $\hat{\mathbf{u}}$. The objective of source coding is to minimize the number of bits in the binary representation of the speech signal s_n for a given level of fidelity requirement, i.e., the distortion $d(s_n, \hat{s}_n)$ between the original speech signal s_n and the reconstructed one $\hat{s}_n$. The distortion $d(s_n, \hat{s}_n)$ can be any distortion measure, such as the MSE (mean squared error), the SNR, or the SSNR of the reconstructed speech. Alternatively, the objective of source coding can be described as minimizing the distortion $d(s_n, \hat{s}_n)$ for a given transmission rate.

2.2.2 Source Coding Techniques

As seen from the previous subsection, data compression applied to the sampled source signal leads to some distortion or a loss of information. Quantization itself is a data compression operation and results in a loss of information. In the following part of the discussion, when we refer to information loss, we actually refer to the information loss after quantization.

Source coding techniques can be classified into reversible coding and irreversible coding. Reversible coding is based on Shannon's source coding theorem. The principal idea of reversible coding is that the bit rate of the encoded source signal can be reduced by variable length coding according to the probability of occurrence of each codeword.

A short code is used for codes with a high occurrence probability, whereas a long code is used for low occurrence probability codes. Thus the coding efficiency is increased. This is also called *entropy coding*. Shannon-Fano coding and Huffman coding are examples of entropy coding [11].

In speech coding applications, a certain amount of distortion is usually allowed as long as the distortion does not impair the auditory comprehensibility of speech. Many coding techniques fall into the category of irreversible coding, which results in information loss but maintains a certain fidelity according to a pre-defined distortion measure. According to Furui [11], the basic irreversible coding methods are classified as follows:

1. Nonlinear quantization
2. Adaptive quantization
3. Predictive coding
4. Time and frequency division
5. Transform coding
6. Vector quantization

The simplest speech coding technique is linear Pulse Code Modulation (PCM). In this method, analog speech signals are uniformly quantized similar to the usual A/D conversion. This method does not reduce the information rate since it does not use the speech-specific characteristics. Each sample of the speech signal is independently quantized to one of the 2^b available amplitude levels.

Since speech signals are nonstationary, no fixed quantizer in PCM is optimal for all segments of speech. If we allow the stepsize in PCM to vary according to the vocal

statistics of the speech signal, we get Adaptive PCM (APCM) [36, 37]. APCM utilizes the nonstationarity of the dynamic characteristics of speech amplitude to improve the SNR of the quantized speech.

Since a speech signal has correlation between adjacent samples as well as distant samples, it is more efficient to encode the difference between adjacent samples or the difference between the actual sample value and a predicted value calculated based on the correlation characteristics (prediction residual). This observation leads to the development of differential PCM (DPCM). In DPCM, linear prediction is performed and the prediction residual, defined as,

$$d_t = x_t + \sum_{i=1}^{p} \alpha_i x_{t-i}$$

is quantized and transmitted. Here, x_t is the sample, α_i is the predictor coefficient, and p is the order of the predictor. In the simplest case of first-order linear prediction, the equation becomes $d_t = x_t + \alpha_1 x_{t-1}$. If the predictor coefficient is set to α_1 = -1, the system merely transmits the difference between adjacent samples. In practice, the prediction order is typically between 1 and 16. Since the speech signal is quasi-stationary in nature, the variance and the correlation of the speech signal vary slowly with time. We can thus make the DPCM encoder adapt to the slowly time-varying statistics of the speech signal. This leads to the adaptive DPCM (ADPCM).

The speech coding techniques mentioned above operate in time domain. There are techniques that operate in the frequency domain. The coding method, in which a speech band is divided into several contiguous bands by a bank of band-pass filters (BPFs), with a specific coding strategy employed for each signal band, is called sub-band coding (SBC). In adaptive transform coding (ATC), a speech signal is divided into several frequency bands in a way similar to that with SBC.

All the speech coding techniques listed above are techniques that attempt to faithfully represent the speech waveform. These methods are classified as waveform coding

techniques.

In contrast to the waveform coding techniques, linear predictive coding (LPC) represents a completely different approach to the problem of speech coding. This technique was first used for speech analysis and synthesis by Itakura and Saito [38] and Atal and Schroeder [39] in 1968. It has had a very large impact on every aspect of speech research. The importance of LPC stems from the fact that the speech spectral characteristics can be efficiently represented by a very small number of parameters. Instead of transmitting the samples of the speech signals to the receiver, we transmit the parameters.

In LPC, the present speech sample is approximated by a linear combination of the previous p samples, i.e.

$$s_n + \alpha_1 s_{n-1} + ... + \alpha_p s_{n-p} = e_n$$

where e_n is an uncorrelated statistical variable having a mean value of 0 and a variance of σ^2. The predicted value $\hat{s}_n$ is then

$$\hat{s}_n = -\sum_{i=1}^{p} \alpha_i s_{n-i}$$

If we define the linear prediction filter as

$$F(z) = -\sum_{i=1}^{p} \alpha_i z^{-i}$$

we can see that

$$S(z) = \frac{E(z)}{1 - F(z)} = \frac{E(z)}{A(z)}$$

where $E(z)$ and $S(z)$ are the Z-transform of e_n and s_n. In other words, the speech signal is modeled as the output of a linear system $H(z) = \frac{G}{1-F(z)}$ excited by an appropriate input signal [5, 37, 36]. Appropriate excitation functions are either sequences of impulses or sequences of white noise with unit variance, depending on whether the speech signal is voiced or unvoiced. The parameter G is the gain of the filter and

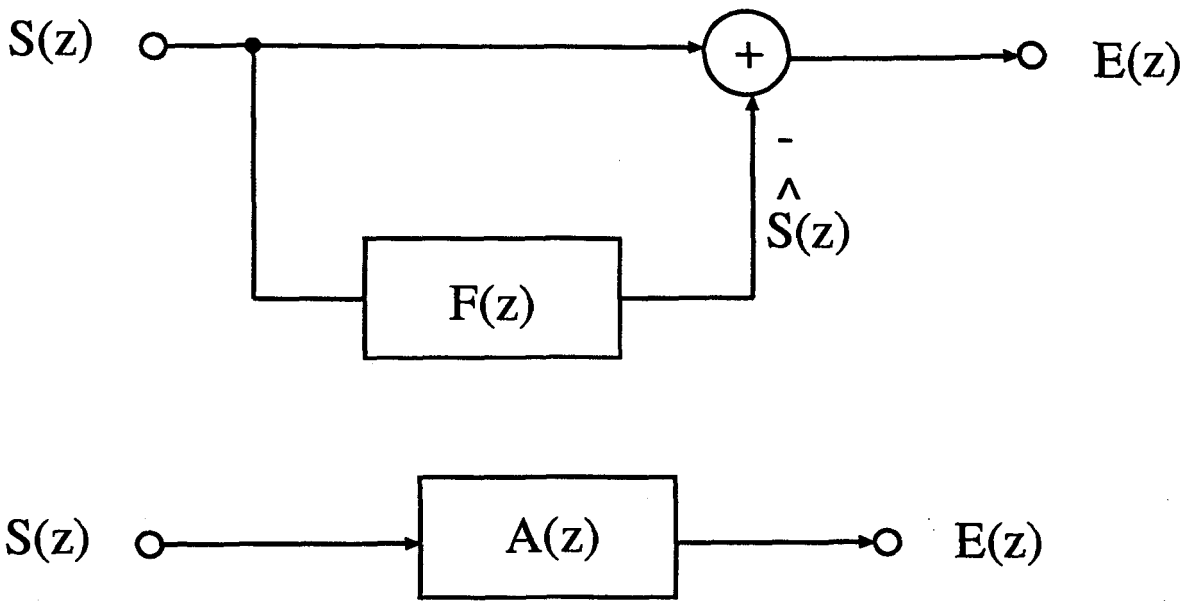


Figure 2.4: Linear prediction model block diagram

e_n is the error between the observed value s_n and the predicted value $\hat{s}_n$. The linear prediction model is shown in Fig. 2.4.

The filter coefficients α_i, i = 1,2,...,p, can be estimated by either forward or backward adaptation. In forward adaptation, the predictor parameters are computed in the speech encoder using the original speech signal, and then transmitted to the decoder as side information. While in the backward adaptation, they are estimated at both the encoder and decoder from the reconstructed speech signal.

The one-dimensional quantization methods described so far are generally classified as forms of scalar quantization. That is, speech signals or residual signals are quantized on a sample-by-sample basis. A fundamental result of the rate distortion theory is that better performance can be achieved by quantizing vectors instead of scalars, i.e., waveform samples or spectral envelope parameters are jointly quantized. This process is called Vector Quantization (VQ) [7]. The combination of vector quantization with linear prediction results in more efficient speech coding algorithms, such

as Adaptive Vector Predictive Coding (AVPC) [40], Code-Excited Linear Predictive Coding (CELP) [4, 5], and Vector Excitation Coding (VXC) [3]. We use the CELP coder in this thesis.

2.2.3 CELP Algorithm

The objective of this subsection is to describe the CELP coder used in the combined speech and channel coding system (see Fig. 2.1). A complexity reduced VQCELP coder is employed as the source coding subsystem in this study. The organization of this section is as follows. First, we give a brief introduction of the CELP coding scheme. The objective measures used, the Signal to Noise Ratio and the Segmental Signal to Noise Ratio, are then introduced. Thirdly, we will discuss the CELP coder structure. The 4k CELP coder used in this thesis is described at last.

2.2.3.1 Introduction

Because the transmission channel bandwidth is strictly limited in many new digital communication systems, speech coding at very low bit rates without sacrificing voice quality is becoming increasingly important. These communication applications include, for example, packet voice transmission, voice encryption, voice mail, and mobile telephony. The speech coding technologies to achieve high quality voice is well developed for bit rates above 8 kbits/s [7, 41]. The major effort now is focussed in encoding speech at rates between 2.4 - 4.8 kbits/s [10, 15, 6].

Code Excited Linear Predictive (CELP) coding is considered to be a successful technique in producing high quality speech at low bit rates. It is also considered as a candidate for encoding speech in mobile radio applications. In fact, the Full Rate North American Digital Cellular Standard employs the Vector-Sum Excited

Linear Predictive (VSELP) coding, which is a variant of CELP. CELP coding is an analysis-by-synthesis process which involves high computational complexity because of the search for the optimum innovation sequence. A few years ago, this technique was economically not feasible for real time implementation because of its complexity. However, in recent years, many low complexity alternatives to the basic CELP coder have been introduced with only a slight degradation in the quality of the reconstructed speech. Also, the rapid progress in the VLSI technology has made it possible to implement more complex systems using advanced DSP processors. These two factors together make it practical to implement CELP speech coders in real time. In this section, we will introduce a complexity reduced CELP coder.

2.2.3.2 Objective Performance Measurements

Human perception mechanism is considered to be the evaluation criterion of the reconstructed speech quality. Therefore, perceptual and subjective testing procedures constitute an integral part of coder design and evaluation. One commonly used subjective testing procedure is known as the Mean Opinion Score (MOS). Unfortunately, subjective evaluations of speech quality or intelligibility are very time consuming and expensive [41]. Compared to the subjective evaluations, the objective measures are much easier and less expensive to use. Since objective measures can be repeatedly computed and are consistent, they are often used in designing the speech coders. There are two objective measures used for evaluating the speech quality [42], e.g., the Signal to Noise Ratio (SNR) and the Segmental Signal to Noise Ratio (SSNR). These two objective measures will be used throughout the course of this thesis. It should be pointed out that one disadvantage of using an objective measure is that it tends to have a poor correlation with the human perception of quality.

Figure 2.3 shows the basic block diagram of a speech coding system. In Fig. 2.3, s_n denotes the original speech signal, $\hat{s}_n$ denotes the reconstructed speech signal at the

receiver. The error between the reconstructed speech signal and the original speech is denoted by q_n, and is defined as:

$$q_n = s_n - \hat{s}_n$$

Let the variance of s_n, $\hat{s}_n$, and q_n be denoted by σ_s^2, $\sigma_{\hat{s}}^2$, and σ_q^2, respectively. The standard *signal to noise ratio* of the reconstructed speech signal is the *ratio of signal variance to reconstruction error variance.* The SNR is usually represented in decibels (dB), and is defined as follows:

$$SNR = 10log_{10}(\frac{\sigma_s^2}{\sigma_q^2}) \qquad dB$$

We do not usually know the true variances of s_n, $\hat{s}_n$, and q_n in practice. The SNRs are actually computed from the speech samples s_n and the reconstructed speech samples $\hat{s}_n$ using the following equation:

$$SNR = 10log_{10}\frac{\sum_{n=1}^{L} s_n^2}{\sum_{n=1}^{L}(s_n - \hat{s}_n)^2} \qquad dB$$

where L is the number of samples used in the estimation.

Usually the SNR of the speech signal is computed over a long speech database. Since the speech signal is non-stationary over the whole time axis, the SNR measure of the speech quality often has a poor correlation with human perception of quality. Since the speech signal can be considered as a stationary process during short periods of time, if we measure the SNR over short segments (frames) of speech samples and average over all segments, the result would be more correlated to the subjective quality measure. We call this objective measure the *Segmental Signal to Noise Ratio* (SSNR). The SSNR is normally used as an objective measure in the analysis of speech coders. The definition of SSNR is shown below:

$$SSNR = \frac{1}{K}\sum_{i=1}^{K} SNR_i \qquad dB \tag{2.1}$$

where K is the number of frames in a speech database. In practice, when there are intervals of silence in the speech utterance, any amount of noise will give rise

to a large negative SNR for that segment, which could considerably bias the overall measure of SSNR. A way of solving this problem is to exclude the silent segments when calculating the SSNR. Silent segments are defined to be segments whose energy levels are 40 dB below the long term energy level.

2.2.3.3 Structure of a CELP coder

The analysis-synthesis structure of a CELP coder is depicted in Fig. 2.5. The encoder consists of a short-term predictor $1/A(z)$, a long-term predictor $1/B(z)$ (adaptive codebook), a weighting filter $W(z)$, and a normalized VQ codebook (CB). The input analog speech signal is sampled and quantized into s_n. The speech signal s_n is then segmented into frames with four subframes in each frame. In Fig. 2.5, the filter $1/A(z)$ models the short term correlation of the speech signal, and has the form

$$A(z) = 1 - \sum_{i=1}^{p} a_i z^{-i} \tag{2.2}$$

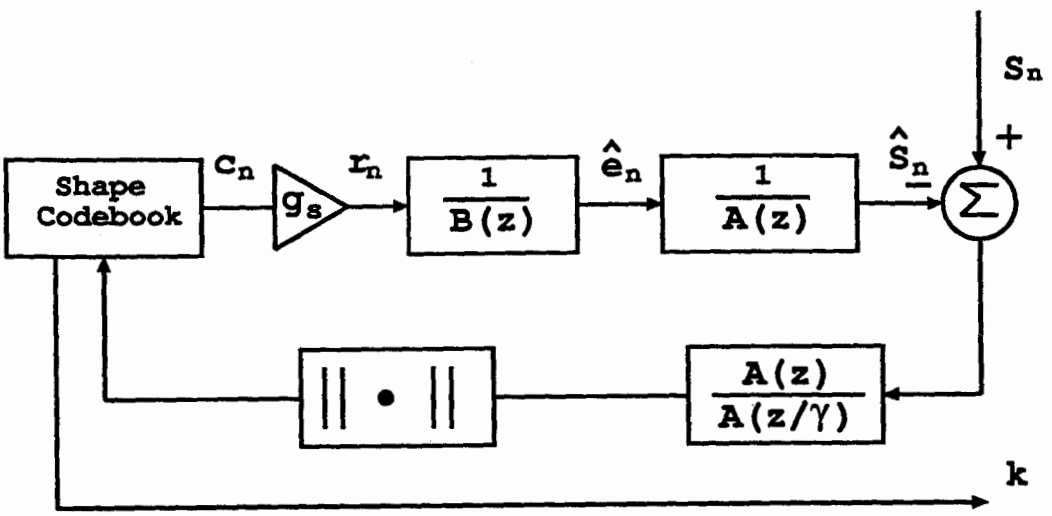
where p is the order of the short term predictor and the a_i are the coefficients of the filter. This filter is also called a linear predictor and the a_i are also called LPC coefficients. This filter is the same as that discussed in Section 2.2.2. The order of the predictor is normally between 10 and 16.

The filter $1/B(z)$ is also called a pitch predictor. It models the long term correlation of the speech signal, which can be used to determine the spectral fine structure. It can also be considered as a linear predictor with the form

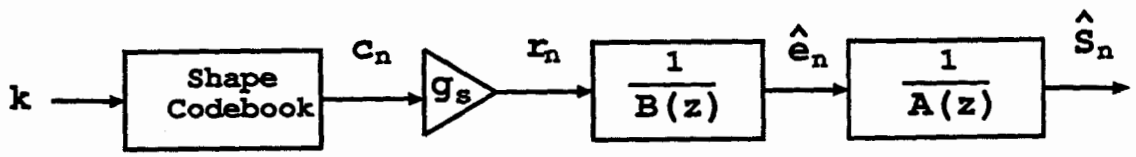
$$B(z) = 1 - \sum_{i=p-m}^{p+m} b_i z^{-i} \tag{2.3}$$

Here, p is the pitch period of the speech signal in samples and the b_i are the predictor coefficients.

The excitation codebook contains M candidate waveforms, and M is usually referred to as the codebook size. In the analysis-by-synthesis process, the difference



(a) Encoder



(b) Decoder

Figure 2.5: CELP coder structure

between the synthesized speech and the original speech is passed through a perceptually weighting filter. The weighted square error is minimized over all possible choices from the codebooks, and the indices corresponding to the best entries are transmitted. The weighting filter $W(z)$ has the form

$$W(z) = \frac{A(z)}{A(z/\gamma)} \tag{2.4}$$

where γ is less than 1. The purpose of the weighting filter is to shape the noise spectrum in the way that the noise level is reduced at the low signal energy region, while the noise level is increased in the high signal energy region. The actually perceived noise level will be reduced by noise weighting.

The speech decoder performs an inverse operation of the encoder. The decoder stores a replica of the excitation codebook, and synthesizes the speech according to the parameters received from the channel.

The basic CELP coder shown in Fig. 2.5 is characterized by high computational complexity because of the search for the optimum excitation sequence. The perceptually weighted error between the original speech and the reconstructed speech is used to select the best excitation codevector from the normalized excitation codebook. The best excitation is found when the following equation is minimized

$$E = \sum_{n=1}^{L} (s_n - \hat{s}_n)^2 \tag{2.5}$$

where L is the length of the excitation sequence. The computational complexity can be reduced with the structure shown in Fig. 2.6. In this structure, the weighting filter is moved before the summation of the original speech s_n and the reconstructed speech $\hat{s}_n$. The Zero Input Response (ZIR) and the Zero State Response (ZSR) are separated. The perceptually weighted reconstructed speech s'_n can be written into three components

$$s'_n = \hat{x}_n + \hat{y}_n + \hat{z}_n \tag{2.6}$$

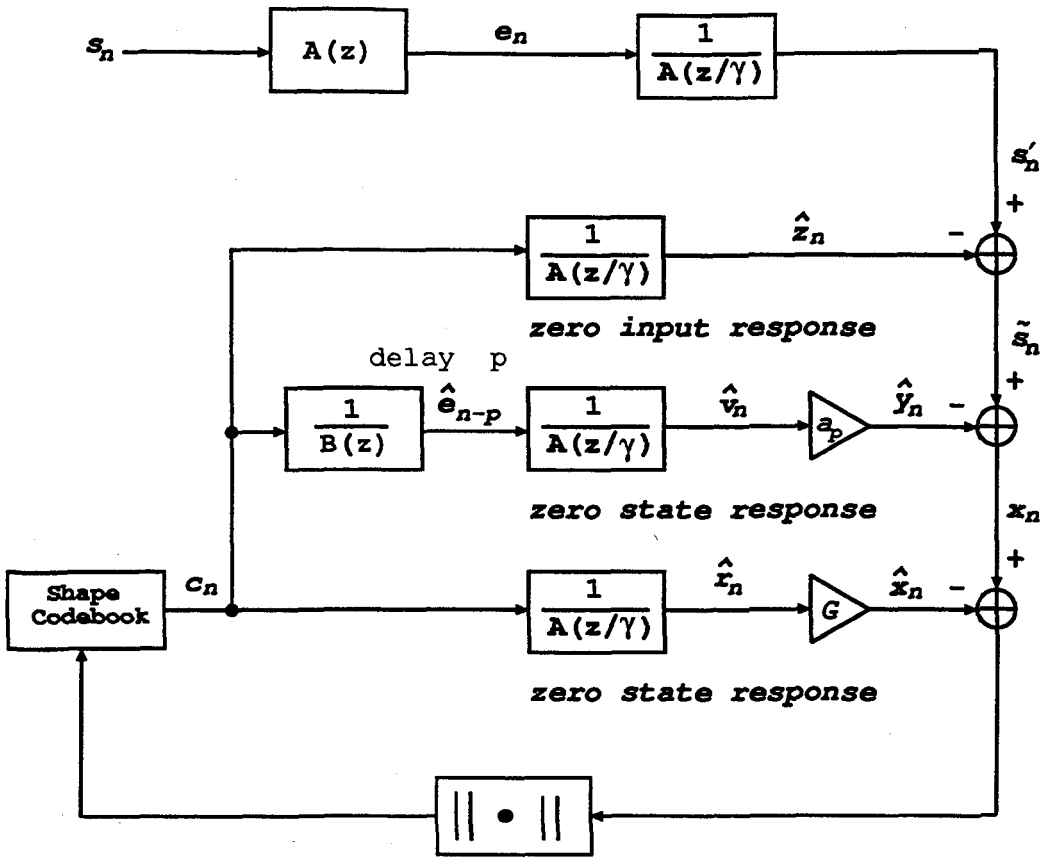


Figure 2.6: Complexity reduced CELP coder

The signal $\hat{z}_n$ is due to the zero input response of the filter $1/A(z/\gamma)$. The signals $\hat{y}_n$ and $\hat{x}_n$ are due to the zero input response of the long-term predictor, and the zero state response of the filter $1/A(z/\gamma)$ to the excitation codewords.

The LPC coefficients are estimated directly from the speech signal using either the autocorrelation method or the covariance method. The autocorrelation method guarantees a stable short term filter and can use the efficient Levinson-Durbin algorithm for solving the Yule-Walker equations, while the covariance method does not necessarily result in a stable short-term filter. The technique for estimating the LPC coefficients a_i is well documented in the literature [37, 36, 11]. The LPC coefficients are calculated once each frame.

The long-term predictor parameters, p and a_p can be jointly optimized with the optimum excitation codeword in a closed-loop search. However, the computational requirements would be extremely demanding since a joint optimization requires the search over all possible excitation vectors and all possible values of the lag d. Three reduced complexity approaches for estimating the long-term predictor parameters were examined by Chan and Cuperman [43]. Typical values of the pitch period range from 20 to 147.

The excitation codeword is chosen after the long-term predictor parameters have been estimated. The initial state of the long-term predictor is subtracted from the speech signal. Then, instead of minimizing Eq. 2.5, we minimize

$$E_x = \sum_{n=1}^{L} (x_n - \hat{x}_n)^2 \tag{2.7}$$

The gain G is given by

$$G = \frac{\sum_{n=1}^{L} \hat{r}_n x_n}{\sum_{n=1}^{L} \hat{r}_n^2} \tag{2.8}$$

Long-term predictor coefficients and the excitation parameters are updated once every frame.

In the 4k CELP codec used in this thesis, the input analog speech signal is sampled at a rate of 8 kHz and is segmented into frames of 35 ms (280 samples), each frame containing 4 subframes of 8.75 ms (70 samples). The short-term predictor is a 10th order linear predictor with coefficients determined by using the autocorrelation method. The LPC coefficients are transformed into LSPs (line spectrum pairs) and scalar quantized. The LSP coefficients are transmitted once each frame. The shape codebook is trained. The index in the shape codebook is transmitted once each subframe. A different approach (delta coding) is used for coding the shape codebook gain, that is, the gain values of the second and fourth subframes are derived from the differential values from the first and third ones respectively. The pitch period ranging from 20 to 147 is quantized using a 7 bit uniform codebook. The adaptive codebook

	LPC	Shape CB	Adaptive CB
Update	35 ms	8.75 ms	8.75 ms
Parameters	10 LSPs	1 gain; 1 index	1 gain; 1 index
Bits / Frame	34	index : 4x9 ; gain : 2x6 2x5	index : 4x7 ; gain : 4x5
Bit Number	1 - 34	index : 35 - 70 ; gain : 71 - 92	index : 93 - 120 ; gain : 121 - 140
Total Number of Bits in a Frame : 140			

Table 2.1: Bit Allocation and Update Rate

index and gain are also transmitted once per subframe. The bit allocation and the rate at which parameters are updated are summarized in Table 2.1 for this 4.0 kbps CELP coder. The typical Segmental Signal to Noise Ratio (SSNR) performance of this codec measured on a data-base containing 28 utterances from 7 male and 7 female speakers (all outside the training sequence used for codec design) is 9.12 dB.

2.3 The Transmission Channel and the Digital Modulation Scheme

Typical transmission channels include telephone lines, microwave links, satellite links, mobile radio links, and so on. These transmission channels are subject to various types of noise disturbances. Noise can be classified as natural noise or man-made noise. The thermal noise from the Earth's surface and the atmosphere plays a significant role in natural noise. Man-made noise includes noise form electrical machinery or ignition systems, or due to interference from signals occupying an adjacent channel in the spectrum, etc.. The overall effect of the noise is to add to the desired message signal

a random component which is generally termed *additive noise*. The received carrier may also exhibit fluctuations in intensity, generally due to the arrival of two or more signal components by different modes of propagation having a randomly varying phase difference. Noise caused by this reason is termed *multiplicative disturbance*.

The channel itself is a waveform channel. As shown in Fig. 2.1, the modulator serves as the interface that transforms the input digital signal, usually binary, into a set of corresponding waveforms. Similarly, the demodulator at the receiving end serves as the opposite of the modulator and acts as the link between the waveform channel and the channel decoder that deals with digital signals. Hence the demodulator accepts waveforms at its input and delivers to the channel decoder a sequence of digital symbols (hard decision decoding) or discrete-time symbols (soft decision decoding). There are many digital modulation schemes that suit different channel and transmission conditions. In this study, we used the $\pi/4$-shifted DQPSK modulation, the modulation scheme adopted for the North American Full Rate Digital Cellular Applications. We assume differential coherent detection [44, 45]. The additive white Gaussian channel and the flat Rayleigh fading channel will be briefly described in this section. The latter is used to model the mobile radio channel. Without loss of generality, the baseband equivalent channel model is used in the analysis of communication systems.

2.3.1 Additive White Gaussian Noise Channel

The additive white Gaussian noise (AWGN) channel is the most commonly used channel model in the analysis of communication systems. *White Gaussian noise* is defined to be a random process, each sample of which is a zero-mean Gaussian random variable and whose power spectral density is flat over the entire frequency range, with a level of $N_0/2$ watts per hertz.

The AWGN channel can be described simply in terms of the input $s(t)$ and the output $r(t)$, which are related by

$$r(t) = s(t) + n(t)$$

where $s(t)$ is the baseband equivalent of the transmitted signal, $r(t)$ is the baseband equivalent of the received signal, and $n(t)$ is a zero-mean, complex Gaussian process with a power spectral density of N_0. With $\frac{\pi}{4}$-DQPSK modulation, $s(t)$can be written as:

$$s(t) = \sum s_k p(t - kT)$$

where s_k is a complex $\frac{\pi}{4}$-DQPSK symbol and $p(t)$ is the transmitted pulse shape. After receiver matched filtering and sampling [44], the kth received sample is simply

$$r_k = s_k + n_k$$

with n_k being a complex Gaussian noise variable with zero mean and a variance of $\sigma^2 = N_0$.

2.3.2 Rayleigh Fading Channel

In mobile radio applications, transmitted signals experience different types of disturbance. The mobile radio signal is received while the mobile unit is in motion. Mobile site receives many reflected waves and one direct wave. All the reflected waves received at the mobile unit result in a multipath fading signal. It is also observed that when the operating frequency becomes higher, the fading signal becomes more severe. The average signal level of the fading signal decreases as the mobile unit moves away from the base-station transmitter. This drop in the average signal level is called propagation path loss.

Normally, the fading process can be considered having a Rayleigh distribution. The mobile channel can then be modeled as a Rayleigh fading channel with a multiplicative

disturbance having a Rayleigh pdf (probability density function). In this situation, let $s(t)$ be the baseband equivalent of the transmitted signal, the baseband equivalent of the received signal $r(t)$ is represented as [1, 46]:

$$r(t) = g(t)s(t) + n(t) \tag{2.9}$$

where $g(t)$ is a zero mean, complex Gaussian process and $n(t)$ is additive complex Gaussian noise with zero mean and a power spectral density of N_0. $g(t)$ is called the Rayleigh fading process with a normalized autocorrelation function

$$\rho_\tau = J_0(2\pi f_D \tau) \tag{2.10}$$

where $J_0(\cdot)$ is the modified Bessel function of the first kind and order zero, and f_D is the maximum Doppler frequency caused by the motion of the mobile unit. The parameter f_D can be expressed in terms of the vehicle speed v and the carrier frequency f_c as

$$f_D = \frac{v}{c} f_c$$

where c is the speed at which the electromagnetic waves propagate.

Similarly as in the AWGN channel, the output of the demodulator at the kth sampling instance can be expressed as:

$$r_k = g_k s_k + n_k$$

where g_k is a complex zero mean Gaussian variate with an autocorrelation function given by

$$\rho(m) = \frac{1}{2} E[g_k^* g_{k+m}] = J_0(2\pi m f_0 T).$$

The s_k's are the data symbols and the n_k's are the noise samples.

2.3.3 Digital Modulation Scheme

The pan-European GSM system uses constant amplitude modulation with a spectral efficiency of 1.35 bits/s/Hz. For the North American standard, linear modulation

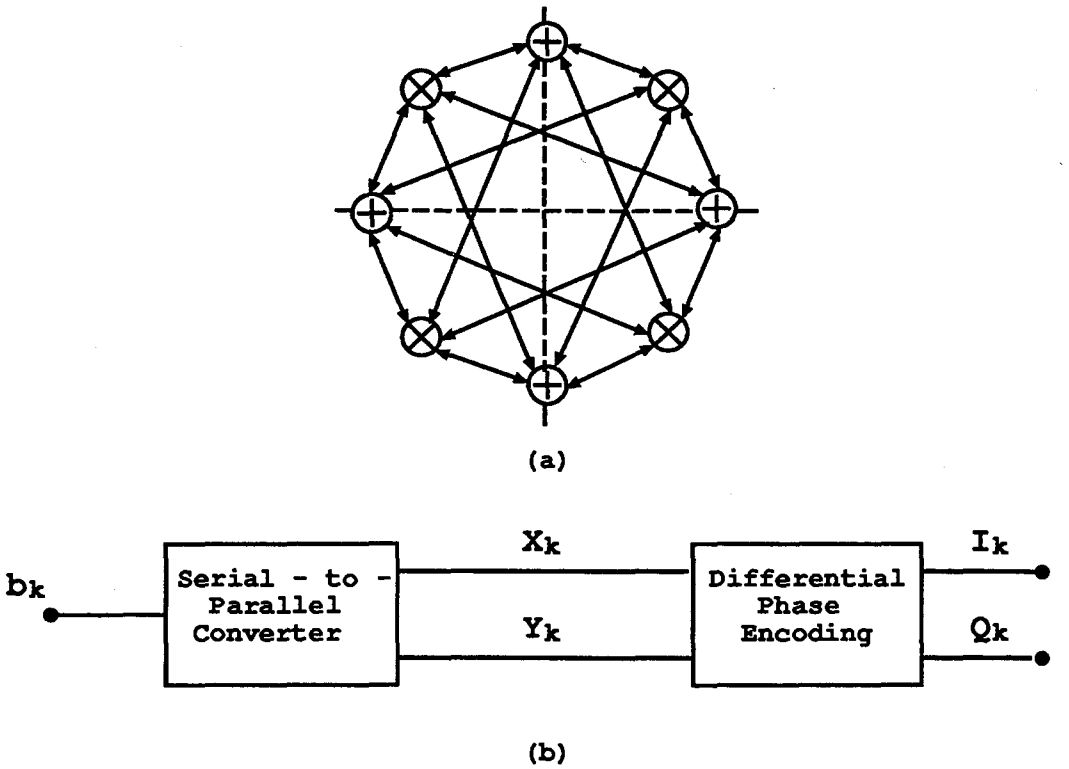


Figure 2.7: Configuration of the DQPSK digital modulation scheme

with a throughput of 1.62 bits/s/Hz has been chosen. The digital modulation method chosen for the new North American digital cellular system is a modified version of differential four phase shift keying with differentially coherent detection, known as $\frac{\pi}{4}$ shifted DQPSK.

The signal constellation for the $\frac{\pi}{4}$ shifted DQPSK modulation is shown in Fig. 2.7 (a), where an 8 signal point constellation is used with the two component QPSK constellations used in alternate symbol intervals. The symbols in the two component constellations are transmitted as changes in phase rather than absolute phase. This differential encoder is shown in Fig. 2.7 (b). The binary data stream entering the modulator is converted into two separate binary streams (X_k) and (Y_k). All odd numbered bits form stream X_k and all even numbered bits form stream Y_k. The

digital data sequences (X_k) and (Y_k) are encoded into (I_k) and (Q_k) according to:

$$I_k = I_{k-1}cos[\Delta\Phi_k(X_k, Y_k)] - Q_{k-1}sin[\Delta\Phi_k(X_k, Y_k)]$$

$$Q_k = I_{k-1}sin[\Delta\Phi_k(X_k, Y_k)] + Q_{k-1}cos[\Delta\Phi_k(X_k, Y_k)]$$

where k means the present time interval and $k-1$ represents the previous time interval. The phase change $\Delta\Phi_k$ is determined according to Table 2.2. Note that a Gray code is used in the mapping.

X_k	1	0	0	1
Y_k	1	1	0	0
$\Delta\Phi_k$	$-3\pi/4$	$3\pi/4$	$\pi/4$	$-\pi/4$

Table 2.2: $\frac{\pi}{4}$-DQPSK phase difference representation

At the receiver, a differential coherent detector is used to estimate X_k and Y_k based on the received phase difference.

2.4 Channel Coding

Channel coding is essentially a signal processing technique that is used to improve the reliability of digital communications by enabling the transmitted signals to better withstand the effect of various channel impairments, such as noise, fading, or jamming. This is achieved by introducing redundancy so that they can accentuate the uniqueness of each message and averaging the noise introduced by the physical channel.

A block diagram of the channel coding system is shown in Fig. 2.8, which is the dashed part in Fig. 2.2. The objective of channel coding is to reduce the bit error probability of the underlying channel at the cost of expanding the transmission

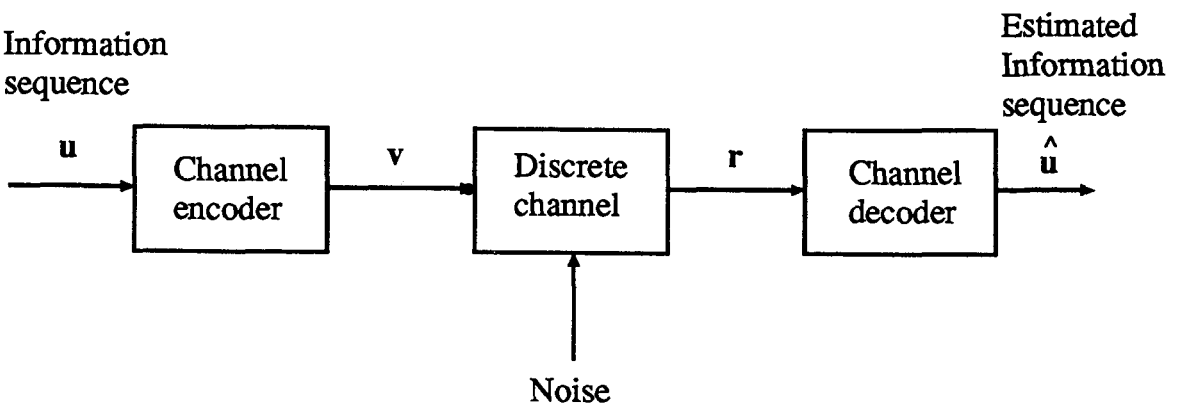


Figure 2.8: Block diagram of a channel coding system

bandwidth. The use of channel coding has grown tremendously in recently years. This is due to the fast progress of IC (integrated circuit) techniques which makes various complex channel decoding algorithms implementable in real time. The utilization of channel coding techniques for real time communications makes it possible to provide a large improvement of the communication system performance without the use of higher power transmitters which are much more costly.

There are basically two types of channel encoding scheme that provide forward error correction. The first is block encoding. In block encoding, the information sequence is divided into blocks of k information bits $\mathbf{u} = (u_1, u_2, ..., u_k)$, which is called a k-tuple. These k-tuples are then encoded by the block encoder into the encoded sequence of n-tuple $\mathbf{v} = (v_1, v_2, ..., v_n)$, where $n > k$. These n-tuples are called codewords. There are a total of 2^k different possible messages. Therefore, corresponding to the 2^k different possible messages, there are 2^k different possible codewords at the encoder output. The code rate is $R_c = k/n$, where R_c is a measure of the redundancy introduced by the encoder and n is the block length of the code. Since the n-symbol output codeword depends only on the corresponding k-bit information input, the encoder is memoryless from block to block. The important performance measurement of the block code is its minimum distance d, which is defined as the

minimum Hamming distance between codeword pairs. The minimum distance d of a block code directly determines the number of errors it can detect or correct. There are many important block codes. We will discuss in detail the class of BCH codes and its subset Reed-Solomon codes in Chapter 3.

Another type of encoding scheme is called convolutional coding. With convolutional codes, the encoded data does not have a simple block structure. Rather, the encoder operates on the source data stream using a "sliding window" to produce a continuous stream of encoded symbols. Each information symbol can affect a finite number of consecutive symbols in the output stream. The encoder of the convolutional code accepts k-bit blocks from the sequence $\mathbf{u}$ and produces an encoded sequence (codeword) of n-symbol blocks $\mathbf{v}$. Each encoded block depends not only on the current k-bit block, but also on v previous message blocks. Hence, the encoder has a memory of order v, which is usually defined as the constraint length of a convolutional code. The convolutional code produced by a k-bit input, n-bit output encoder of memory order v is call an (n, k, v) convolutional code. The ratio $R_c = k/n$ is again defined as the code rate. Many of the concepts in block codes, such as Hamming distance, minimum distance, parity check, linearity, and the syndrome, carry over to convolutional codes, although the definitions must be modified somewhat to account for the lack of a simple block structure. Details about convolutional codes will also be given in Chapter 3.

For both block codes and convolutional codes, there are two types of decoding algorithms, i.e., hard decision decoding and soft decision decoding. If the received data fed into the decoder is quantized into two levels as 0 or 1, the decoding process is called *hard decision*. If a measure of reliability is provided for each received bit as part of the demodulation function, the resulting decoding process is called *soft decision* decoding. In soft decision decoding, unlike that in hard decision decoding, the received data are unquantized, the decoder makes use of the additional information contained in the unquantized samples to recover the information with higher reliability. In practical

communication systems, we rarely have the luxury of being able to process the actual unquantized samples. The normal practice is to quantize the received analog samples and to make available a set of numbers which represent the quantization levels. For block codes, soft-decision decoders are substantially more complex than hard-decision decoders, and the complexity increases with the levels of quantization. Thus, block codes are usually implemented with hard decision decoding. For convolutional codes, with the maximum likelihood Viterbi decoding algorithm, soft decision decoding can be accomplished much easier than that with block codes. Therefore, both hard and soft decision decoding implementations are equally popular for convolutional codes. Normally, soft decision decoding offers an approximate 2 - 3 dB decoding gain over hard decision decoding.

2.5 Combined Source and Channel Coding

In most existing communication systems, the source codec and the channel codec are designed separately. When designing the source codec, the standard approach is to assume that ideal channel encoders and decoders are employed so that the link between the source encoder and decoder is noiseless. On the other hand, in designing channel codecs, attention is normally paid to that part of Fig. 2.2 consisting of the channel encoder, channel, and the channel decoder. This separation in designing the source and channel codecs is supported by Shannon's work on information theory [22] which indicates that the source and channel coding functions are fundamentally separable. Viterbi and Omura [35] also clearly stated that the assumption that source and channel encoders can be considered separately can be justified on the basis that, in the limit of arbitrarily complex overall encoders and decoders, no loss in performance results from separating source and channel coding in this way. Or in other words, as the dual facets of the same fundamental problem, source and channel coding can be separated in the way that the entropy rate reduction takes place in the source

encoder and the channel error protection in the channel encoder. The advantage of separating the design of source and channel codecs is that it allows channel codecs to be designed independently of the actual source statistics and user. In practical situations where there are always limitations on the complexity of the system, severe performance degradation is expected. Consequently, we choose the designing of the combined source and channel codec.

Generally, there are two approaches in designing combined source and channel coding system. One is to jointly optimize the source and channel coders. The other approach is to match existing source and channel coding schemes.

An example of the first approach has been presented by Kurtenbach and Wintz [47] in the design of an optimal uniform quantizer for transmission over a noisy channel. It is determined that the structures of the optimum quantizer, under the criterion of the mean-squared error(MSE), depend on the input data through its probability density function and on the channel through its transition matrix. Rydbeck and Sundberg addressed in [48] the issue of codebook indices assignment and found out that Gray code assignment to codebook indices results in a more robust source coding system to channel errors. The Gray ordering used in this manner is called protection without introducing redundancy. The Gray coding scheme performs well with a low channel noise level, however, when the channel bit error rate gets higher than 10^{-3}, the performance of the source coder degrades significantly. For acceptable system performance under this situation, error protection with redundancy introduced is necessary. Farvardin and Vaishampayan [32] have presented an analysis of the zero-memory quantization of a memoryless source when the quantizer output is to be encoded and transmitted across a noisy channel. They found the necessary conditions for the joint optimization of the quantizer and the encoder/decoder pair and developed an iterative algorithm for obtaining a locally optimum system.

Another application for joint optimization of the source and channel coding system

can be seen in the design of Vector Quantization (VQ) systems. Zeger and Gersho [12, 49] have studied the effect of transmission errors on the performance of VQ by incorporating a channel index assignment function into a source/channel model of the VQ. Gray coding algorithm, also referred to as the zero redundancy channel coding, was used to find the optimal assignment of a unique b-bit codeword to each of the 2^b code vectors in a VQ codebook. Marca et al. [14] and Kleijn [50] have applied simulated annealing techniques to improve the index assignment function of a vector quantizer on noisy channels. It is shown that about 4.5 dB SNR gain can be achieved over a random assignment.

In contrast, the second approach (matching existing source and channel coding schemes) is more practical for more complex source coders. In the study of an error protected 16 kbits/s voice transmission for land mobile radio channel, Suda and Miki [18] presented a design procedure for their bit selective forward error correction (BS-FEC) scheme. Goodman and Sundberg [17] conducted a study on combined source and channel coding for variable-bit-rate speech transmission. In their study, the speech coder was an embedded DPCM coder and punctured convolutional codes were used for error protection. The rate assignment between source and channel coding can be changed in response to the changing transmission quality for a given transmission rate. Back in 1979, Modestino and Daut [33] presented a combined source-channel coding system for still images. The source encoder employed two-dimensional (2-D) differential pulse code modulation (DPCM). By providing error protection to those encoded bits which contribute more significantly to the quality of the reconstructed image, they achieved the minimum degradation without sacrificing transmission bandwidth.

In our study, we employ the CELP coder as the source coder. As mentioned in the introductory chapter, CELP coding is a promising technique for synthesizing high quality speech at low bit rates. The structure of a CELP coder is much more complex than the speech coders mentioned above. Thus, it is not feasible to use the

joint optimization stated above. In this study, we use the second approach to find the matching channel coders to protect the speech elements from transmission errors. The channel codes used are punctured Reed-Solomon codes and punctured convolutional codes.

CHAPTER 3

Error Correction Scheme

For most low rate speech coders, quality degrades significantly due to transmission errors. As discussed briefly in Chapter 1, there are several ways of improving the robustness of the speech transmission, such as efficient index assignment, parameter smoothing with error detection, and Forward Error Correction [15, 16, 21]. Efficient index assignment and parameter smoothing with error detection are only suitable for the case when the channel is assumed to only introduce a single error to a binary code representing a speech parameter. In mobile radio applications, the physical communication channel is characterized by severe disturbances. In this case a more powerful protection scheme has to be employed so as to efficiently mitigate the effect of errors.

Forward Error Correction is used in this thesis as the channel protection scheme. For mobile radio applications, the channel can be accurately modeled as a Rayleigh fading channel [46, 21]. Such a channel introduces bursty errors. As such, two different error protection methods may be employed. One is to use interleaving to spread out the bursty errors in a nearly random fashion, and then use a good random-error-correcting code, such as a convolutional code, as the forward error protection code.

The other is the use of the burst-error-correcting code such as Reed-Solomon (RS) codes.

In this chapter, we discuss in detail the channel coders we used in our combined system. The basic channel coding theorem is outlined in Section 3.1. The Reed-Solomon codes will be discussed in Section 3.2. Section 3.3 focuses on the convolutional codes. Rate-compatible punctured Reed-Solomon codes and rate-compatible punctured convolutional codes and their performance in Rayleigh fading channels will be discussed in Section 3.4. As mentioned in Chapter 2, these are the FEC codes used in this study.

3.1 Channel Coding Fundamentals

When digital data are transmitted over a noisy channel, there is always a chance that the received data will contain errors. Error-correction coding is being used on an almost routine basis in most communication systems. The utility of coding was demonstrated by the work of Shannon in 1948, which states that if the data source rate is less than a quantity called the channel capacity, communication over a noisy channel with an error probability as small as desired is possible. Essentially, Shannon's work states that signal power, channel noise, and available bandwidth set a limit only on the communication rate and not on the accuracy.

Since Shannon's pioneering work, many important error-control coding techniques have been developed. Error-correction coding is essentially a signal processing technique that is used to improve the reliability of communication on digital channels. Although individual coding schemes take on many different forms and have their roots in diverse mathematical disciplines, they all have two common ingredients. One is the use of redundancy. Coded digital information always has extra symbols other

than the original information symbols. These redundant symbols are used to accentuate the uniqueness of each message. They are always chosen so as to make it very unlikely that the channel noise will corrupt enough of the symbols in a message to destroy its uniqueness. The second ingredient is noise averaging. This averaging effect is obtained by making the redundant symbol depend on a span of several information symbols.

Generally, there are two kinds of codes. They are *block codes* and *tree codes*, depending on whether their encoders have *memory* or not. Conceptually, the encoder for a block code is a *memoryless* device that maps a k-symbol input sequence into an n-symbol output sequence. The k-symbol input sequence is the information sequence and the n-symbol output sequence is the encoded channel codeword with $(n - k)$ symbols of redundancy added. The term "memoryless" indicates that each n-symbol block is purely determined by a specific k-symbol block. The usual distinguishing parameters for a block code are the codeword blocklength n, the number of information symbols k, the coding rate $R = k/n$, and the minimum Hamming distance d_{min}. The latter parameter determines the number of correctable symbol errors. The encoder for a tree code is a device with memory. It accepts binary symbols in sets of m and outputs binary symbols in sets of n. Each set of n output symbols is determined by the current input set and a span of v of the preceding input symbols. The memory span of the encoder is then $v + m$ input symbols. The constraint length is defined as v in our study, according to Forney [51]. Tree codes are also characterized by parameters such as the coding rate $R = m/n$ and a minimum free distance d_{free}. Linear tree codes are usually referred to as the *convolutional codes*.

Codes can also be classified as either random-error-correcting codes or burst-error-correcting codes. In the next two sections, we will describe in more detail the Reed-Solomon codes which is a class of burst-error-correcting block codes and the random-error-correcting convolutional codes.

3.2 Reed-Solomon Codes

According to the mobile radio channel condition, the errors introduced by the channel usually appear in bursts. To apply suitable error protection to the speech signal transmitted over the channel, one approach is to utilize a powerful multi-error-correcting code.

3.2.1 Properties of Reed-Solomon Codes

Reed-Solomon codes are known as a subclass of Bose-Chaudhuri-Hocquenghem codes (BCH codes). The class of BCH codes is a large class of multiple-error-correcting codes that occupies a prominent place in the theory and practice of error correction. This prominence is a result of at least four reasons [27]: (1). Provided that the block length is not excessive, there are good codes in this class. (2). Relatively simple and instrumentable encoding and decoding techniques are known. (3). The popular nonbinary subclass of Reed-Solomon codes has certain optimality properties and a well-understood distance structure. (4). A thorough understanding of BCH codes is probably the best single starting point for studying many other classes of codes.

Since BCH codes (including Reed-Solomon codes) belong to cyclic code, they are easy to encode and decode and have a well-defined algebraic structure. BCH codes include codes defined on both binary and nonbinary symbol alphabets. Also the generator polynomial of BCH code is well related to the Hamming distance d. We can easily construct BCH codes according to the requirement for d. Both the encoding and decoding circuits are simpler for engineering implementation. With medium and short block lengths, the properties of BCH codes are very close to the optimal values.

Reed-Solomon codes operate on symbols in $GF(q)$ (Galois Field). These are BCH

codes in which the block length divides the multiplicative order of the symbol alphabet, that is, they are BCH codes with block length $n = q - 1$. For a t-error-correcting RS code, the generator polynomial is :

$$g(x) = (x - \alpha)(x - \alpha^2).....(x - \alpha^{2t}) \tag{3.1}$$

where α is the primitive element of $GF(q)$. $g(x)$ is always a polynomial of degree $2t$. Hence, a Reed-Solomon code satisfies $n - k = 2t$.

The RS codes are optimal in the sense of the Singleton Bound. The Singleton Bound states that for any linear code

$$d^* \leq n - k + 1$$

where d^* is the actual minimum distance. Since $d^* = 2t + 1 = n - k + 1$ for RS codes, we have $d^* = d$, which is the designed distance. Consequently, Reed-Solomon codes are maximum-distance-separable (MDS) codes. In other words, for a fixed codeword (n, k), no code can have a larger minimum distance than a RS code. This is often a strong justification for using RS codes. RS codes always have relatively short block length as compared to other cyclic codes over the same alphabet.

Generally, the parity-check matrix of a RS code (n, k, d) on $GF(q)$ is :

$$H = \begin{bmatrix} \alpha^{n-1} & \alpha^{n-2} & \dots & \alpha & 1 \\ (\alpha^2)^{n-1} & (\alpha^2)^{n-2} & \dots & \alpha^2 & 1 \\ \dots\dots\dots\dots & & & & \\ (\alpha^{d-1})^{n-1} & (\alpha^{d-1})^{n-2} & \dots & \alpha^{d-1} & 1 \end{bmatrix} \tag{3.2}$$

where α is the primitive element of $GF(q)$. In practice, $q = 2^m$, and consequently $n = 2^m$ -1 and $k = 2^m - d$. This means that if each symbol in $GF(2^m)$ is represented as a binary m-tuple, the size of H is $d \times (n - 1)$. In other words, the RS code can be treated as a $(m(2^m - 1), m(2^m - d), d)$ binary code. For example, the (7, 3, 5) RS code on $GF(2^3)$ can correct any two 8-ary random errors in the codeword, while the

corresponding (21, 9, 5) binary code can correct two arbitrary random errors, burst errors with length less than or equal to 4 and also many other error patterns. This method of translating 2^m-ary RS codes into binary codes is called the mapping from $GF(2^m)$ to $GF(2)$. Although the original code is cyclic, the binary code obtained through the mapping is not necessarily cyclic.

3.2.2 Decoding Algorithm for RS codes

RS codes are cyclic codes and can be decoded by any technique for decoding cyclic codes. Here we present the method of syndrome decoding. For a block code, suppose that the parity check matrix is H and the received codeword is R, then the syndrome S is obtained by:

$$S = RH^T$$

Details about syndrome decoding for block codes can be found in [26, 27].

The following is the procedure for syndrome decoding of a nonbinary code:

1. Calculate the syndrome values S_k, $k = 0, 1, ..., 2t-1$.
2. Determine the error-locator polynomial $\sigma(x)$ from syndrome values S_k.
3. Solve for the roots of $\sigma(x)$, which are the error locators.
4. Given the error locators, calculate the error values.

For a t-error-correcting RS code, the parity check matrix in Eq. 3.2 can be written as:

$$H = \begin{bmatrix} \alpha^{n-1} & \alpha^{n-2} & \dots & \alpha & 1 \\ (\alpha^2)^{n-1} & (\alpha^2)^{n-2} & \dots & \alpha^2 & 1 \\ \dots\dots\dots\dots & & & & \\ (\alpha^{2t})^{n-1} & (\alpha^{2t})^{n-2} & \dots & \alpha^{2t} & 1 \end{bmatrix} \tag{3.3}$$

where α is the primitive element of $GF(q)$. The error polynomial is:

$$e(x) = e_{n-1}x^{n-1} + e_{n-2}x^{n-2} + ... + e_1x + e_0$$

We assume here that at most t coefficients are nonzero. Suppose that v errors actually occur. $0 \leq v \leq t$, and that they occur in unknown locations i_1, i_2,..., i_v. The error polynomial can be written as:

$$e(x) = e_{i_1}x^{i_1} + e_{i_2}x^{i_2} + ... + e_{i_v}x^{i_v}$$

where e_{i_l} is the magnitude of the lth error ($e_{i_l} = 1$ for binary codes). We now want to solve for v, i_1,...,i_v, e_{i_1},..., e_{i_v}. We denote the received codeword R and the actual transmitted codeword C. Evaluate the received polynomial at α^j to obtain the syndrome S_j:

$$\begin{aligned} S_j = RH^T = R(\alpha^j) = C(\alpha^j) + e(\alpha^j) = e(\alpha^j) \\ = e_{i_1}(\alpha^j)^{i_1} + e_{i_2}(\alpha^j)^{i_2} + ... + e_{i_v}(\alpha^j)^{i_v} \\ j = 1, 2, ..., 2t \end{aligned}$$

To streamline it, we define the error magnitude $Y_l = e_{i_l}$ for $l = 1, ..., v$, and the error-location numbers $X_l = \alpha^{i_l}$ for $l = 1, ..., v$, where i_l is the actual location of the lth error and X_l is the field element associated with this location. Then we have the following set of $2t$ simultaneous equations in the v unknown error locations X_1,...,X_v and the v unknown error magnitudes Y_1,...,Y_v.

$$\begin{cases} S_1 = Y_1X_1 + Y_2X_2 + ... + Y_vX_v \\ S_2 = Y_1X_1^2 + Y_2X_2^2 + ... + Y_vX_v^2 \\ S_3 = Y_1X_1^3 + Y_2X_2^3 + ... + Y_vX_v^3 \\ ..... \\ S_{2t} = Y_1X_1^{2t} + Y_2X_2^{2t} + ... + Y_vX_v^{2t} \end{cases} \tag{3.4}$$

This set of equations must have at least one solution because of the way in which the syndromes are defined. The set of nonlinear equations is too difficult to solve directly. We introduce:

$$\sigma(x) = \sigma_v x^v + \sigma_{v-1}x^{v-1} + ... + \sigma_1 x + 1 \tag{3.5}$$

known as the error-locator polynomial and defined to be the polynomial with zeros at the inverse error locations X_l^{-1} for $l = 1, ..., v$. That is,

$$\sigma(x) = (1 - xX_1)(1 - xX_2)...(1 - xX_v)$$

3.2.2.1 Finding the Error-Locator Polynomial $\sigma(x)$

The main task in decoding the RS codes lies in finding the error-locator polynomial $\sigma(x)$. There are several ways of solving this problem, among which the Massey-Berlekamp Algorithm [26] is the simplest and mostly frequently used .

Let's look at the error-locator polynomial shown in Eq. 3.5. Multiply both sides of the Eq. 3.5 by $Y_l X_l^{j+v}$ and set $x = X_l^{-1}$, then the left hand side is zero, and we have:

$$0 = Y_l X_l^{j+v}(1 + \sigma_1 X_l^{-1} + ... + \sigma_v X_l^{-v})$$

or

$$\begin{aligned} Y_l(X_l^{j+v} + \sigma_1 X_l^{j+v-1} + ... + \sigma_v X_l^j) = 0 \\ j = 1, 2, ..., 2t; \\ l = 1, 2, ..., v \end{aligned} \tag{3.6}$$

Sum up these equations in Eq. 3.6 from $l = 1$ to $l = v$. This gives, for each j,

$$\sum_{l=1}^{v} Y_l(X_l^{j+v} + \sigma_1 X_l^{j+v-1} + ... + \sigma_v X_l^j) = 0$$

or

$$\sum_{l=1}^{v} Y_l X_l^{j+v} + \sigma_1 \sum_{l=1}^{v} Y_l X_l^{j+v-1} + ... + \sigma_v \sum_{l=1}^{v} Y_l X_l^j = 0 \tag{3.7}$$

The individual sums are recognized as syndromes, then Eq. 3.7 becomes:

$$S_{j+v} + \sigma_1 S_{j+v-1} + \sigma_2 S_{j+v-2} + ... + \sigma_v S_j = 0$$

Because $v \leq t$, the subscripts always specify known syndromes if j is in the interval $1 \leq j \leq v$. Hence, we have the set of equations

$$\begin{aligned} \sigma_1 S_{j+v-1} + \sigma_2 S_{j+v-2} + ... + \sigma_v S_j &= -S_{j+v} \\ j &= 1, ..., v \end{aligned} \tag{3.8}$$

Writing these equations in matrix form gives:

$$\begin{bmatrix} S_1 & S_2 & ... & S_{v-1} & S_v \\ S_2 & S_3 & ... & S_v & S_{v+1} \\ ..... & & & & \\ S_v & S_{v+1} & ... & S_{2v-2} & S_{2v-1} \end{bmatrix} \begin{pmatrix} \sigma_v \\ \sigma_{v-1} \\ ... \\ \sigma_1 \end{pmatrix} = \begin{pmatrix} -S_{v+1} \\ -S_{v+2} \\ ... \\ -S_{2v} \end{pmatrix} \tag{3.9}$$

$$[S][\sigma] = [-S'] \tag{3.10}$$

Here $[S]$ is the syndrome matrix and $[-S']$ is simply the right-hand side of Eq. 3.9

At this point, it becomes clear that if v is known and if the matrix $[S]$ is nonsingular, then the vector $[\sigma]$ or the error locations are known. The following theorem is invoked when the decoder tries to determine $\sigma(x)$.

Theorem: The matrix of syndromes

$$M = \begin{bmatrix} S_1 & S_2 & ... & Su \\ S_2 & S_3 & ... & S_{u+1} \\ ..... & & & \\ S_u & S_{u+1} & ... & S_{2u-1} \end{bmatrix}$$

is nonsingular if u is equal to v, the number of errors that actually occurred. The matrix is singular if u is greater than v.

A proof of this is shown in Appendix A.

The above theorem provides the basis of the decoding algorithm. First find the correct value of v as follows. As a trial value, set $v = t$. If $[S]$ has full rank, the actual

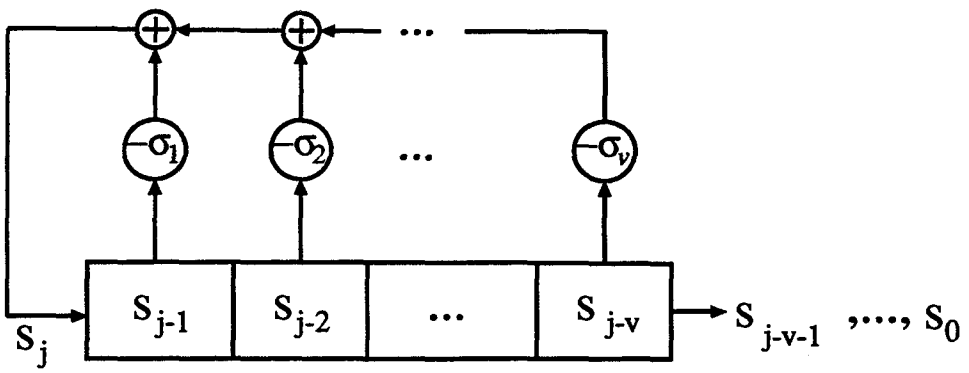


Figure 3.1: The Linear Feedback Shift Register

number of error is t. If the determinant of $[S]$ is zero, reduce the trial value of v by 1 and repeat. Continue in this way until a nonzero determinant is obtained. The actual number of errors that occurred is then known. Then we can get the error locators in Eq. 3.10 by solving

$$[\sigma] = [S]^{-1}[-S']$$

This algorithm is called the Peterson-Gorenstein-Zierler algorithm. It requires that a t by t matrix be inverted in computing $\sigma(x)$. The number of computations necessary is proportional to v^3. The high complexity involved in this inversion can be circumvented by using the Massey-Berlekamp algorithm.

The Massey-Berlekamp algorithm relies on the fact that the matrix equation in Eq. 3.9 is not arbitrary in its form, rather it is highly structured. This structure is used to advantage in obtaining the vector σ by a method that is conceptually more intricate than direct matrix inversion, but computationally much simpler. For those readers who are familiar with linear prediction theory, Eq. 3.9 is similar to that used to obtain the LPC coefficients that model the underlying random process. Consequently, determining the coefficients of the error-locator polynomial from S_k's is equivalent to synthesizing a linear feedback shift register (LFSR) with minimum length that generates the given S_k's, see Fig. 3.1, where $\sigma(x) = \sigma_v x^v + \sigma_{v-1}x^{v-1} + ... + \sigma_1 x + 1$ is

called the connection polynomial.

For a given sequence of syndrome values, there is a determinable number of connection polynomials of various lengths that will generate the syndromes. With bounded-distance decoding, the task is to find the lowest weight connection polynomial. The iterative procedure used to find such a polynomial is outlined below without proof. It begins by postulating the shortest possible shift register and then attempts to generate the entire sequence of given syndrome values in order. The actual syndrome sequence is continually compared with the output of the postulated LFSR until either the entire sequence of given syndrome values is reproduced or a discrepancy is encountered. The steps in this iterative procedure are the following [27].

1. Compute the syndrome values $S_n, 1 \leq n \leq d-1$

2. Initialize algorithm variables

   $\sigma(x) = 1 \qquad D(x) = x \quad$ (D(x) is a correction term)

   $l = 0 \qquad n = 1$

3. Take in a new syndrome value and compute discrepancy

$$\delta = S_n + \sum_{i=1}^{l} \sigma_i S_{n-i}$$

4. Test discrepancy

   If $\delta = 0$, go to step 9. Otherwise, go to step 5.

5. Modify connection polynomial

   Let $\sigma^*(x) = \sigma(x) - \delta D(x)$

6. Test register length

   If $2l \geq n$, got to step 8 (i.e. do not extend register). Otherwise, go to step 7

7. Change register length and update correction term

   Let $l = n - l$, and $D(x) = \sigma(x)/\delta$

8. Update connection polynomial
   Let $\sigma(x) = \sigma^*(x)$

9. Update correction term
   Let $D(x) = xD(x)$

10. Update syndrome counter
    Let $n = n + 1$

11. Test syndrome count
    If $n < d$, go to step 3. Otherwise, stop.

By following all the steps above, we will end up with the error-locator polynomial:

$$\sigma(x) = \sigma_v x^v + \sigma_{v-1} x^{v-1} + ... + \sigma_1 x + 1$$

3.2.2.2 Chien Search

After finding the error-locator polynomial $\sigma(x)$, the next task is to find the roots of this polynomial which are the reciprocals of the error locations. Since there are usually only a finite number of field elements to check, the simplest way to find the zeros of $\sigma(x)$ is by trial and error, a process known as a Chien Search. One simply computes in turn $\sigma(\alpha^j)$ for each j and checks for zero. A computationally simpler way to evaluate the polynomial $\sigma(x)$ at β is by Horner's rule:

$$\sigma(\beta) = (...(((\sigma_v\beta + \sigma_{v-1})\beta + \sigma_{v-2})\beta + \sigma_{v-3})\beta + ... + \sigma_0)$$

Horner's rule needs only v multiplications and v additions to compute $\sigma(\beta)$.

3.2.2.3 Computation of the Error Magnitudes

If the code we deal with is nonbinary, we have to compute the error magnitude from the error locations. There are two ways to do so. One is the direct matrix inversion method. The other is known as the Fast Forney algorithm.

From Eq. 3.4, we have

$$\begin{bmatrix} X_1 & X_2 & \dots & X_v \\ X_1^2 & X_2^2 & \dots & X_v^2 \\ \dots\dots\dots & & & \\ X_1^v & X_2^v & \dots & X_v^v \end{bmatrix} \begin{pmatrix} Y_1 \\ Y_2 \\ \dots \\ Y_v \end{pmatrix} = \begin{pmatrix} S_1 \\ S_2 \\ \dots \\ S_v \end{pmatrix}$$

Then, the error magnitude Y_l can be found by

$$[Y] = [X^{-1}][S]$$

The Forney algorithm starts with the error-locator polynomial which was defined to have zeros at X_l^{-1} for $l = 1, ..., v$, i.e.:

$$\sigma(x) = \prod_{l=1}^{v}(1 - xX_l)$$

Define the syndrome polynomial

$$S(x) = \sum_{j=1}^{2t} S_j x^{j-1} = \sum_{j=1}^{2t}\sum_{i=1}^{v} Y_i X_i^j x^{j-1},$$

and define the error-evaluator polynomial $\Omega(x)$ in terms of these known polynomials:

$$\Omega(x) = S(x)\sigma(x) \qquad (mod\ x^{2t})$$

The error-evaluator polynomial will play a minor role from time to time. It can be related to the error-locations and error magnitudes as follows.

Theorem : The error-evaluator polynomial can be written as:

$$\Omega(x) = \sum_{i=1}^{v} Y_i X_i \prod_{l \neq i}(1 - X_l x).$$

Theorem: (The Forney Algorithm) The error magnitudes are given by

$$Y_l = (X_l^{-1}\Omega(X_l^{-1}))/(\prod_{j\neq l}(1 - X_jX_l^{-1})) = -(\Omega(X_l^{-1}))/(\sigma'(X_l^{-1}))$$

where $\sigma'(x)$ denotes the derivative of $\sigma(x)$.

The proofs of these theorems can be found in [27].

3.3 Convolutional Codes

As mentioned at the beginning of this chapter, good random-error-correcting codes can also be utilized together with interleaving to combat the error bursts in the mobile communication channel. Convolutional codes are well known as good random error correcting codes. We have also employed convolutional codes with Viterbi decoding algorithm in this thesis.

With *convolutional* codes, which were originally called the *recurrent* codes, the encoded data does not have a simple block structure as in the case for the Reed-Solomon codes. Rather, the encoder for a convolutional code operates on the source data stream using a "sliding window" and produces a continuous stream of encoded symbols, usually binary. Each information symbol in turn can affect a finite number of consecutive symbols in the output stream.

A simple encoder for a rate 1/2 convolutional code is illustrated in Fig. 3.2. Convolutional codes can be either systematic or nonsystematic. For this systematic rate 1/2 code in Fig. 3.2, each information symbol to be transmitted is associated with two channel symbols, where one is the information symbol itself and the other is the parity symbol, given by

$$p_j = \sum_{l=0}^{k-1} G_l i_{j-l}, \qquad j = 1, 2, 3, ...$$

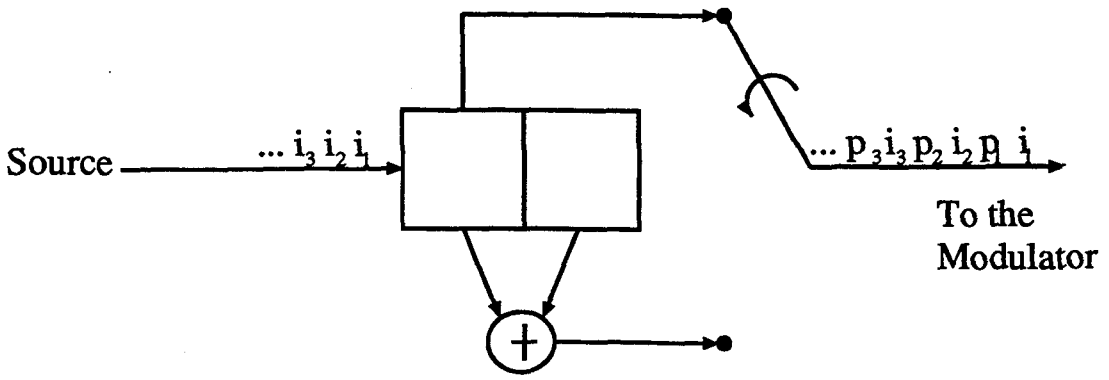


Figure 3.2: Encoder for a simple convolutional code

where $G_0 = 1$, $G_1 = 1$ are the generator coefficients, and $k = 1$.

In general, a code is completely defined by the set of coefficients G_l, $l = 0, 1, ..., k-1$. Any convolutional code can be represented by its generator matrix. The generator matrix of the code shown in Fig. 3.2 is simply:

$$G(D) = \begin{bmatrix} G_1(D) \\ G_2(D) \end{bmatrix} = \begin{bmatrix} D \\ 1 + D + D^2 \end{bmatrix}$$

Our discussion of the convolutional code structure uses the notation proposed by Forney [51]. The code constraint length is defined to be the number of memory elements, denoted by v, which is 2 in Fig. 3.2.

There are many decoding methods for convolutional codes, such as Viterbi decoding, syndrome decoding, and sequential decoding, etc.. Convolutional coding with Viterbi decoding has become one of the most widely used forward-error-correction techniques, due to both the simplicity of implementation and the relatively large coding gain. The achievement of such coding gains results principally from the ease with which this algorithm can utilize demodulator soft decision and thereby provide approximately 2 - 3 dB gain over hard decision decoding. The Viterbi algorithm is derived from the finite-state machine (FSM) nature of a convolutional encoder [26].

3.4 Rate-compatible Punctured Reed-Solomon Codes and Rate-Compatible Punctured Convolutional Codes

The design of an error protection scheme usually consists of selecting a fixed channel code with a certain rate, complexity, and correction capability that is uniform for all the data to be transmitted. However, in many cases, the data to be transmitted have different levels of protection requirement, such as the speech data in this study. Under this requirement, the idea of punctured codes is utilized.

The idea of punctured codes is well described by Clark et al. [26], where it is suggested that higher rate convolutional codes can be constructed from lower rate "mother codes" by periodically puncturing (deleting) symbols from the mother code. Since puncturing does not change the trellis structure, the same decoder can be used for the punctured code and the mother code. Punctured convolutional codes are used in this study to provide unequal error protection because of their implementation simplicity. Another class of punctured codes considered are the punctured RS codes obtained from the Chinese remainder theorem. Unlike conventional RS codes, punctured RS codes have flexible word lengths. The distance property however is the same as the mother code. Like punctured convolutional codes, a single encoder-decoder pair can be used to decode the punctured and the mother codes. In the next two subsections, we describe these codes in detail.

Both punctured RS codes and convolutional codes can be encoded and decoded with one encoder-decoder pair. Thus they make it possible to apply unequal error protection to the speech output without the use of multiple encoder and decoder pairs. By use of either of these two FEC codes, unequal error protection can be accomplished without increasing the complexity of the combined speech and channel coding system.

We will discuss the codeword properties and the encoding and decoding algorithm for both punctured RS codes and rate-compatible punctured convolutional codes. Their bit error performance on Rayleigh fading channel will also be presented.

3.4.1 Punctured Reed-Solomon Codes

The idea of constructing multiple error correcting codes with the Chinese remainder theorem was first introduced by J. J. Stone [28] in his paper in 1963. D. Mandelbaum [29] introduced a simpler operation in the detection of errors in 1968. Bossen and Yau [54] also published a paper in 1968 on this subject. In 1971, Mandelbaum [30] showed that Reed-Solomon codes encoded by means of the Chinese remainder theorem can also be decoded using the Berlekamp algorithm.

3.4.1.1 The Fundamentals of Encoding Multiple Error Correcting Codes Via the Chinese Remainder Theorem

According to Stone [28], the Chinese remainder theorem of number theory describes conditions under which a number may be recaptured by the knowledge of its residues when divided by certain moduli. Thus under conditions which allow recapturing, a number can be communicated from a sender to a receiver by the transmission of its residues. If additional residues were sent, the number might be communicated despite some distortion in the transmission. If we extend the Chinese remainder theorem to deal with polynomials over the Galois fields $GF(q)$, then codes can be constructed to combat multiple errors.

Theorem 1 (Chinese Remainder Theorem): Let $m_1, m_2, ..., m_r$ be integers which are relatively prime in pairs. Let $M = \Pi_{i=1}^{r} m_i$ denote their product. If $a_1, a_2, ..., a_r$ are any given integers, there exists one and only one number f such that $0 \leq f < M$

and

$$f \equiv a_i \qquad (mod\ m_i) \qquad i = 1, 2, ..., r.$$

Extending this theorem from numbers to Galois fields, we get Theorem 2:

Theorem 2: Let $m_1(x)$, $m_2(x)$, ..., $m_r(x)$ be in $GF(q)[x]$ (polynomials with coefficients in $GF(q)$) and relatively prime in pairs. Let $M(x) = \Pi_{i=1}^{r} m_i(x)$ denote their product. If $a_1(x)$, $a_2(x)$, ..., $a_r(x)$ are any given polynomials of $GF(q)[x]$, then there exists one and only one polynomial $f(x)$ of $GF(q)[x]$ such that the degree of $f(x)$ is less than the degree of $M(x)$ and

$$f(x) \equiv a_i(x) \qquad mod\ m_i(x).$$

Theorem 2 is the foundation of a method for designing multiple error correction codes in a way that is different from using the generator polynomial. Let $m_0(x), m_1(x), ..., m_{h-1}(x)$ be pairwise relatively prime polynomials over $GF(q)$, each of degree c. Any polynomial $f'(x)$ with coefficients in $GF(q)$ and having degree less than ch can be encoded uniquely by a sequence $A(x)$ of h residues:

$$A(x) = \{a_0(x), a_1(x), ..., a_{h-1}(x)\} \tag{3.11}$$

where $a_i(x) \equiv f'(x)\ mod\ m_i(x)$. $f'(x)$ will be called the *information polynomial*, and $a_i(x)$ is the residue modulo $m_i(x)$. The degree of $f'(x)$ must be less than the degree of $M(x)$, where

$$M(x) = \prod_{i=0}^{h-1} m_i(x)$$

Because the $m_i(x)$ are relatively prime, there must exist a unique polynomial $z_i(x)$ such that

$$(M(x)/m_i(x))z_i(x) \equiv 1 \qquad mod\ m_i(x) \tag{3.12}$$

for all i since $M(x)/m_i(x)$ and $m_i(x)$ are relatively prime. This results in the same conclusion that

$$(M(x)/m_i(x))z_i(x)a_i(x) \equiv a_i(x) \qquad mod\ m_i(x) \tag{3.13}$$

As a result of the above, $f^{'}(x)$ can be written as:

$$f^{'}(x) = \sum_{i=0}^{h-1}(M(x)/m_i(x))|z_i(x)a_i(x)|_{m_i(x)} \tag{3.14}$$

where $|z_i(x)a_i(x)|_{m_i(x)}$ stands for $z_i(x)a_i(x)$ mod $m_i(x)$. This sum is unique by the Chinese remainder theorem since $f^{'}(x)$ written in the above way has degree less than that of $M(x)$.

From the channel coding point of view, the residues $a_i(x)$ can be treated as a code-word whose coefficients are transmitted in binary form. The recapture of $f^{'}(x)$ shown in Eq. 3.14 is the same as the decoding operation at the receiver. The redundancy is defined as $ch - 1 - deg(f^{'}(x))$ symbols. Thus if the redundancy is $2t$ residues or $2ct$ symbols over $GF(q)$, random errors in t residues can be corrected. The distance of this code is $2t+1$. We will then talk about the decoding method of codes constructed in this way.

Let's use the information polynomial:

$$f(x) = x^{2t}f^{'}(x) \tag{3.15}$$

There is no information in the lower $2ct$ symbols of $f(x)$. Given a received word (sequence of residues) without errors, the original information polynomial can be recaptured by

$$f(x) = \sum_{i=0}^{h-1}(M(x)/m_i(x))|z_i(x)a_i(x)|_{m_i(x)} \tag{3.16}$$

If t errors have occurred in transmission, the received sequence of residues will have the form:

$$V(x) = A(x) + 0, ..., e_{i_1}(x), 0, ..., 0, e_{i_2}(x), 0, e_{i_t}(x), 0, ..., 0$$

Performing the operation on $V(x)$, we obtain

$$v(x) = f(x) + (M(x)/m_{i_1}(x))y_{i_1}(x) + (M(x)/m_{i_2}(x))y_{i_2}(x) + ... \tag{3.17}$$
$$+(M(x)/m_{i_t}(x))y_{i_t}(x)$$

where $(M(x)/m_{i_k}(x))y_{i_k}(x) \equiv e_{i_k}(x) \ mod \ m_{i_k}(x)$ and $deg(y_{i_k}(x)) < deg(m_{i_k}(x))$

The error terms are given by

$$S(x) = v(x) - f(x) \tag{3.18}$$

Since $f(x)$ has zero coefficients at its lowest $2ct$ positions, $S(x)$ can be written as:

$$S(x) = s_0 + s_1 x + s_2 x^2 + ... + s_{2ct-1} x^{2ct-1} + ... \tag{3.19}$$

where

$$s_i = v_i, \qquad 0 \leq i \leq 2ct - 1$$

v_i are the coefficients of $v(x)$. We will use these encoding and decoding methods in constructing RS codes.

3.4.1.2 Construction of (Punctured) Reed-Solomon Codes

Reed-Solomon codes can be constructed by means of the Chinese remainder theorem, if

$$m_0(x) = x - 1, m_1(x) = x - \beta, ..., m_{h-1}(x) = x - \beta^{q^n-2},$$

where $h = q^n - 1$ and β is a primitive element of $GF(q^n)$. It is well-known that

$$M(x) = \prod_{i=0}^{h-1} m_i(x) = x^h - 1$$

From the decoding procedure mentioned above, the error term will be:

$$S(x) = \sum_{j=1}^{t} (M(x)/m_i(x))y_i(x) \tag{3.20}$$

where $y_i(x)$ is a member of $GF(q^n)$. It is seen from the above equation that each term of the right-hand side has non-zero value for only one non-zero element of $GF(q^n)$. Then, we can get:

$$\sum_{j=1}^{t} y_j/(x - \beta^k) = S(x)/(x^h - 1) \tag{3.21}$$

where β^k is an element of $GF(q^n)$. Using Taylor series,

$$1/(x^h - 1) = -(1 + x^h + x^{2h} + ...)$$

then, Eq. 3.21 can be written as:

$$\sum_{j=1}^{t} y_j/(x - \beta^k) = -S(x) - S(x)x^h - S(x)x^{2h} - ... \tag{3.22}$$

The $2t$ lowest order symbols of the right-hand side of Eq. 3.22 are known as:

$$\sum_{j=1}^{t} y_j/(x - \beta^k) = -(s_0 + s_1x + s_2x^2 + ... + s_{2t-1}x^{2t-1} + ...) \tag{3.23}$$

We write the left-hand side of Eq. 3.23 as:

$$\sum_{j=1}^{t} \beta^{-k} y_j/(-1 + \beta^{-k}x)$$

Representing β^{-k} with w_j, Eq. 3.23 becomes

$$\sum_{j=1}^{t} w_j y_j/(1 - w_jx) = s_0 + s_1x + s_2x^2 + ... \tag{3.24}$$

The left-hand side of Eq. 3.24 is the partial fraction expansion of $P(x)/C(x)$, where

$$C(x) = \prod_{j=1}^{t}(1 - w_jx) \tag{3.25}$$

and

$$P(x) = \sum_{j=1}^{t} w_j y_j \prod_{k=1,k\neq j}^{t} (1 - w_kx) \tag{3.26}$$

Therefore

$$\frac{P(x)}{C(x)} = s_0 + s_1x + s_2x^2 + ... \tag{3.27}$$

where $C(0) = 1$. $P(x)$ and $C(x)$ are relatively prime. Examination shows that Eq. 3.25 - Eq. 3.27 are exactly the forms of functions used in the Berlekamp iterative algorithm

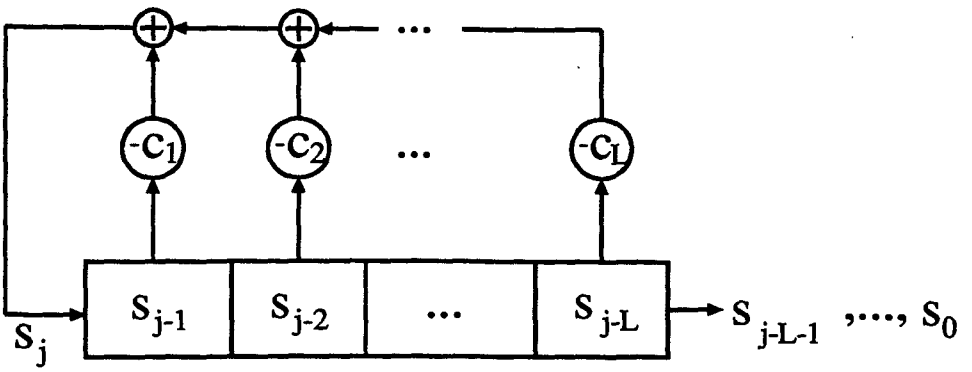


Figure 3.3: General L-stage Linear Feedback Shift-Register (LFSR)

for decoding nonbinary BCH codes [55]. This decoding algorithm is called the linear feedback shift register (LFSR) decoding. The block diagram of the LFSR synthesis algorithm is depicted in Fig. 3.3.

The known $2t$ lowest order coefficients of $S(x)$ are used as the initial states of the LFSR. Suppose $P(x)$ can be expressed in the form:

$$P(x) = p_0 + p_1 x + ... + p_{t-2} x^{t-2}$$

we can find the matrix equation

$$\begin{bmatrix} p_0 \\ p_1 \\ . \\ . \\ . \\ p_{t-2} \end{bmatrix} = \begin{bmatrix} 1 & 0 & ... & 0 & 0 \\ c_1 & 1 & ... & 0 & 0 \\ ..... & & & & \\ ..... & & & & \\ ..... & & & & \\ c_{t-2} & c_{t-3} & ... & c_1 & 1 \end{bmatrix} \begin{bmatrix} s_0 \\ s_1 \\ . \\ . \\ . \\ s_{t-2} \end{bmatrix}$$

which relates the coefficients of $P(x)$ to the coefficients of the connection polynomial $C(x)$ and the initial contents of the LFSR. Since the matrix is nonsingular, there will be a unique corresponding assignment of initial conditions for every $P(x)$. C(x) is the connection polynomial of the shortest LFSR that generates the sequence s whose

transform is $S(x)$. C(x) can be found using the Massey-Berlekamp iterative procedure described in Section 3.2.

The Berlekamp algorithm yields the $C(x)$ having smallest degree, which satisfies Eq. 3.20, resulting in a unique $P(x)/C(x)$ since it is assumed that not more than t errors occurred. Since the error term is $S(x) = (P(x)/C(x))M(x)$, decoding is complete and the information sequence is $f(x) = v(x)$ - $S(x)$.

This RS decoding algorithm is different from that stated in the previous section. Here the decoder uses the Berlekamp algorithm as well. However, the roots of the error locator polynomial do not need to be found as in the Chien search nor the values of the errors determined. Instead, a polynomial division is performed.

We can also observe that RS codes constructed by means of the Chinese remainder theorem have the interesting property that each symbol in the encoded word is determined solely by the information symbols. Symbols can be discarded from an encoded word, reducing the distance of the codeword. This observation leads to the realization of punctured Reed-Solomon codes. The codeword length does not need to be the full length, which is $n = q^n$ -1 as in the case of a normal RS code. If the distance of the full length codeword is $2t + 1$, up to $2t$ symbols can be discarded without changing the distance properties of the corresponding full length code. Thus, codes with arbitrary rates and lengths can be obtained.

To address the rate compatibility property of punctured RS codes based on the Chinese remainder theorem, we assume that there are two punctured RS codes with code rates R_1 and R_2, where $R_1 < R_2$. The rate compatibility restriction requires that the rate R_1 RS code retains all the residues used in the rate R_2 RS code.

3.4.1.3 Performance of Punctured Reed-Solomon Codes in Rayleigh Fading Channels

As we discussed in the previous subsections, punctured RS codes constructed by means of the Chinese remainder theorem have flexibility in codeword length and consequently are suitable for providing variable protection to the speech parameters. Punctured RS codes are also easy to implement, see [30].

For burst error correcting codes, such as the RS codes in our study, there are several factors that will affect the bit error performance apart from the code rate. These are the blocklength and the Galois field in which the RS codes is built. Since RS codes are a class of block codes, it follows that RS codes with longer blocklength have better error performance compared to those with shorter blocklength. Also, for RS codes with same code rate but constructed in different Galois fields $GF(p)$ and $GF(q)$, if $p > q$, then the RS codes built on $GF(p)$ protects better than that constructed on $GF(q)$.

The error performance of some of the RS codes on $GF(2^5)$ are depicted in Fig. 3.4 (hard decision decoding). To assess the effect that interleaving might have on RS codes, we also include the performance under the condition of perfect interleaving. The Rayleigh fading channel considered is a slowly fading channel with a fade rate of 0.003 (Hz). This corresponds to a vehicle speed of 12.5 km/h in the 800 - 900 Mhz band. An important observation is that the use of interleaving brings much improvement in the performance of RS codes. This implies that even though RS codes are meant to correct bursty errors, when channel fading is slow and the error burst is long, interleaving is still necessary to disperse the errors.

We also tested the error performance of punctured RS codes on the faster fading channel with a fading rate of 0.012. This is corresponding to a vehicle speed of 50 km/h in the 800 - 900 MHz band and is considered to be the normal transmission

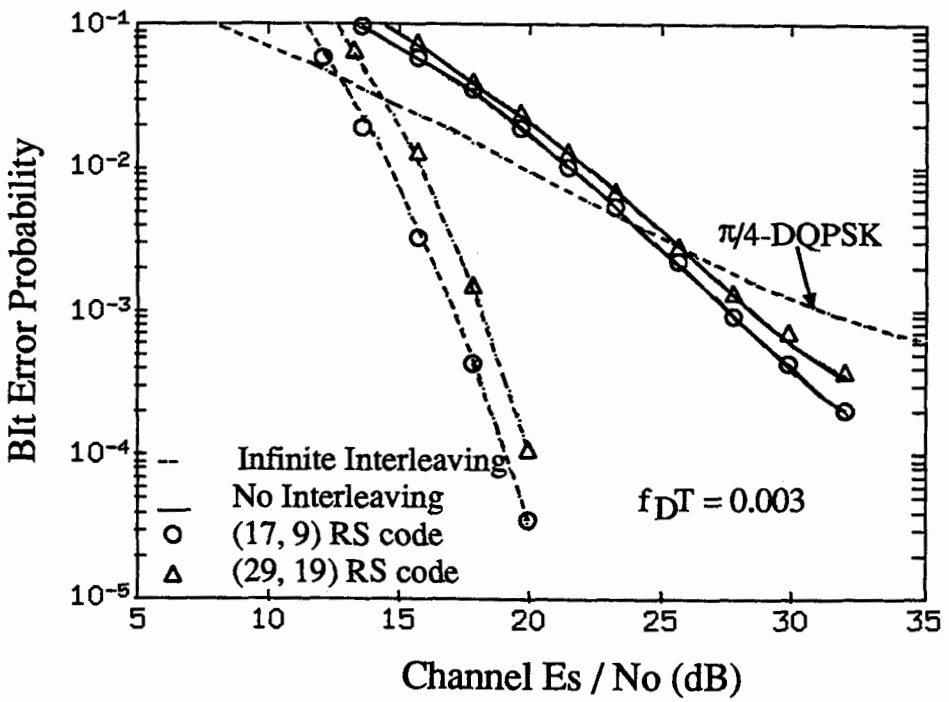


Figure 3.4: The error performance of some RS codes on $GF(2^5)$ with different interleaving at a $f_D T = 0.003$

condition in mobile radio communications. In this case, the improvement brought by interleaving is not as significant as that achieved in a slow fading condition, which coincides with the theoretical assumption.

3.4.2 Rate-Compatible Punctured Convolutional Codes

Convolutional coding with Viterbi decoding is considered to be a good random-error-correcting algorithm and widely used for forward error correction. As we discussed in Section 3.3, for convolutional codes with moderate constraint length, simple implementation with Viterbi decoding is possible. It can also achieve relatively large coding gains by using demodulator soft decision with Viterbi decoding and therefore provide approximately 2 dB more gain than the corresponding hard decision decoding [26, 56]. That is why convolutional codes with Viterbi decoding are widely used for error protection in various applications. General convolutional encoding and decoding schemes have been covered in Section 3.2.

To satisfy the need to protect the speech parameters with respect to their different significance to channel errors, we will use rate-compatible punctured convolutional codes (RCPC codes). The main advantages for using the RCPC codes is summarized as following:

1. They can be decoded by the maximum likelihood Viterbi algorithm.
2. Viterbi decoding is relatively easy to implement for codes with small constraint lengths.
3. Soft decision decoding is readily implementable in the Viterbi algorithm.
4. Only one flexible pair of encoder and decoder circuit is needed.

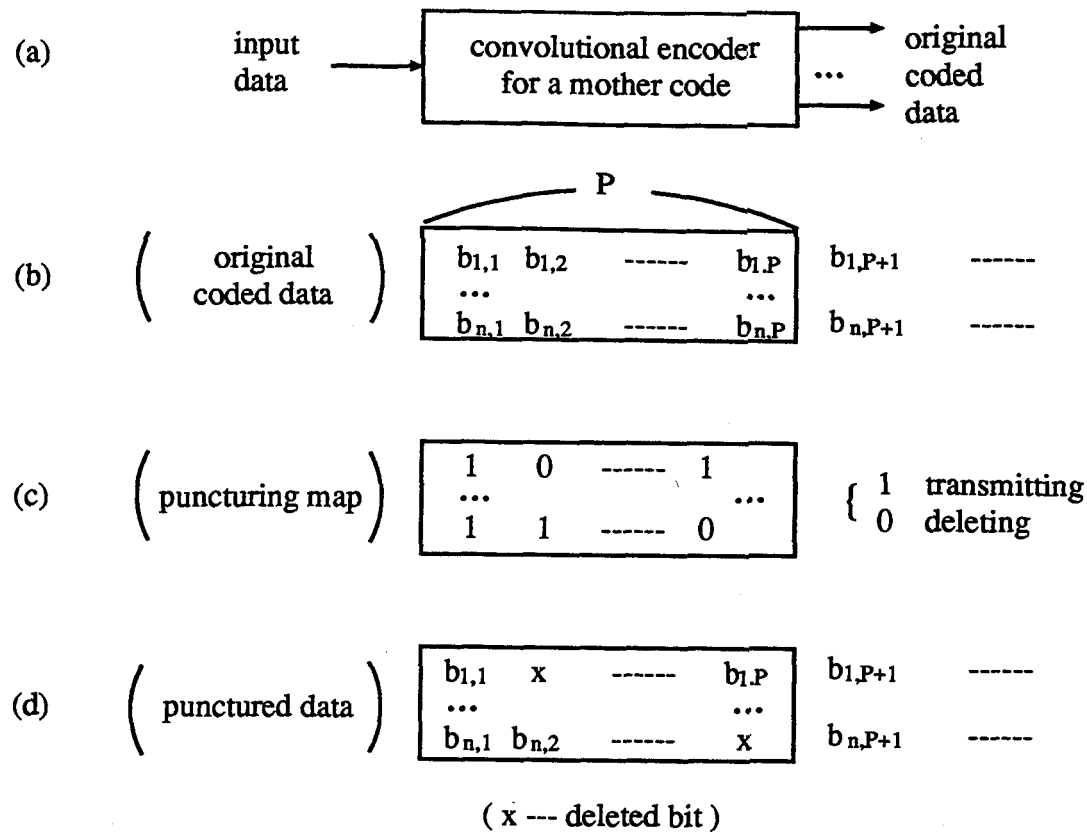


Figure 3.5: Basic Procedure for Constructing Punctured Codes from a rate 1/n convolutional code

3.4.2.1 Constructing Rate-Compatible Punctured Convolutional Codes

Puncturing is a method of obtaining higher rate convolutional codes from a mother code with a lower rate. As shown in [26], these punctured convolutional codes can perform almost as well as the best known conventional convolutional codes.

In order to explain how to construct a high rate punctured code from the low rate mother code, we show the procedure in Fig. 3.5. The encoder of the mother code is also included in Fig. 3.5. For a rate $1/n$ mother code with constraint length v, the

generator matrix is of the following form:

$$G(D) = \begin{bmatrix} g_{10} + g_{11}D + ... + g_{1v}D^v \\ g_{20} + g_{21}D + ... + g_{2v}D^v \\ ... \\ g_{n0} + g_{n1}D + ... + g_{nv}D^v \end{bmatrix}$$

There is a puncturing matrix $a(l)$ of size $n \times p$ associated with each punctured convolutional code:

$$a(l) = \begin{bmatrix} a_{11} & a_{12} & ... & a_{1p} \\ a_{21} & a_{22} & ... & a_{2p} \\ ... & & & \\ a_{n1} & a_{n2} & ... & a_{np} \end{bmatrix}$$

Here p is the puncturing period, and $a_{ij} \in (0, 1)$. Having $a_{ij} = 0$ means that the bit at that position in the encoded code sequence will be punctured, while $a_{ij} = 1$ means that the bit at that position in the encoded sequence will be transmitted. Suppose m positions in the puncturing matrix are zeros, then we obtain a punctured convolutional code with a higher rate :

$$R_c = \frac{p}{np - m}$$

Since the value of m can be 0, ..., $(n-1)p - 1$, we can obtain a family of punctured codes with code rates $R_c = 1/n, ..., p/(p+1)$.

Table 3.1 shows an example of a family of punctured convolutional codes presented by J. Hagenauer [53]. The mother code here is the rate 1/2 convolutional code with the generator matrix shown in the Table 3.1. The constraint length v is 4 and the puncturing period p is 4. Same as the notation used in Fig. 3.5, a "0" in the puncturing matrix means a puncturing operation at that position. The parameter l describes the order of the puncturing. The number of bits that is punctured is $(n-1)p - l$. As we can see from the table, when we go to a lower rate code in the table, we retain all the bits used by the higher rate code in the puncturing matrix and allow new "1's" only

RCPC Code M = 4, p=4		Generator Matrix $g = \begin{pmatrix} 1\,0\,0\,1\,1 \\ 1\,1\,1\,0\,1 \end{pmatrix}$		
l	1	2	3	4
Punctured Code Rate	4/5 4/5	4/6 2/3	4/7 4/7	4/8 1/2
Puncturing Map $a(l)$	1 1 1 1 1 0 0 0	1 1 1 1 1 0 1 0	1 1 1 1 1 1 1 0	1 1 1 1 1 1 1 1
d_f	3	4	5	7

Table 3.1: Rate-Compatible Punctured Convolutional Codes with M = 4, p = 4

in previously punctured positions. That is to say, all the bits of the higher rate code are used by the low rate code. If $a_{ij}(l_0) = 0$, then:

$$a_{ij}(l) = 1 \qquad for\ all\ l \geq l_0 \geq 1 \tag{3.28}$$

This equation is called the rate compatibility restriction.

At the receiver, the Viterbi algorithm [26, 56] is used for decoding. For all members in the RCPC code family, the Viterbi decoding operates on only one trellis. This can be seen in Fig. 3.6. A "x" on the trellis branches means that no bit has been transmitted for this position, in accordance with the puncturing map, and consequently the distance caused by the error at this bit position will not be counted.

The rate compatibility restriction given in Eq. 3.28 is very important. In a transitional phase between the two metric $a(l_1)$ and $a(l_2)$ where $l_2 > l_1$, we have to be sure that despite the transition, the distance properties of all paths originating in code l_1 do not suffer a loss of distance due to transitions, thus guaranteeing at least

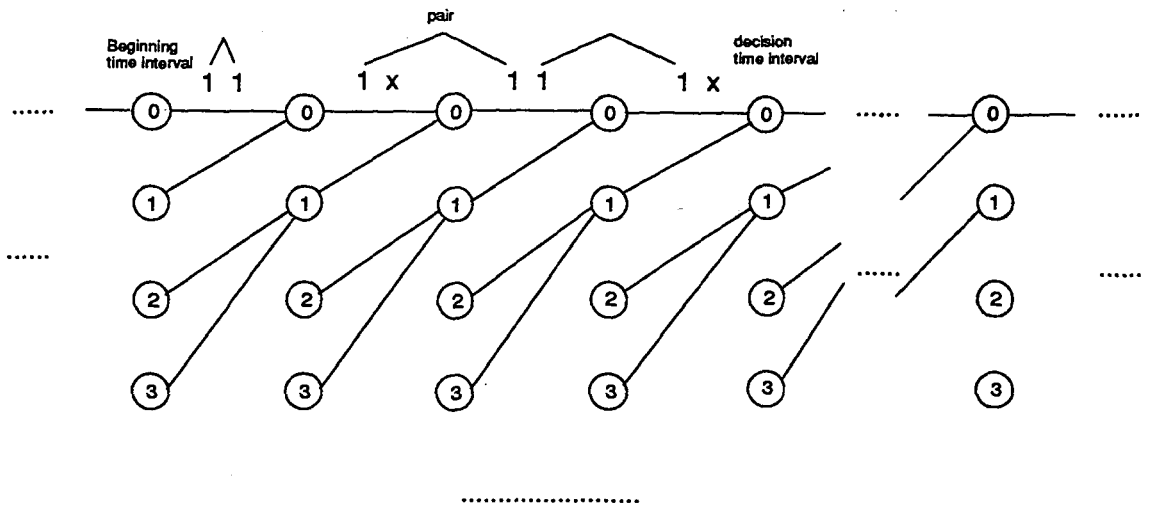


Figure 3.6: Viterbi decoding for a rate 2/3 punctured convolutional code from a rate 1/2 mother code

the designed performance [53]. Thus, it is advisable, although not necessary that the designed punctured convolutional codes satisfy the rate compatibility rule.

There is no known constructive way to determine the generator matrix $G(D)$ of the mother code and the rate-compatible punctured convolutional code. Good punctured convolutional codes are basically found by searching over all possible mother codes and puncturing matrices. Details can be found in [57].

3.4.2.2 Performance of Rate Compatible Punctured Convolutional Codes in Rayleigh Fading Channels

Viterbi algorithm with both hard and soft decision decoding for punctured convolutional codes are used in this study. A description of the decoding operation is given below.

The Viterbi decoding algorithm can be applied to punctured convolutional codes just like in the case of a regular convolutional code, but with the minor modification

of assigning a metric of zero to each of the deleted bits, i.e., only the transmitted bits count in the distance calculation. This is done by equipping the Viterbi decoder with a copy of the puncturing rule a. Since puncturing is done periodically, then $a_{ij+p} = a_{ij}$, where p is the puncturing period. With hard decision decoding, channel decoding is done after the demodulation and the Viterbi decoding deals with the demodulated binary sequence $\mathbf{r} = (r_1, r_2, ...)$. With soft decision decoding, the Viterbi decoder operates on the complex numbers generated by the demodulator.

Assume the $\pi/4$-shifted DQPSK modulation with differential coherent detection is used. Suppose that s_k is the kth transmitted $\pi/4$-DQPSK symbol. In the Rayleigh fading channel, the kth received sample is:

$$r_k = g_k s_k + n_k$$

For unquantized soft decision decoding of a mother code having rate 1/2, there is a symbol of two bits associated with each branch in the trellis. The metric associated with each branch of the trellis is simply:

$$\begin{aligned} m_j &= |r_k - r_{k-1} C_i|^2 \\ i &\in (1,2,3,4) \\ j &= 1,2,... \end{aligned}$$

where r_{k-1} is the received signal at the previous time interval and C_i is the phase shift shown in Table 2.2 which is determined by the two bits combination on each branch. For the case of punctured convolutional codes, the situation is a bit different. Let's refer to the trellis structure shown in Fig. 3.6. We consider the case of a rate 2/3 punctured convolutional code which is obtained by puncturing the rate 1/2 mother code. A "1" means that the bit at that position is transmitted and a "x" means the bit at that position will be punctured. With normal DQPSK modulation, we can actually separate the effect of the real and imaginary part of the complex modulation signal. This means that calculating the metric for a symbol generated by a bit pair

occupying adjacent transitions in the trellis is straightforward. However with $\frac{\pi}{4}$-DQPSK signaling, this is not the case. For this modulation format, the trellis diagram must be modified accordingly before soft decision decoding can be performed. In the case shown in Fig. 3.6, we have to pair the bits in three consecutive time intervals to make a decision on metric calculation at the third time interval.

As for error performance, the Viterbi upper bound on the bit error rate P_b is given by [57, 52, 58]

$$P_b \leq \frac{1}{p} \sum_{d=d_f}^{\inf} C_{d_f} P_d$$

where d_f is the free distance of the punctured code, C_{d_f} is the total number of bits on all free-distance paths [52, 53], and P_d is the probability that one incorrect path with free distance to the transmitted word is selected in the Viterbi decoding process. A good code should have a large d_f and a small C_{d_f}. Unfortunately no constructive method is known for determining the optimum puncturing map. Thus, the optimum puncturing matrix has to be searched among all possible puncturing maps and mother codes. In our study, we have used Hagenauer's punctured convolutional codes with constraint length $v = 4$ [53]. Table 3.2 shows the rate 2/3, 3/4, 4/5 punctured convolutional codes and the rate 1/2 mother code with constraint length $v = 2$, 3, 4, 5. Their corresponding d_f and C_{d_f} are also included. Due to implementation complexity, codes with constraint lengths greater than 5 are not considered. As shown in Table 3.2, punctured codes with $v = 5$ have large C_{d_f} although they have relatively large d_f. They are also more complex. So, we employed the punctured convolutional codes with $v = 4$.

We have also examined the bit error performance of punctured convolutional codes with both hard and soft decision decoding, and the results are shown in Fig. 3.7 to Fig. 3.9. Figure 3.7 shows the error performance of the rate 1/2 mother code in fading channels with $f_D T = 0.003$. Both soft and hard decision decoding results are included. Also the effect of interleaving was considered. It is seen that with no interleaving, the

(a) Map of deleting bits for punctured codes derived from 1/2 codes with $v = 2, \ldots, 5$

(1: transmitting, 0: deleting)

code rate \ v	2	3	4	5
1/2	1 (5) 1 (7)	1 (15) 1 (17)	1 (23) 1 (35)	1 (53) 1 (75)
2/3	10 11	11 10	11 10	10 11
3/4	101 110	110 101	101 110	100 111
4/5	1011 1100	1011 1100	1111 1000	1000 1111

v : Constraint length defined by Forney
(*) : Generator polynomial (octal notation) of original 1/2 code with maximum free distance

(b) d_f and C_d for punctured codes listed in (a)

code rate \ v	2		3		4		5	
	d_f	C_{d_f}	d_f	C_{d_f}	d_f	C_{d_f}	d_f	C_{d_f}
1/2	5	1	6	2	7	4	8	2
2/3	3	1	4	10	4	1	6	96
3/4	3	15	4	124	3	1	4	3
4/5	2	1	3	14	3	42	4	10

Table 3.2 Punctured convolutional codes with different constraint lengths

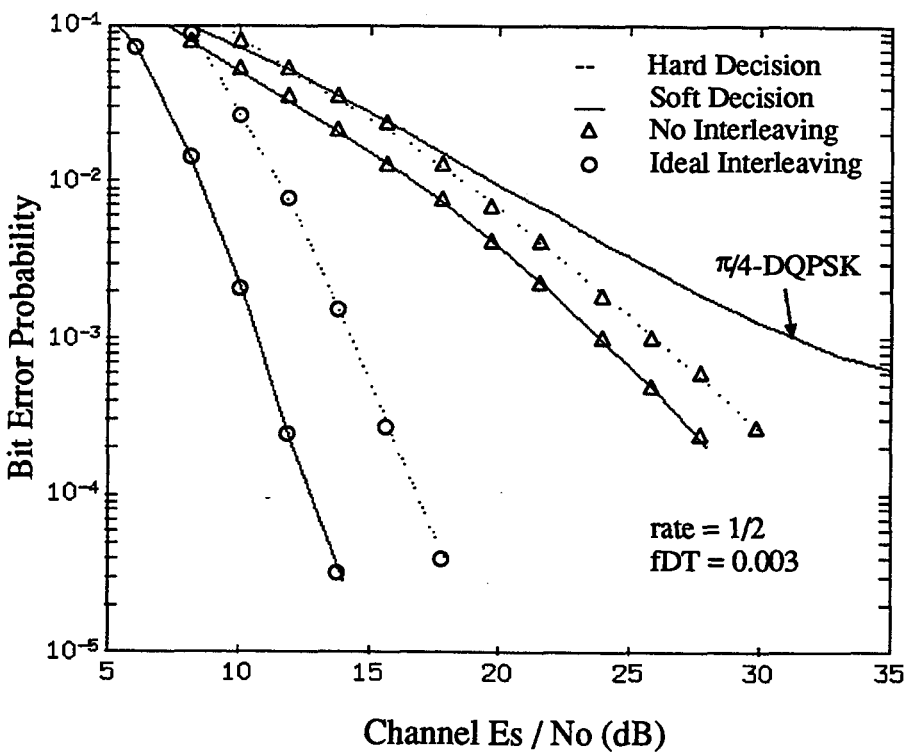


Figure 3.7: Error performance of the rate 1/2 RCPC code with both hard and soft decision decoding on Rayleigh fading channel at $f_D T = 0.003$

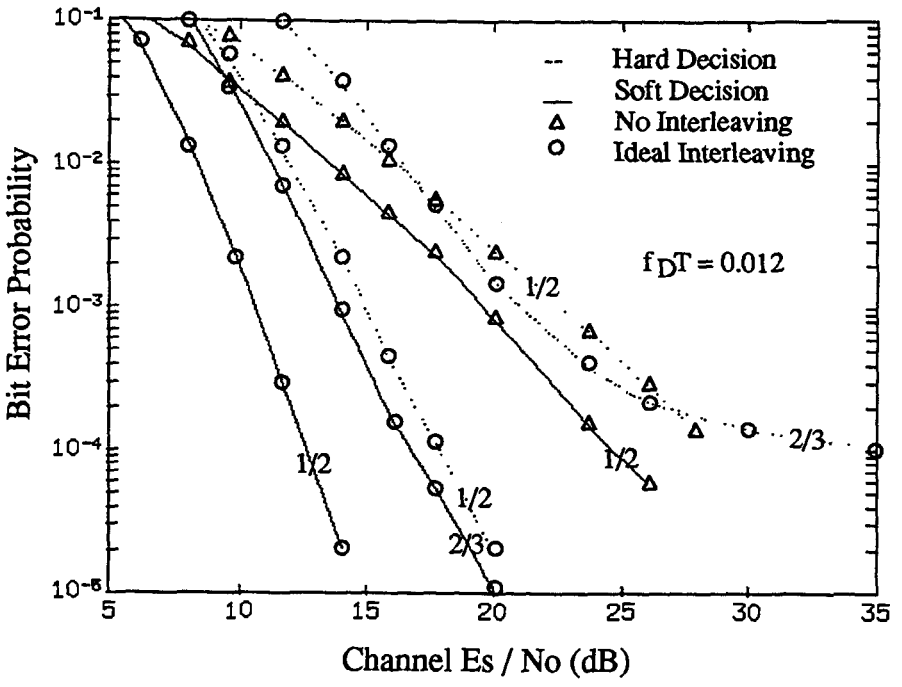


Figure 3.8: Error performance of the rate 1/2 and rate 2/3 RCPC codes for both hard and soft decision decoding on Rayleigh fading channel at a fade rate of 0.012

performance for soft decision decoding is more than 2 dB better than that for hard decision decoding. The soft decision decoding performs even better for the case with interleaving. It is clear that for random-error-correcting codes, interleaving brings more improvement than for RS codes.

Figure 3.8 contains the error performance of the rate 2/3 punctured convolutional code with ideal interleaving. Both soft and hard decision decoding results are included. The performance of the rate 1/2 mother code with both noninterleaving and infinite interleaving is also added for reference purposes. The fading rate here is $f_DT = 0.012$ compared to Fig. 3.7. We can see that the interleaving for convolutional codes has less effect with faster fading rate. The same observation was made earlier for RS codes.

In our experiments with punctured convolutional codes, we observed that the decoding window delay for the Viterbi algorithm plays a very important role in the

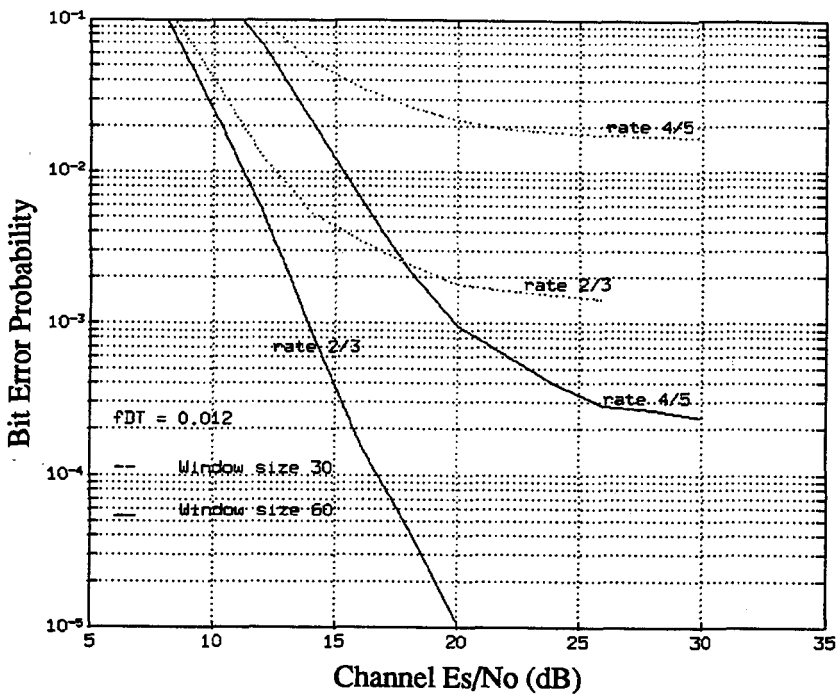


Figure 3.9: Error performance of the rate 2/3 and rate 4/5 RCPC codes for soft decision decoding with different decoding window delay in fully interleaved Rayleigh fading channel

correction capability of the soft decision decoder. Figure 3.9 shows the performance of the rate 2/3 and the rate 4/5 punctured convolutional codes with different decoding window sizes (30 and 60). Full interleaving in a Rayleigh fading channel with a fade rate of 0.012 is assumed. It is observed that the window length has to be more than 10 times the constraint length. In our experiment with punctured convolutional codes, the window size is chosen to be 60 intervals.

In this chapter, we discussed the general Forward-Error-Correction scheme and addressed in detail the burst-error-correction Reed-Solomon codes and the random-error-correcting convolutional codes. Detailed information about the error protection coding applied for the combined speech and channel coding system in this thesis, that is, the rate-compatible punctured convolutional code and the rate-compatible punctured convolutional code, is also provided. We will discuss in the next chapter the design strategy of our combined speech and channel coding system.

CHAPTER 4

Combined Speech and Channel Coding System

We discussed in the previous chapters separately the basics of the speech and channel coding techniques used in this study. In this chapter, we will show how to design a combined speech and channel codec based on the techniques described earlier. As the readers will find out, the channel codec used in this study is matched to the characteristics of the source codec. Section 4.1 describes some observed characteristics of the CELP coder when operating over a noisy channel and the motivation behind the idea of combined speech and channel coding. An evaluation of the bit error sensitivities of the 4k CELP coder is presented in Section 4.2. A combined speech and channel coding configuration that employs the unequal error protection is introduced in Section 4.3. In Section 4.4, we present the procedure for the search of optimal code rate allocation.

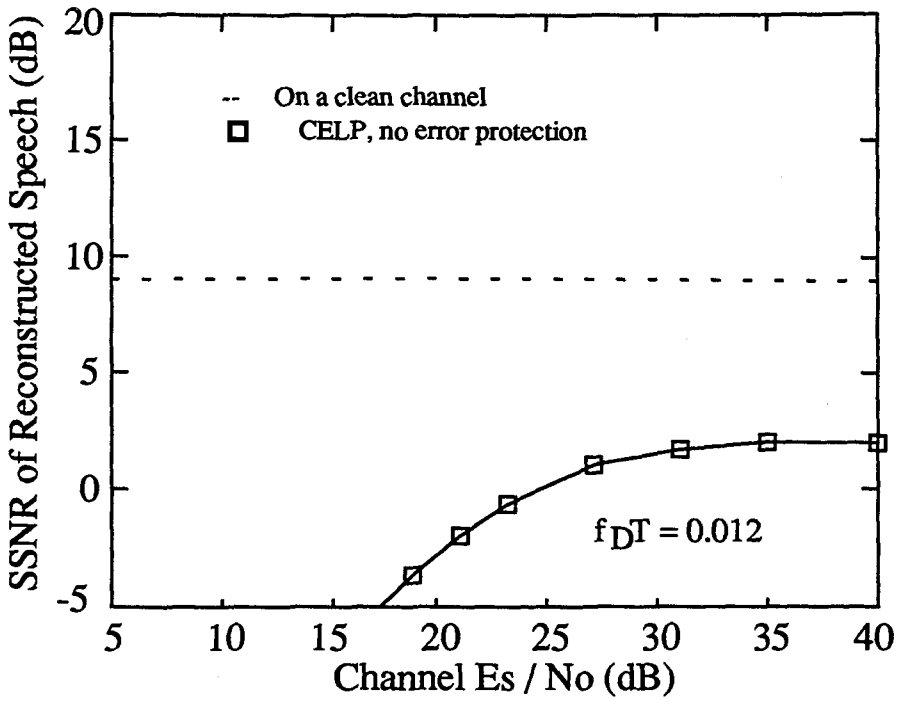


Figure 4.1: The performance of the 4k CELP coder on Rayleigh fading channel with fade rate $f_DT = 0.012$

4.1 Observation and Motivation

As discussed in Chapter 2, the CELP coder can produce good quality speech at rates as low as 4 kbits/s in the absence of channel errors. In the presence of channel errors, however, the reconstructed speech quality degrades dramatically. In the mobile radio application we consider in the study, the transmission channel is very noisy due to the adjacent channel interference and multi-path fading.

Figure 4.1 illustrates the performance of our 4k CELP coder in the Rayleigh fading channel. The performance of the 4k CELP coder in a clean channel is also included. In a clean channel, this codec produces the reconstructed speech at a Segmental Signal to Noise Ratio (SSNR) of 9.12 dB. The test was done using a speech segment containing 28 utterances from 7 male and 7 female speakers. We can see that the performance

of the speech coder degrades dramatically compared to that achieved under a clean channel condition. The performance does not change much when the channel SNR gets higher. This is due to the irreducible error floor of the DQPSK modulation on the Rayleigh fading channel. The fading condition $f_D T = 0.012$ shown in Fig. 4.1 is considered to be normal with vehicle speed of 50 km/h in the 800 – 900 MHz band.

Our experiments also show that some output bits of the 4k CELP coder are very sensitive to channel errors, while the others are not (will be presented in the next section). It should be pointed out that a bit is sensitive to channel errors if a transmission error in that particular bit causes a large degradation in the reconstructed speech quality. In contrast, errors in the insensitive bits do not cause a large degradation in speech quality. Experimental results show that the quality of the reconstructed speech improves significantly if the most sensitive bits are protected.

In mobile radio communications, channel bandwidth is a very scarce resource. Forward Error Correction, if applied, should be done in the most efficient way. In the Half Rate Digital Cellular Application for North America, the total transmission rate assigned to source and channel coding is about 6.5 kbits/s. In this study, we assign 4 kbits/s to the speech coder, and use the remaining 2.5 kbits/s for channel protection. Since, as mentioned above, different parameters of the speech output have different levels of sensitivities to channel errors, we introduce in this thesis a combined speech and channel coding scheme which makes use of this information provided by the speech coder. This leads to the unequal error protection scheme described in Section 4.3. Below we will present a detailed analysis of the different bit error sensitivities.

4.2 Evaluation of the Bit Error Sensitivity

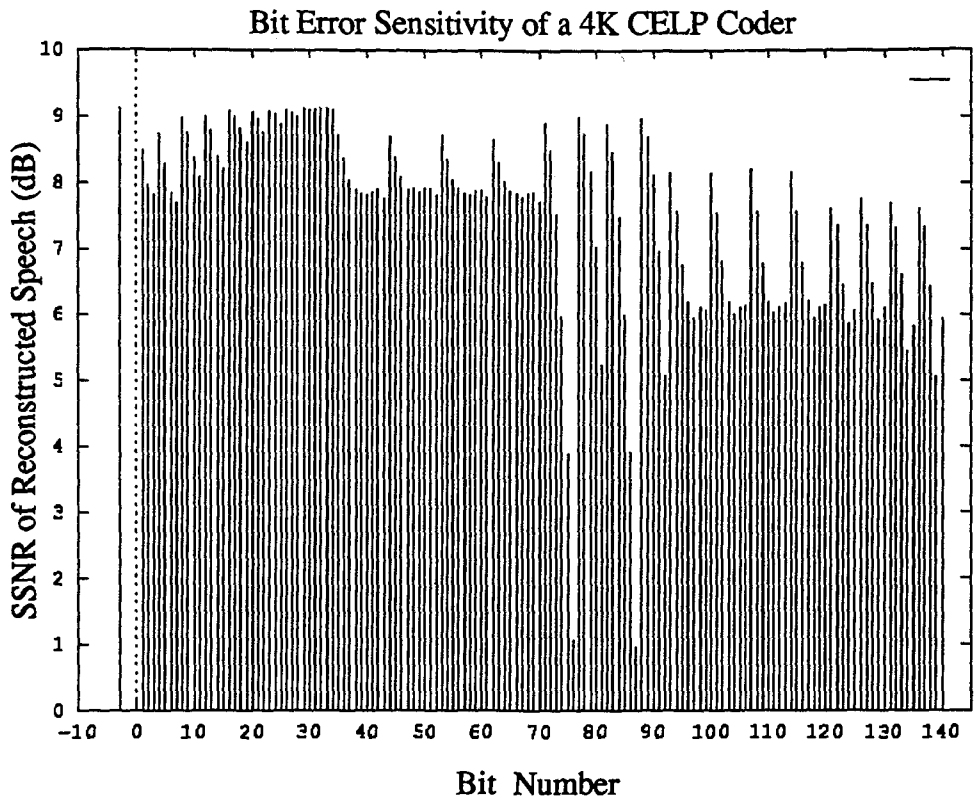
As explained in Chapter 2, in a CELP coding system, the speech signal is analyzed frame by frame and represented by the LPC coefficients, the pitch period, the pitch

gain, the excitation codeword, and the excitation gain. These parameters are either scalar quantized or vector quantized, and the indices of the chosen codewords are represented by the natural binary code (NBC). The binary representation will be transmitted to the physical channel and corrupted by the channel disturbance. In the 4k CELP coder used in this study, all parameters except the scalar quantized LPC coefficients are vector quantized.

The codebook indices from the speech coder are transmitted to the receiver in binary form. At the receiver, the source decoder stores a replica of each codebook in the source encoder. On a clean channel, the transmitted and received indices will be identical. However, on a noisy channel, the received indices will be different from the transmitted ones, and consequently the reconstructed parameters will be different from the chosen ones at the encoder. This can cause a large distortion between the original speech and the reconstructed speech. We observed that errors occurred at different bit positions have different levels of influence on the quality of the reconstructed speech. That is, some bits in CELP are more sensitive to channel errors than the others. Consequently, a bit error sensitivity test is necessary before applying error protection.

The speech signal is analyzed and synthesized frame by frame in the CELP coder. As seen in Table 2.1 (also included in Fig. 4.2 (b)), there are 140 bits per frame in the 4k CELP coder. An efficient procedure for evaluating the bit error sensitivity is as following. We introduce a single bit error to a specific bit of a frame with a probability p_c. We measure the SSNR between the original speech and the reconstructed speech obtained from the distorted speech. The SSNR of the reconstructed speech will be used to determine the bit error sensitivity.

The bit error sensitivity for different parameters in one speech frame for the 4k CELP is depicted in Fig. 4.2(a), for an average bit error rate of 10^{-2}. This test is done on 2000 frames of speech spoken by both male and female speakers. In Fig. 4.2(a),



(a) Bit Error Sensitivity of the 4K CELP coder

	LPC	Shape CB	Adaptive CB
Update	35 ms	8.75 ms	8.75 ms
Parameters	10 LSPs	1 gain; 1 index	1 gain; 1 index
Bits / Frame	34	index : 4x9 ; gain : 2x6 2x5	index : 4x7 ; gain : 4x5
Bit Number	1 - 34	index : 35 - 70 ; gain : 71 - 92	index : 93 - 120 ; gain : 121 - 140
Total Number of Bits in a Frame : 140			

(b) Bit Allocation and Update Rate for the 4K CELP coder

Figure 4.2: Bit Error Sensitivity of the 4k CELP coder

along the x axis is the bit number of one speech frame. Along y axis is the resulting averaged SSNR value of the reconstructed speech. More sensitive bits have lower SSNR values. There are 140 bits in a frame time of 35 ms for the 4k CELP coder. The correspondence between the bit numbers and the parameters is also included in Fig. 4.2(b). It is shown that there exists a large dynamic range of different bit error sensitivities. The bits representing the excitation gains, the pitch gains and the pitch periods are more sensitive than the bits representing the LPC coefficients and the codebook indices. The most significant bit (MSB) in each parameter is also more sensitive than the others.

The above bit error sensitivities are obtained with a single bit error randomly introduced at specific positions in a speech frame. For the same bit error rate (BER), this sensitivity would be slightly different from the sensitivity when multiple errors occurred in a frame since there are some correlation among the distortions caused by the different bit errors. Accurate bit error sensitivities can be obtained by complicated simulations. However, the bit error sensitivities evaluated above are adequate to inform us the degree of significance of different bits.

4.3 The Combined Speech and Channel Coding Configuration

It is shown in the previous sections that the speech coder's output bits have a large dynamic range of bit error sensitivity. An effective way to protect the output speech information is to apply unequal forward error protection. This means that more powerful FEC codes should be used to protect more sensitive output bits. We discuss in this section the configuration of such a combined speech and channel coding system.

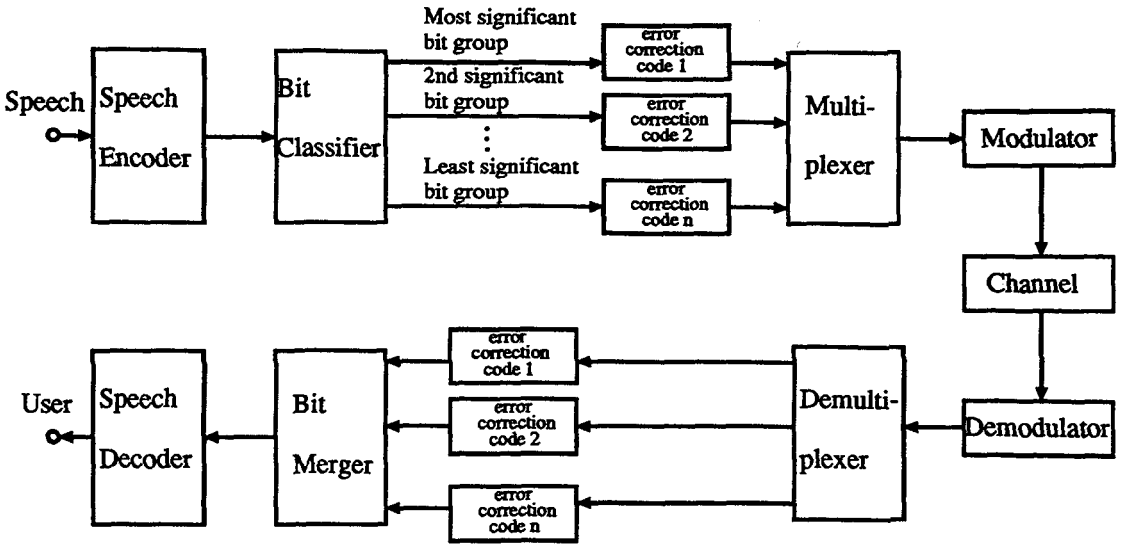


Figure 4.3: Combined speech and channel coding configuration

The block diagram of our combined speech and channel coding system configuration is shown in Fig. 4.3. The speech coder is the 4k CELP coder discussed in Chapter 2. In the 4k CELP coder, the speech signal is analyzed frame by frame. Assume that the output data stream of the CELP coder contains N_s bits in each frame. The bit error sensitivities of the N_s bits can be determined from the procedure described in the previous section. Then the N_s bits are rearranged in the order shown in Fig. 4.4 according to the relative bit error sensitivity. The first bit that appears in Fig. 4.4 is more sensitive to channel errors than the second, which in turn is more sensitive than the third, and so on. The bit classifier divides the N_s bits into n groups with n_k bits in the kth group, where

$$\sum_{k=1}^{n} n_k = N_s$$

The n different groups of bits are then protected by n different FEC codes. The most sensitive bit group is protected by the most powerful FEC code. The least powerful FEC code is used to protect the least sensitive bit group. In practice, the least sensitive group can be left unprotected according to the channel and transmission conditions. The encoded bits from different encoders are then multiplexed, modulated,

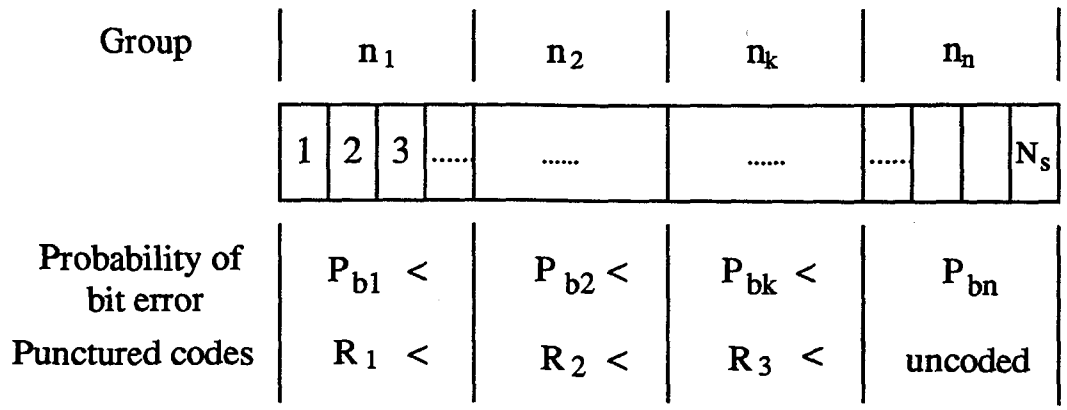


Figure 4.4: Grouping of information bits in each speech frame according to their relative sensitivities

and sent over the communication channel. At the receiving end, the inverse operation is performed. The output bits from n different decoders are rearranged to the original form and fed to the speech decoder to reconstruct the original speech.

From the combined speech and channel coding configuration shown in Fig. 4.3, it seems that n different channel codecs are necessary to provide unequal error protection and the system complexity increases accordingly. Rate-compatible punctured Reed-Solomon codes (PRS) and rate-compatible punctured convolutional codes (RCPC) are considered in this study to circumvent this problem. Punctured Reed-Solomon codec and rate-compatible punctured convolutional codec are able to generate FEC codes of different code rates but with only one pair of encoder and decoder. As a result, the complexity of the combined system does not increase. Reed-Solomon codes are chosen for this study since they are meant to correct multiple errors – a scenario that is very probable to happen in mobile radio channels. Convolutional codes are selected since they can achieve a relatively large decoding gain by using the soft decision Viterbi decoding algorithm [52, 53]. Also, certain level of interleaving can be applied to convolutional codes to help combat the channel error bursts. These two FEC codes have been discussed in detail in Chapter 3. Their bit error rate performance on Rayleigh fading channels has also been presented.

Suppose that the total transmission rate R for speech and channel coding is fixed, which is the case in our study. If the speech coding rate is R_s, then the channel coding rate R_c will be $R_c = R/R_s$. In other words, if there are altogether $N = RT$ bits that can be used to represent one error-protected speech frame, and if $N_s = R_sT$ bits are used for speech coding, then the number of bits available for channel coding in each frame will be $N_c = N - N_s$. Note that T is the duration of a speech frame. As discussed above, the N_s bits in the speech frames have different levels of requirement for FEC. They are divided into n groups and protected by n different FEC codes according to their bit error sensitivities to channel errors. The issue of dividing the N_c channel coding bits among the n different levels of FEC codes is an optimization problem. It will be addressed in Section 4.4.

4.4 Optimal Code Rate Allocation

The objective of this section is to discuss the issue of deciding the optimal channel code rate allocation strategy. A full search method was used to obtain the optimal rate allocation using the SSNR criterion. As indicated in Chapter 2, the SSNR of the reconstructed speech is an objective performance measure that is more correlated to the reconstructed speech quality than the normal SNR criterion.

Recall that the total transmission bit rate for combined speech and channel codec is around 6.5 kbits/s, of which 4 kbits/s are assigned to speech coding and the remaining 2.5 kbits/s are used for channel protection. For the 4k CELP codec we used, there are 140 bits in each speech frame of 35 ms. In other words, there are 228 bits altogether in each error protected speech frame, of which 88 bits are used for error protection. We arrange the 88 bits among the 140 speech output bits according to their bit error sensitivities, as shown in Fig. 4.5. The leftmost bit is the most sensitive bit, and the rightmost bit is the least sensitive bit. The sensitivity level of the bits from left to

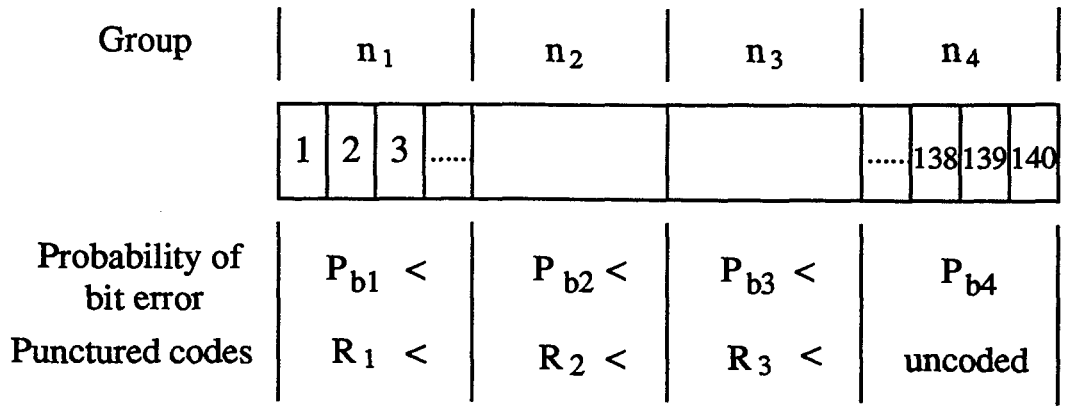


Figure 4.5: Bit arrangement of the 4k CELP coder according to bit error sensitivities

right is decreasing. We considered four levels of error protection in this study, which are provided by the rate 1/2, 2/3, 3/4, and 1 punctured RS codes, or the rate 1/2, 2/3, 4/5, and 1 rate-compatible punctured convolutional codes with $v = 4$. More levels of error protection and other codes could be considered. Under the restriction of the number of redundant bits, we found out that these groups of rates are more practical.

Our objective here is to find out the code rate allocation that gives the best combined system performance under the SSNR criterion. A direct way of doing this is through an exhaustive search of all the possible combinations of the code rates. We will present the results of this search for both the combined CELP and punctured Reed-Solomon (CELP/PRS) and the combined CELP and RCPC (CELP/RCPC) systems.

4.4.1 Optimal Code Rate Search for the CELP/PRS system

As shown above, the source bits are arranged according to their sensitivities, the possible bit allocation should satisfy the condition:

$$R_{c_1} \leq R_{c_2} \leq ... \leq R_{c_{140}}$$

where R_{c_i} is the code rate corresponding to the ith bit in each speech frame. As shown in Fig. 4.5, n_1, n_2, n_3, and n_4 number of bits are assigned to each protecting group, where n_i, i = 1, 2, 3, 4, can be only positive integers or zeros. They satisfy:

$$n_1 + n_2 + n_3 + n_4 = 140 \tag{4.1}$$

Since n_1 bits are protected by the rate 1/2 punctured RS code, n_2 bits by the rate 2/3 punctured RS code, etc., and since the total number of redundant bits is 88, we have:

$$n_1 + \frac{n_2}{2} + \frac{n_3}{3} = 88 \tag{4.2}$$

The Galois field where the RS codes are constructed is $GF(2^5)$. That is, there are 5 bits in each nonbinary symbols in the codeword. Therefore, it is easier if we rewrite the above equations as:

$$l_1 + l_2 + l_3 + l_4 = 28 \tag{4.3}$$

$$l_1 + \frac{l_2}{2} + \frac{l_3}{3} = 18 \tag{4.4}$$

where l_i, i = 1, 2, 3, 4, is the number of information symbols in group i with each symbol containing 5 bits. Notice that the division of 88 by 5 would give 17.6 and we round it to 18.

Eq. 4.3 and Eq. 4.4 can be written as:

$$l_4 = 28 - (l_1 + l_2 + l_3) \tag{4.5}$$

with

$$l_3 = 18 - (3l_1 + \frac{3}{2}l_2) \tag{4.6}$$

We can see from Eq. 4.5 and Eq. 4.6 that l_3 and l_4 are dependent on l_1 and l_2. Therefore, the problem of finding the optimal rate allocation is equivalent to finding the optimum values for l_1 and l_2.

To assist in the search, we can make use of the following bounds from Eq. 4.5 and Eq. 4.6:

$$8 \leq l_1 \leq 18 \tag{4.7}$$

$$max(52 - 4l_1, 0) \leq l_2 \leq 36 - 2l_1 \tag{4.8}$$

Since a $(n,\ k)$ RS code has the property that $n - k = 2t$, i.e., $n - k$ has to be an even number, certain adjustment has to be done when searching for the optimal code rate allocation. This often results in a RS code that does not match exactly the designed code rate. For example, we might get a (27, 13) RS code to approximate a rate 1/2 RS code.

According to the combined speech and channel coding configuration shown in Fig. 4.3, the SSNR of the reconstructed speech can be obtained for each of the possible rate allocations. The optimal rate allocation is simply the one that gives the largest SSNR. The optimal rate allocation of our experiment is: a (17, 9) punctured RS code protects the 45 most sensitive speech bits, a (29, 19) punctured RS code is found to protect the rest 95 less sensitive bits in a speech frame. No rate 3/4 RS code is necessary. Also there is no bit that does not need any protection.

4.4.2 Optimal Code Rate allocation Search for the Combined CELP/RCPC codec

Following the same strategy presented above for the combined CELP/PRS codec, we also optimized the bit allocation for the combined CELP/RCPC codec. Since we are only interested in the rate compatible codes, we do not use the rate 3/4 punctured convolutional code. Thus, we chose the four levels of protection provided by the rate 1/2, 2/3, 4/5 convolutional codes with $v = 4$, and no coding. The optimal rate allocation search is similar to that for punctured RS codes, but it is simpler since convolutional codes deal with binary representation.

Using the same notation as in the case of punctured RS codes, we have:

$$n_1 + n_2 + n_3 + n_4 = 140 \tag{4.9}$$

Since the data in the 4th group are uncoded, no redundant bits are assigned in this group. There are 88 redundant bits in each speech frame, out of which up to 4 bits are used for termination bits at the end of each trellis (We terminate the trellis on the basis of each speech frame), then:

$$n_1 + \frac{n_2}{2} + \frac{n_3}{4} = 84 \tag{4.10}$$

From Eq. 4.9 and Eq. 4.10, we have:

$$36 \leq n_1 \leq 84 \tag{4.11}$$

$$196 - 3n_1 \leq n_2 \leq 168 - 2n_1 \tag{4.12}$$

After examining all the possibilities, we obtain the bit allocation that gives the highest SSNR value of the reconstructed speech. This bit allocation is: 88 most sensitive bits need to be protected by the rate 1/2 mother code, 38 bits in the next sensitive group are coded with rate the 2/3 punctured convolutional code, the next 8 bits are assigned to the third class protected by the rate 4/5 punctured convolutional code, and the remaining 6 bits are left unprotected.

The full search method for finding the bit allocation presented above leads to a optimal result. Another approach would be a suboptimal search which provides similar performance but consumes less computing time. This work can be found in [20].

CHAPTER 5

Experimental Results

In this chapter, we present the experimental results on the combined speech and channel coding system operating in the mobile environment. We simulated the combined system under different channel and transmission conditions. Specifically, the normalized Doppler frequencies in the Rayleigh fading channel studied were $f_D T = 0.012$ and 0.003. At first, we briefly describe the system model used in our experiments in Section 5.1. Section 5.2 gives the interleaving strategy used in the experiment. The experimental results for the combined speech and channel codecs from different aspects are given in Section 5.3, 5.4, and 5.5 respectively.

5.1 System Model

Our target application area is the North American Half Rate Digital Cellular Standard [24]. In this standard, the channel separation is 30 kHz. The total bit rate is 8 kbps/user. The gross bit rate for speech and channel coding is altogether 6.5 kbits/s, and the remaining 1.5 kbits/s are available for system overhead. In our design, we assigned 4 kbits/s to the speech information and 2.5 kbits/s for error protection. In

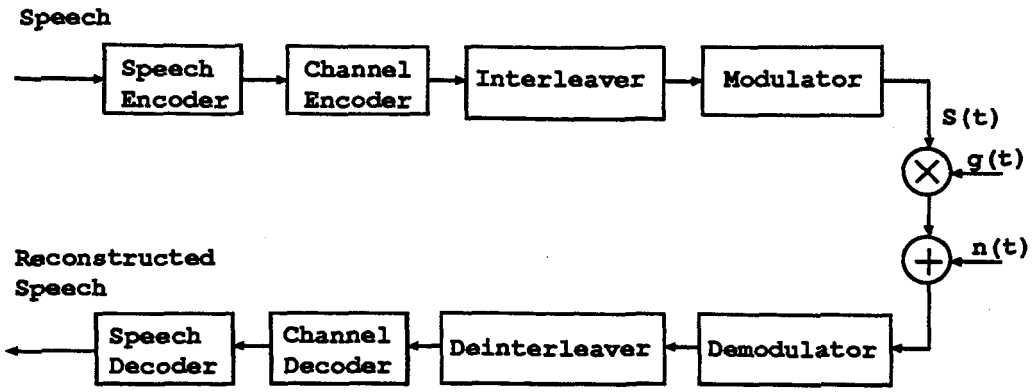


Figure 5.1: Combined speech and channel coder

this application, the channel is considered to be the harsh Rayleigh fading channel with large burst of errors. The requirement is that the speech codec be able to provide good quality speech at a channel bit error rate as high as 10^{-2}.

We show in Fig. 5.1 the combined speech and channel coding system considered in our study. The analog speech is sampled at a rate of 8 KHz and each sample is digitized into 12 bits. The 4k CELP coder takes the digitized speech as the input. The digitized speech, at a rate of 96 kbps, is compressed by the CELP coder and transformed into a binary bit stream of 4 kbits/s. This bit stream is then protected by a bank of error correction codes in an optimal fashion. Both rate-compatible punctured Reed-Solomon codes and rate-compatible punctured convolutional codes are considered as forward error correction codes in this thesis. Interleaving is then applied to the protected bit stream with a rate of 6.5 kbits/s. Prior to entering the channel, the bit stream is modulated via $\pi/4$ shifted DQPSK modulation scheme, the modulation scheme adopted by the North American Full Rate Digital Cellular Standard. The modulated signal will be transmitted to the physical channel and corrupted by the channel noise. Since we are considering the mobile radio application, the physical channel is modeled as a flat Rayleigh fading channel.

At the receiving end, under the condition of hard decision channel decoding, the

differential demodulation is first conducted and the demodulated bit stream is then decoded by the channel decoder. In the case of soft decision decoding of punctured convolutional codes, the demodulation and channel decoding is done at one step. Deinterleaving is done before channel decoding. The output binary stream of the channel decoder will be processed by the CELP decoder to reconstruct (a distorted version of) the original analog speech.

In this chapter, we will present the results of the software simulation of the performance of the combined system in Fig. 5.1 from different aspects.

5.2 Interleaving Strategy

For the combined speech and channel coding system considered in this thesis, the total transmission rate is 6.5 kbits/s. As discussed in detail in the previous chapters, the speech codec and the channel codec are combined in an optimal fashion according to the different bit error sensitivities of the speech element shown in Fig. 4.2 (a).

For error protection, first we optimize the bit allocation for unequal error protection according to the Segmental SNR (SSNR) of the reconstructed speech and the interleaving delay. For punctured RS codes, the field in which RS codes were built is also a factor. We consider four levels of error protection for each speech frame of 140 bits. For RS codes, they are provided by codes with rates 1/2, 2/3, 3/4 and 1. With $f_D T = 0.012$ and channel E_s/N_0 at 23 dB, we found out that with no interleaving delay, the best system uses punctured RS codes in $GF(2^5)$ with two levels of protection. Specifically, a (17,9) code was used to protect the most sensitive 45 bits, and a (29,19) code was used to protect the remaining 95 bits in one speech frame. The total bit rate for this combined CELP/PRS codec is roughly 6.5 kbps.

For punctured convolutional codes with hard decision decoding, four levels of error

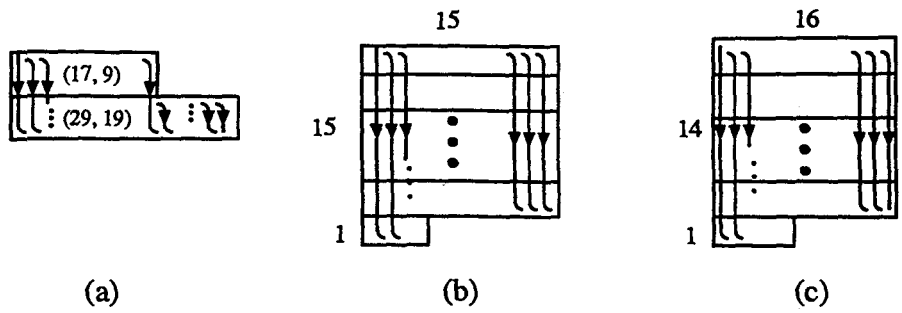


Figure 5.2: Interleaving strategy of the combined systems

protection are provided by codes with rates 1/2, 2/3, 4/5 and 1. Also, with $f_D T = 0.012$, a channel SNR of 23 dB, and no interleaving delay, the best bit allocation turned out to be: 38 most sensitive bits needed rate 1/2 protection, 88 less sensitive bits were protected by the rate 2/3 code, 8 bits were encoded with a rate 4/5 code, and the least sensitive 6 bits were left unprotected. The total transmission rate is also about 6.5 kbits/s.

Note that with $\pi/4$-DQPSK modulation and a Rayleigh fading channel with a fade rate of 0.012, a channel E_s/N_0 of 23 dB translates into a channel bit error rate of 10^{-2}.

In order to improve the performance of the combined system without increasing the interleaving delay, we utilized intra-frame interleaving for both the combined CELP/PRS codec and the combined CELP/RCPC codec. The strategy of intra-frame interleaving is shown in Fig. 5.2. For the combined CELP/PRS codec, the (17, 9) and (29, 19) punctured RS codes in each speech frame were interleaved symbol by symbol (each symbol contains 5 bits in $GF(2^5)$), which is shown in Fig. 5.2 (a). The arrows indicate the order of transmission. For the combined CELP/RCPC codec with hard decision decoding, since there are altogether 229 bits in each error protected speech frame, a 15x15 block interleaver (225 bits) was used and the remaining 4 bits are also interleaved, see Fig. 5.2 (b). For the combined CELP/RCPC codec with

soft decision decoding, since $\pi/4$-DQPSK modulation scheme was used and there are two bits in each received complex signal, we used a 16x14 block interleaver (224 bits) together with the remaining 5 bits, see Fig. 5.2 (c).

In the following sections, when we talk about the performance of the combined speech and channel coding system with no interleaving delay, we actually mean that no inter-frame interleaving is used.

5.3 Performance of the Combined Speech and Channel Coding System

We present in this section the performance of the:

1. 4k CELP codec in a clean channel.
2. 4k CELP codec, without error protection, in a Rayleigh fading channel.
3. Combined CELP/PRS codec, with no interleaving delay, in a Rayleigh fading channel.
4. Combined CELP/RCPC codec with hard decision decoding, with no interleaving delay, in a Rayleigh fading channel.
5. Combined CELP/RCPC codec with soft decision decoding, with no interleaving delay, in a Rayleigh fading channel.

5.3.1 4k CELP Coder in Rayleigh Fading Channels

The 4k CELP coder is able to reconstruct the analog speech with a SSNR of 9.12 dB in a clean channel. Informal listening tests indicate that the reconstructed speech is

of good quality.

Without error protection, the output data stream from the 4k CELP coder is directly modulated into a $\pi/4$-DQPSK signal and transmitted over the Rayleigh fading channel. The channel fading rates considered are $f_DT = 0.012$ and $f_DT = 0.003$. As mentioned above, $f_DT = 0.012$ is considered to be a normal fading channel which corresponds to a vehicle speed of 50 km/h in the 800 - 900 MHz band, while a channel with $f_DT = 0.003$ is a slower fading channel. The performances of the 4k CELP coder in both channels are presented in Fig. 5.3. The performance of the 4k CELP in a clean channel is also included in Fig. 5.3 as a reference. Along the x axis is the signal to noise ratio E_s/N_0 per channel symbol in decibels, where E_s is the average received signal energy per symbol and $N_0/2$ is the power spectral density of the bandpass Gaussian noise. Along the vertical axis is the SSNR of the reconstructed speech. It is observed that the performance of the 4k CELP coder in the faster fading channel degrades dramatically from that achieved in a clean channel. For the slower fading channel, the performance approaches that in a clean channel when channel E_s/N_0 is about 35 dB or greater. The difference in performance in the two cases is due primarily to the modulation scheme used, which in this case is $\frac{\pi}{4}$-DQPSK with differential detection. This kind of modulation technique exhibits irreducible error floor. The error floor is larger in channels with larger fade rate.

Our objective is to transmit good quality speech at a BER as high as 10^{-2} with no interleaving delay. Note that a channel BER of 10^{-2} is equivalent to a channel SNR of 23 dB in a Rayleigh fading channel with a fade rate of 0.012. Informal listening tests were conducted for the 4k CELP coder in the Rayleigh fading channel with fade rate of 0.012 and a channel SNR of 23 dB. The reconstructed speech is quite noisy and the speech quality is not acceptable. We will show in the next subsection the improvement brought by applying the error protection.

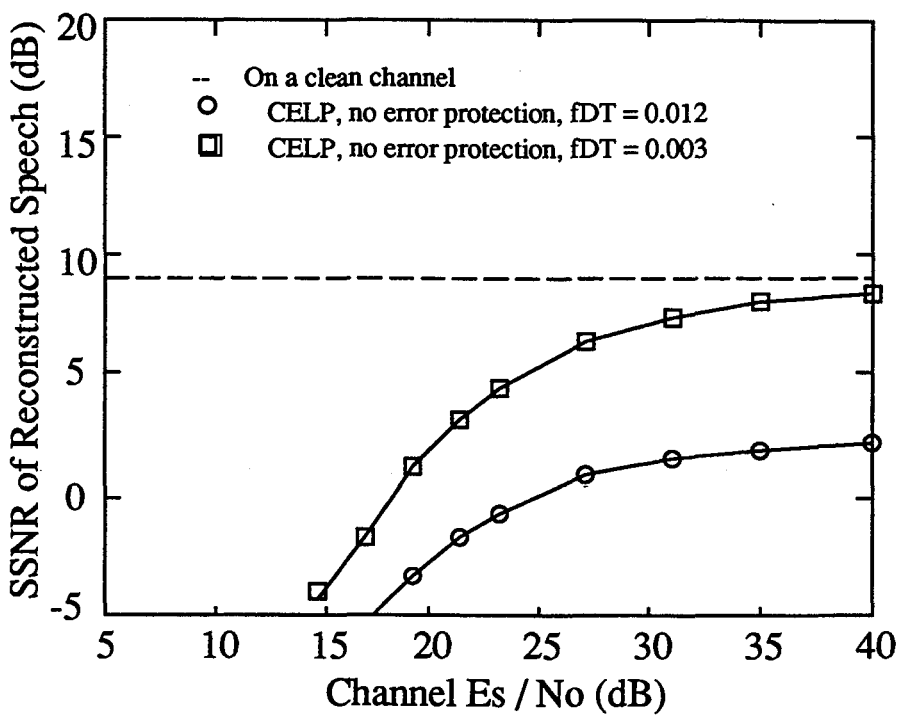


Figure 5.3: The performance of the 4k CELP coder in fading channels with different fading rates

5.3.2 Performance of the Combined Speech and Channel Coding systems

We will present in this subsection the performance of the combined CELP/PRS codec and the combined CELP/RCPC codec with both hard and soft decision decoding. No interleaving delay is assumed.

As described in Section 2.3, the fading variables $g_k's$ are correlated. The correlation depends on the fading spectrum given in Eq. 2.10. The purpose of the interleaving is to de-correlate the fading experienced by the modulated symbols. Although interleaving can combat the error bursts in the Rayleigh fading transmission environment, it also introduces transmission delay. The deeper the interleaving, the longer the interleaving delay. In real time speech communication, only small delays are tolerable. As an example, the Digital Cellular Communication Standard allows a 50 ms delay. Since speech coding handles speech frame by frame, a delay of one speech frame time is inevitable, which is 35 ms in our study. So, we would try to provide good quality speech without introducing additional interleaving delay.

The block diagram of the combined speech and channel coding system in the Rayleigh fading environment is shown in Fig. 5.1. We assigned 4 kbps to speech coding and 2.5 kbits/s are available for error protection. For both the combined CELP/PRS and the combined CELP/RCPC codec, we considered four levels of error protection. The code rate allocations are optimized at a channel SNR of 23 dB when $f_DT = 0.012$. Fig. 5.4 shows the SSNR of the reconstructed speech versus channel E_s/N_0 of the combined speech and channel coding systems at $f_DT = 0.012$ with no interleaving delay. Again, no interleaving delay means no inter-frame interleaving is introduced. Specifically, Fig. 5.4 includes the performance of the combined CELP/PRS codec, the combined CELP/RCPC codec with hard decision decoding, and the combined CELP/RCPC codec with soft decision decoding (unquantized). The performance of

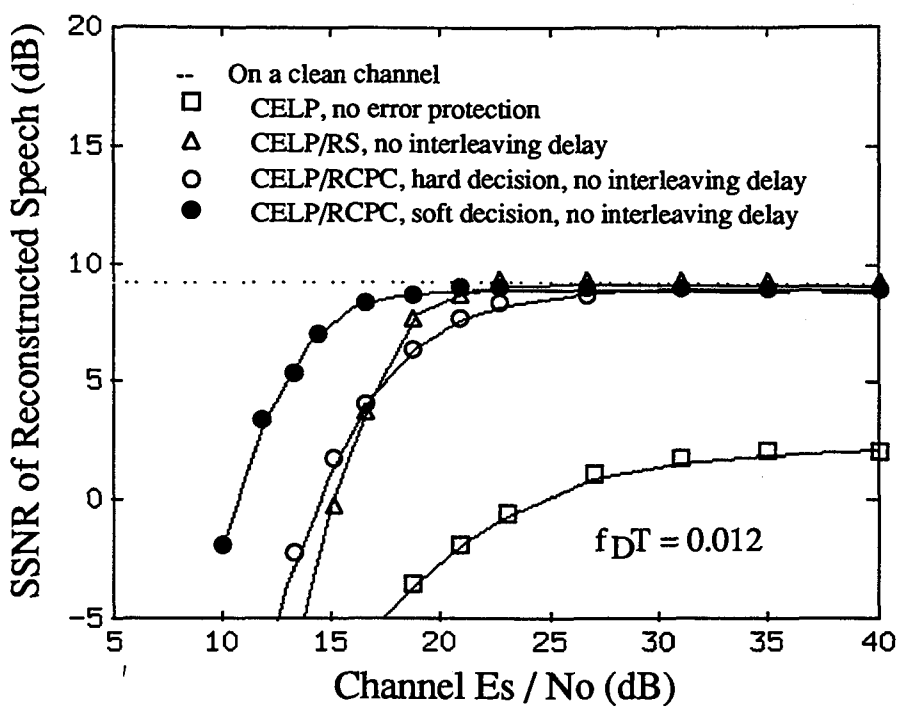


Figure 5.4: Performance of the combined CELP/PRS codec and the combined CELP/RCPC codec with both hard and soft decision decoding in a Rayleigh fading channel with $f_DT = 0.012$

the 4k CELP coder with no error protection and that in a clean channel are also included. As seen from Fig. 5.4, error protection provides large improvement on the reconstructed speech quality. Specifically, at a channel SNR of 23 dB, the combined CELP/PRS codec provides more than 9.7 dB improvement in SSNR. For the combined CELP/RCPC codec with hard decision decoding, error protection brings about 8.94 dB improvement in SSNR. For combined CELP/RCPC codec with unquantized soft decision decoding, the improvement will be about 9.58 dB. Objectively, at a channel SNR of 23 dB, punctured RS codes provide best protection. Subjectively, informal listening tests show that the three combined CELP/FEC codecs perform roughly the same. The difference between the reconstructed speech quality of the combined systems in the Rayleigh fading condition given above and that of the 4k CELP in a clean channel is imperceptible.

At the low channel SNR region, the performance of the combined CELP/PRS codec degrades more rapidly than the combined CELP/RCPC codec, especially when compared to the combined CELP/RCPC codec with soft decision decoding. It shows that punctured convolutional codes with either hard or soft decision decoding perform better than the punctured Reed-Solomon codes.

From the point of view of the system complexity, according to the computer simulation time consumed by the different combined systems, the combined CELP/PRS codec is the most complex one, while the combined CELP/RCPC codec with hard decision decoding has the least complexity. One point that is worth mentioning is that the decision depth of the Viterbi decoding algorithm is very important for the performance of the combined CELP/RCPC system. As mention in Chapter 3, the effect of the decision depth is more obvious for punctured convolutional codes than that for mother codes, especially in the case of soft decision decoding. There is a large difference between the system performance with decision depth of 60 and that of 30. The decision depth of the Viterbi decoder is chosen to be 60 intervals in our experiments.

In the next section, we will address the effect of the Doppler frequency on the system performance.

5.4 Effect of the Doppler Frequency

In the previous section, we presented the simulation results for the combined speech and channel coding systems in a channel with a normal fade rate of 0.012. In this section, we will address the effect of Doppler frequency on the system performance. Specifically, we will compare the system performance at normalized Doppler frequencies of 0.012 and 0.003.

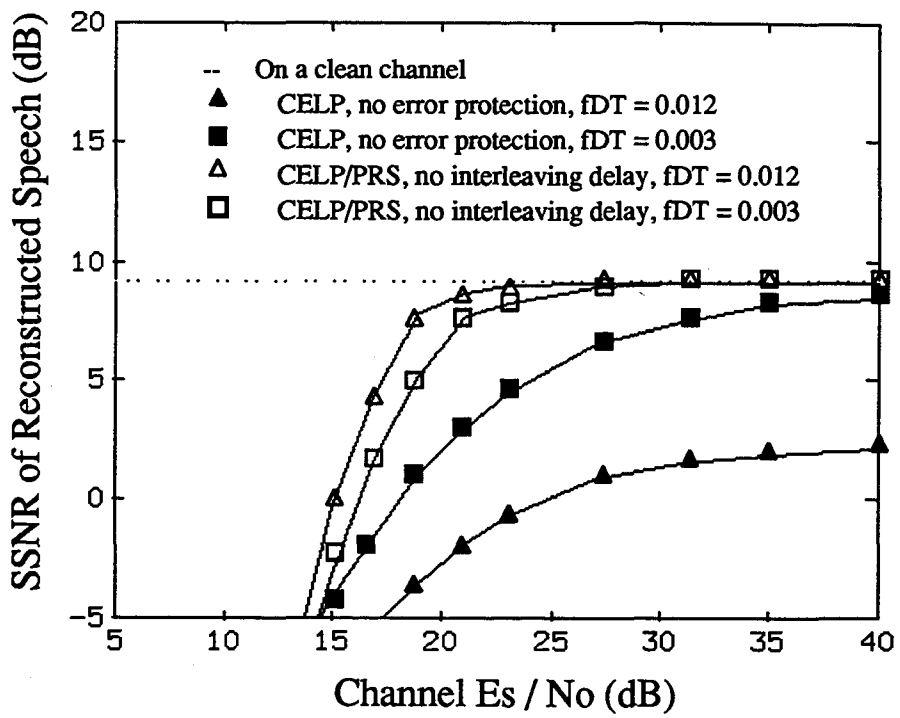


Figure 5.5: Performance of the combined CELP/PRS codec on different fading channels with zero delay

As seen from the previous section, the unprotected 4k CELP coder performs better in a slower fading channel because of the modulation scheme used. We show in Fig. 5.5 the performance of the 4k CELP coder after protected by the punctured Reed-Solomon codes. Fig. 5.5 consists the performance of the combined CELP/PRS codec with no interleaving delay and the performance of the 4k CELP coder with no error protection in Rayleigh fading channels with different fade rates. It is seen that for the 4k CELP coder without error protection, the speech coder performs better in a slow Rayleigh fading channel than in a fast Rayleigh fading channel. On the other hand, for the system with error protection, i.e., the combined CELP/PRS codec, it appears that faster fading provides higher SSNR of the reconstructed speech at lower channel SNR region. This is because that in a Rayleigh fading channel with a low channel SNR, the signal energy level is comparably low and the error burst is accordingly long. This is especially true for signals in slower fading channels. The intra-frame interleaving scheme we introduced can help improving the ability of combating the error burst, but it is still not enough to efficiently mitigate the error burst when the channel SNR is low, which is more obvious for slow fading case. We can conclude that for the system without error protection, a slower fade rate is preferable, while for combined system, we prefer a faster fading channel.

We also examined the effect of Doppler frequencies on combined CELP/RCPC codec with either hard or soft decision decoding. These results are included in Fig. 5.6 and Fig. 5.7 respectively for hard and soft decision decoding. We can see the same tendency as discussed above for the combined CELP/PRS codec shown in Fig. 5.5. Accordingly, same conclusion can be drawn.

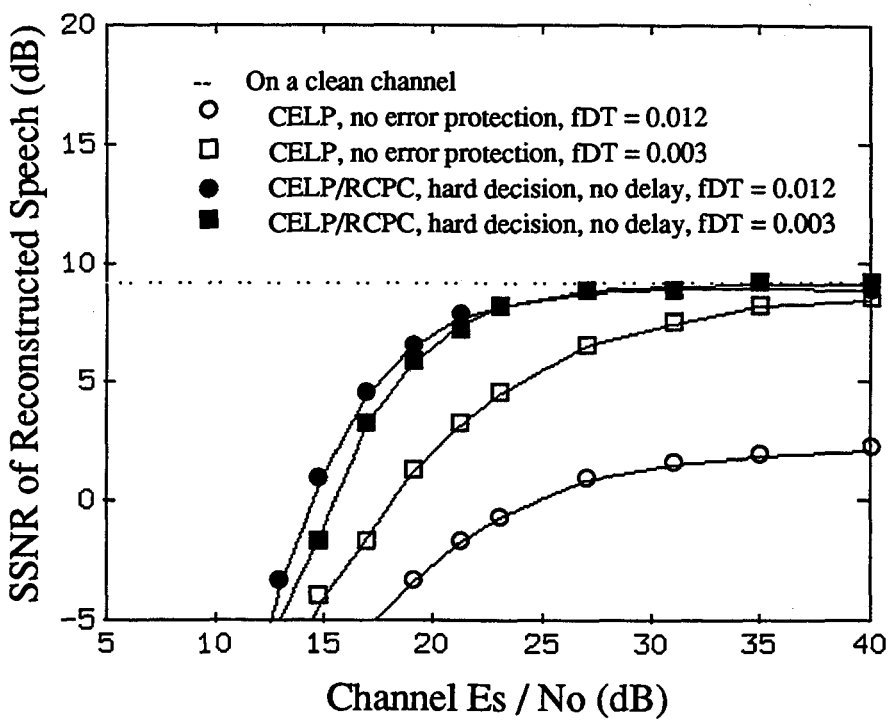


Figure 5.6: Performance of the combined CELP/RCPC codec with hard decision decoding in different fading channels with zero interleaving delay

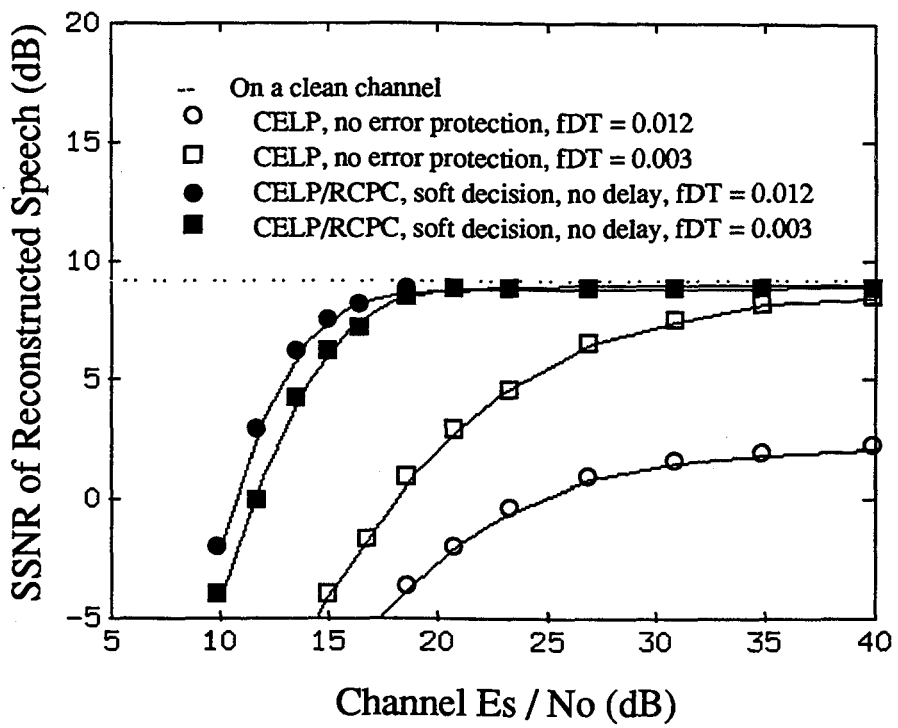


Figure 5.7: Performance of the combined CELP/RCPC codec with soft decision decoding in fading channels with different Doppler frequencies with no interleaving delay

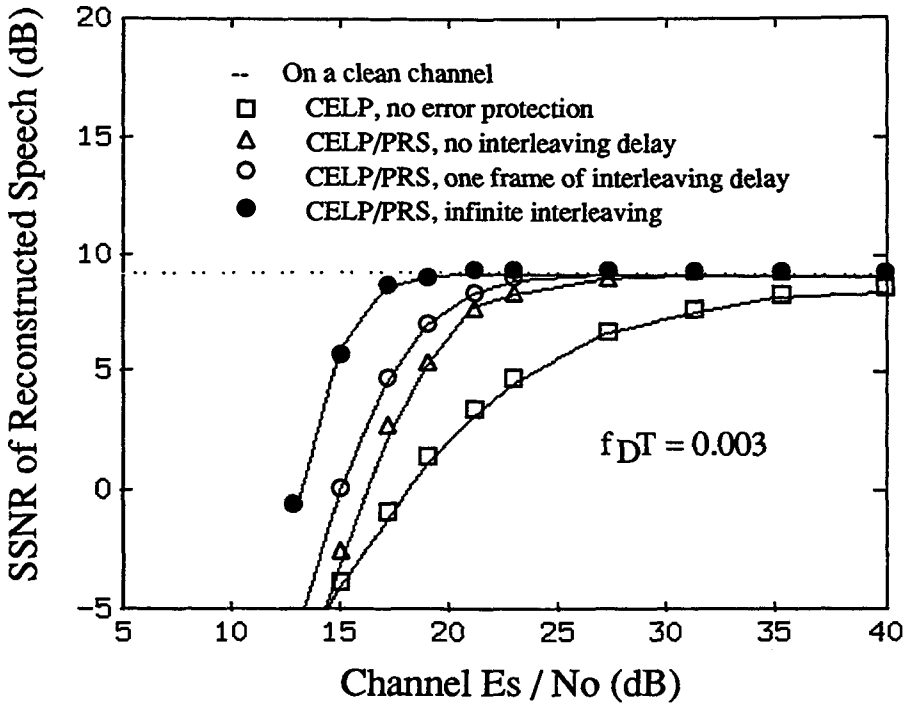


Figure 5.8: Interleaving effect for the combined CELP/PRS codec for $f_DT = 0.003$

5.5 Effect of Inter-frame Interleaving

To give another perspective of the combined speech and channel coding systems studied in this thesis, we now show the interleaving effect on the performance of the combined systems.

As discussed in Section 5.2, intra-frame interleaving is employed to efficiently combat error bursts without increasing the interleaving delay. Details of the interleaving strategy can be found in Fig. 5.2. It is also suggested in the previous section that when the channel SNR is low or when fading is slow, the resulting error bursts are long. In this case, intra-frame interleaving itself may not be sufficient to randomize the channel errors. Interleaving with two or more frames of speech data will be helpful to improve the system performance.

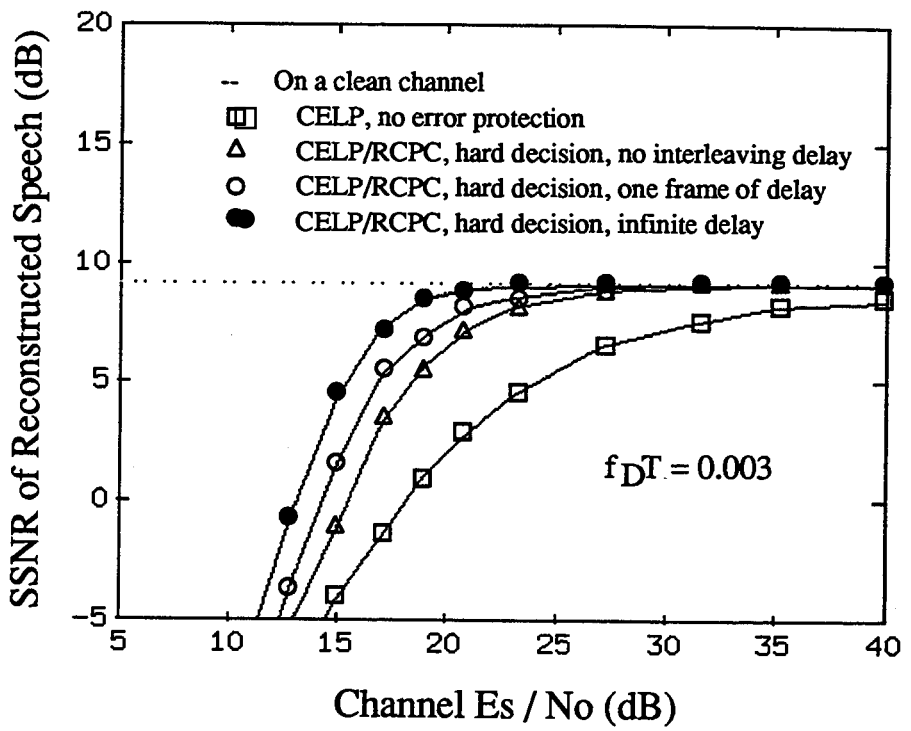


Figure 5.9: Interleaving effect for the combined CELP/RCPC codec with hard decision decoding for $f_DT = 0.003$

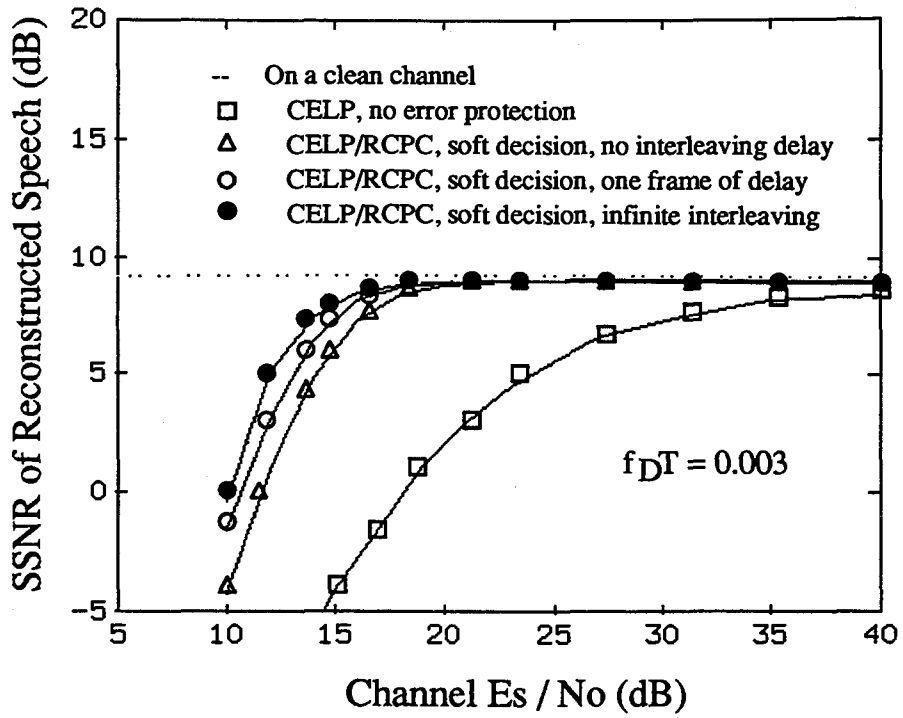


Figure 5.10: Interleaving effect for the combined CELP/RCPC codec with soft decision decoding for $f_D T = 0.003$

We show the interleaving (inter-frame interleaving) effect for the combined CELP/PRS codec, combined CELP/RCPC codec with hard decision decoding, and the combined CELP/RCPC codec with soft decision decoding in Fig. 5.8, Fig. 5.9, and Fig. 5.10 respectively. The channel fade rate is $f_D T = 0.003$. The levels of interleaving delay considered are: 0 delay, a delay of one frame, and infinite interleaving delay. The performance of the 4k CELP coder in a clean channel and that in a Rayleigh fading channel with no error protection are also included for reference. It is clear from the three figures that inter-frame interleaving is definitely required at low channel SNR region, whereas for high channel SNR, there is no need to use any inter-frame interleaving (the three curves representing different levels of interleaving delay converge).

On the other hand, at a normal vehicle speed where $f_D T = 0.012$, we observed

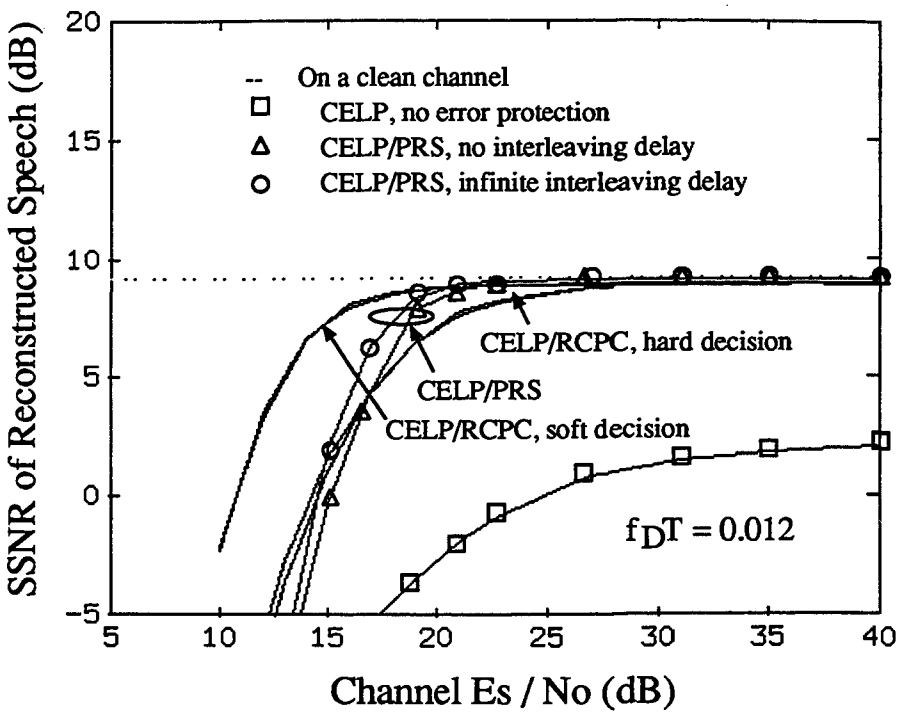


Figure 5.11: Interleaving effect for the combined speech and channel coding systems for $f_D T = 0.012$

from Fig. 5.11 that the performance improvement provided by increasing the interleaving intensity is very minimal. Specifically, Fig. 5.11 includes the performance of the three combined codecs with both zero interleaving delay and a delay of infinity. The 4k CELP coder's performance in a clean channel and that in a Rayleigh fading channel with no error protection is also included. It is clear from Fig. 5.11 that only the interleaving effect in the combined CELP/PRS codec is identifiable. Even that, the difference between no interleaving and infinite interleaving is small. For the combined CELP/RCPC codec with either hard or soft decision decoding, inter-frame interleaving is absolutely not necessary. These results lead us to believe that our combined speech and channel coding systems can provide good quality speech and require little interleaving.

The computer simulation time for the combined CELP/RCPC codec is less than

that for the combined CELP/PRS codec. Also, among the three systems, the combined CELP/RCPC codec with soft decision decoding seems dominantly better than the other two.

CHAPTER 6

Conclusions and Future Work

This thesis has considered the error protection of digital speech transmitted over mobile radio channels. The speech codec is based on Code-Excited Linear Predictive (CELP) coding and it operates at a rate of 4 kbps. Three channel coding systems, all employing unequal error protection (UEP) and all adding about 2.5 kbps of redundant bits, are proposed. The first system uses rate compatible punctured Reed-Solomon (RS) codes constructed from the Chinese remainder theorem. On the other hand, the second and third error protection system use the rate compatible punctured convolutional codes of Hagenauer. The difference between these last two systems is that one uses the simpler hard decision decoding while the other uses soft decision decoding. Given a family of rate compatible codes, the procedure for finding the optimal UEP strategy is also presented in the thesis.

We found that with zero interleaving delay and larger channel signal-to-noise ratio (SNR), RS codes provides the best performance in terms of the segmental signal-to-noise ratio (SSNR) of the reconstructed speech. Informal listening tests, however, indicates that the three systems exhibit no perceptible difference, compared to the reconstructed speech transmitted over a clean channel. At lower channel SNR though,

convolutional codes are definitely better than RS codes, especially when soft decision decoding is employed.

The effects of fade rate and/or interleaving depth have also been studied in the thesis. In general, it is fair to say that given a certain interleaving strategy, the effect of the fade rate is minimal in all three systems. However with slow fading (a normalized fade rate of 0.003), all three systems show significant improvement in performance when the degree of interleaving is increased. With fast fading (a fade rate of 1.2 percent the baud rate), the performance of three systems is insensitive to the degree of interleaving.

In the authors' opinion, convolutional codes are preferred over RS codes, due to mainly their better performance at the low channel SNR region. Because of the frame format considered, the combined speech and channel codec used in this study can be applicable to the half rate digital cellular system.

For the future work, there are several directions that can be approached.

The work in this thesis is done in a flat Rayleigh fading channel. However, further tests of the system performance in the frequency selective fading environment need to be considered.

In this thesis, the optimization of the combined speech and channel coding system is done with a predefined source coding rate and a channel coding rate. The trade-off of source coding against channel coding can be a way to further improve the system performance. When the channel is in a good condition, fewer bits can be assigned to channel coding. While when the channel is in a noisy condition, more bits can be assigned to channel coding. This will result in a combined system with variable coding rates. The channel condition may be judged by testing the correlation between the LPC coefficients in two or more consecutive frames at the receiver. There is a high correlation between LPC coefficients in adjacent frames for the speech signal. If

this correlation is detected at the receiver, the channel is then considered to be in a good condition. Otherwise, the channel is noisy.

Another possible approach is using the combined modulation and coding scheme. The combination of modulation and channel coding can reduce the bit error rate without increasing the bandwidth and power. Combined source coding, channel coding, and modulation can be a solution to a more efficient communication system.

APPENDIX A

Theorem: The matrix of syndromes

$$M = \begin{bmatrix} S_1 & S_2 & \ldots & Su \\ S_2 & S_3 & \ldots & Su+1 \\ \ldots\ldots\ldots & & & \\ S_u & S_{u+1} & \ldots & S2u-1 \end{bmatrix}$$

is nonsingular if u is equal to v, the number of errors that actually occurred. The matrix is singular if u is greater than v.

Proof: Let $X_u = 0$ for $u > v$. Let A be the Vandermonde matrix

$$A = \begin{bmatrix} 1 & 1 & \ldots & 1 \\ X_1 & X_2 & \ldots & X_u \\ \ldots\ldots\ldots & & & \\ X_1^{u-1} & X_2^{u-1} & \ldots & X_u^{u-1} \end{bmatrix}$$

with elements $A_{ij} = X_j^{i-1}$, and let B be the diagonal matrix

$$B = \begin{bmatrix} Y_1X_1 & 0 & \ldots & 0 \\ 0 & Y_2X_2 & \ldots & 0 \\ 0 & 0 & \ldots & Y_uX_u \end{bmatrix}$$

then the matrix product ABA^T has elements

$$(ABA^T)_{ij} = \sum_{l=1}^{u} X_l^{i-1} \sum_{k=1}^{u} Y_l X_l \delta_{lk} X_k^{j-1} = \sum_{l=1}^{u} X_l^{i-1} Y_l X_l X_L^{j-1} = \sum_{l=1}^{u} Y_l X_l^{i+j-1}$$

which is the ijth element of the matrix M. Therefore $M = ABA^T$. Hence the determinant of M satisfies

$$det(M) = det(A)det(B)det(A)$$

If u is greater than v, then $det(B) = 0$. Hence $det(M) = 0$, and M is singular. If u is equal to v, then $det(B) \neq 0$. Further, the Vandermonde matrix A has a nonzero determinant if the columns are different and nonzero, which is true if u is equal to v. Hence $det(M) \neq 0$ QED

REFERENCES

[1] William C. Y. Lee, *Mobile Communications Engineering*, McGraw-Hill Book Company, 1982.

[2] William C.Y. Lee, *Mobile Communications Design Fundamentals*, Howard W. Sams & Co., 1986.

[3] J. H. Chen, G. Davidson, A. Gersho and K. Zeger, 'Speech Coding for the Mobile Satellite Experiment', Proc. IEEE Int. Comm. Conf., 1987, pp756-763.

[4] B. Atal and M. Schroeder, 'Stochastic Coding of Speech at Very Low Bit Rates', Proc. IEEE Int. Comm. Conf., 1984, pp1610-1613.

[5] M.R. Schroeder and B. Atal, 'Code-Excited Linear Predictor (CELP): High Quality Speech at Very Low Bit Rates', ICASSP 85, vol. 1, 25.1.1.

[6] Peter Kroon and Ed F. Deprettere, 'A Class of Analysis-by-Synthesis Predictive Coders for High Quality Speech Coding ar Rates Between 4.8 and 16 kbits/s', IEEE J. Select. Areas on Commun., vol. SAT-6, No.2, pp.353-363, Feb. 1988.

[7] V. Cuperman and A. Gersho, 'Vector Predictive Coding of Speech at 16kbits/s', IEEE Trans. on Commun., vol COM-33, No.7, pp.685-696

[8] S. Yeldener, A.M. Kondoz, and B.G. Evans, 'High Quality Multiband LPC Coding of Speech at 2.4kbit/s', Electronics Letters, Vol. 27, No. 14, pp. 1287-1289.

[9] F.F. Tzeng 'Analysis-by-Synthesis Linear Predictive Speech Coding at 2.4 kbit/s', Proc. Int. Conf. Acoust., Speech, Signal Processing, Glasgow, UK. May 1989, pp. 34.4.1-34.4.5.

[10] T. Tremain, J. Campbell, and V. Welch, 'A 4.8 kbps Code Excited Linear Predictive Coder', Proc. of the Mobile Satellite Conf., 1988, pp491-496.

[11] Sadaoki Furui, *Digital Speech, Processing, Synthesis, and Recognition*, Marcel Dekker, Inc., New York and Basel, 1988.

[12] K.A. Zeger and A. Gersho, 'Zero Redundancy Channel Coding in Vector Quantization', Electronic Letters, Vol. 23, pp654-655, May 1987

[13] J.R.B. Marca and N.S. Jayant, 'An Algorithm for Assigning Binary Indices to the Codevectors of a Multi-Dimentional Quantizer', Proc. of ICC'87, pp.32.2.1-32.2.5, 1987.

[14] J.R.B. De Marca, N. Farvardin, and N.S. Jayant, 'Robust Vector Quantization for Noisy Channels', Proc. of ICASSP'88, pp515-520, 1988.

[15] J.R. Campbell, Jr., V.C. Welch, T.E. Tremain, 'An Expandable Error-Protected 4800 BPS CELP Coder (U.S. Federal Standard 4800 BPS)', Proc. of ICASSP, 1989, pp735-738.

[16] Cox R., W. Kleijn, and P. Kroon, 'Robust CELP Coders for Noisy Background and Noisy Channels', Proc. of ICASSP, 1989 pp739-742.

[17] D.J. Goodman and C.E. Sundberg, 'Combined Source and Channel Coding for Variable-Bit-rate Speech Transmission', B.S.T.J., Vol.62, No.7, September 1983, pp2017-2036.

[18] H. Suda and T. Miki, 'An Error Protected 16 kbits/s Voice Transmission for Land Mobile Radio Channel', IEEE J. of Selected Areas on Commun., Vol.6, No.2, 1988, pp346-352.

[19] Cox R., J. Hagenauer, N. Seshadri, and C. Sundberg, 'A Sub-Band Coder Designed for Combined Source and Channel Coding', Proc. of ICASSP, 1988, pp235-238.

[20] G. Yang, P. Ho, and V. Cuperman, 'Error Protection for A 4.8 KBPS VQ Based CELP Coder', Proc. of IEEE Vehicular Technology Conf., Florida, pp.726-731, May 1990.

[21] H. Shi, P. Ho, and V. Cuperman, 'A Combined CELP/Reed-Solomon Codec for Mobile Radio Applications', Proc. of IEEE Vehicular Technology Conf., Denver, pp187-191, May 1992.

[22] C.E. Shannon, 'A Mathematical Theory of Communication', Bell System Tech., J.27(1948): 379-423, 623-656.

[23] C.E. Shannon, 'Coding Theorems for a Discrete Source with a Fidelity Criterion', IRE Nat. Conv. Rec., pp142-163, 1959.

[24] Electronic Industries Association EIA, 'Cellular System', Report IS-54, December 1989.

[25] R.G. Vaughan, 'Signals in Mobile Communications: A Review', IEEE Trans. on Vehicular Technology, Vol. VT-35, No.4, pp133-145.

[26] G.C. Clark and Jr.J.B. Cain, *Error Correction Coding for Digital Communications*, Plenum Press, New York, 1981.

[27] A.M. Michelson and A.H. Levesque, *Error-Control Techniques for Digital Communication*, John Willy Sons, Inc, 1985.

[28] J.J. Stone, 'Multiple-Burst Error Correction with the Chinese Remainder Theorem', SIAM J. Appl. Math., Vol.11, Mar. 1963, pp74-81.

[29] D. Mandelbaum, 'A Method of Coding for Multiple Errors', IEEE Trans. Inform. Theory, Vol. IT-14, May 1968, pp518-521.

[30] D. Mandelbaum, 'On Decoding of Reed-Solomon Codes', IEEE Trans. Inform. Theory, Vol. IT-17, No.6, 1971, pp707-712.

[31] J.P. Campbell, Jr. T.E. Tremain, and V.C. Welch, 'The DOD 4.8 KBPS Standard (Proposed Federal Standard 1016)'.

[32] Nariman Farvardin and Vinay Vaishampayan, 'Optimal Quantizer Design for Noisy Channels: An Approach to Combined Source-Channel Coding', IEEE Trans. on Information Theory, Vol. 33, No.6, November 1987, pp827-838.

[33] James W. Modestino and David G. Daut, 'Combined Source-Channel Coding of Images', IEEE Trans. on Commun., Vol. Com-27, No.11, November 1979, pp1644-1659.

[34] Richard E. Blahut, *Principles and Practice of Information Theory*, Addison-Wesley Publishing Company, 1987.

[35] Andrew J. Viterbi and Jim K. Omura, *Principles of Digital Communication and Coding*, MaGraw-Hill Inc., 1979.

[36] N.S. Jayant and P. Noll, *Digital Coding of Waveforms*, Englewood Cliffs, NJ: Prentice-Hall, 1984.

[37] L.R. Rabiner and R.W. Schafer, *Digital Processing of Speech Signals*, Englewood Cliffs, NJ: Prentice-Hall, 1978.

[38] Itakura, F. and S. Saito, 'Analysis synthesis telephony based on the maximum likelihood method', Proc. 6th Int. Cong. Acoust., 1968.

[39] B.S. Atal and M.R. Schroeder, 'Predictive Coding of Speech Signals', Proc. Conf. Commun. Pressing, Nov 1967, pp306-361.

[40] V. Cuperman and A. Gersho, 'Adaptive differential vector coding of speech', Conf. Rec., IEEE Global Commun. Conf., Miami, FL, 1982, E6.6.1 - E6.6.5.

[41] W.R. Daumer, 'Subjective Evaluation of Several Efficient Speech Coders', IEEE Trans. Commun., Vol. Com-30, pp655-662, April 1982.

[42] Schuyler R. Quackenbush, Thomas P. Barnwell III, Mark A. Clements, *Objective Measurements of Speech Quality*, Printice Hall, Englewood Cliffs, New Jersey, 1988.

[43] S. Chan and V. Cuperman, 'The Noise Performance of A CELP Speech Coder Based on Vector Quantization', Conf. Records, Canadian Conference on Electrical and Computer Engineering, pp795-798, Nov. 1988.

[44] John G. Proakis, *Digital Communications*, Second Edition, McGraw-Hill Book Company, 1989.

[45] John M. Wozencraft and Irwin M. Jacobs, *Principles of Communication Engineering*, John Wiley & Sons, 1965.

[46] Y. Wu and P. Ho, 'Trellis-Coded DPSK for Mobile Fading Channels', Proc. of IEEE Pacific Rim Conf. on Comm., Computer and Signal Processing, June 1989, pp400-403.

[47] A.J. Kurtenbach and P.A. Wintz, 'Quantizing for Noisy Channels', IEEE Trans. on Commun. Tech., Vol.COM-17, N0.2, April 1969.

[48] Nils Rydbeck and Carl-Erik W. Sundberg, 'Analysis of Digital Errors in Nonlinear PCM Systems', IEEE Trans. on Commun., Vol.COM-24, No.1, Jan.1976, pp59-65.

[49] Kenneth A. Zeger and Allen Gersho, 'Vector Quantizer Design for Memoryless Noisy Channels', Proc. of ICASSP, 1988, pp1593-1597.

[50] W.B. Kleijn, 'Optimal Codes to Protect CELP Against Channel Errors', Program and Abstracts of IEEE Workshop on Speech Coding for Telecommunications', Vancouver, Sept. 1989, pp33-34.

[51] G.D. Forney, Jr., 'Convolutional Codes I: Algebraic Structure', IEEE Trans. Inform. Theory, Vol. IT-16, pp720-738, Nov. 1970.

[52] J. Hagenauer, 'Rate-Compatible Punctured Convolutional Codes (RCPC Codes) and their Applications', IEEE Trans. on Commun., Vol. 36, No.4 pp389-400, Apr. 1988.

[53] Joachim Hagenauer, Nambirajan Seshadri, and Carl-Erik W. Sundberg, 'The Performance of Rate-Compatible Punctured Convolutional Codes for Digital Mobile Radio', IEEE Trans. Commun., Vol.38, No.7, pp966-980, July 1990.

[54] D.C. Bossen and S.S. Yau, 'Redundant Residue Polynomial Codes', Inform. Contr., Vol.13, Dec. 1968, pp597-681.

[55] James L. Massey, 'Shift-Register Synthesis and BCH Decoding', IEEE Trans. on Inform. Theory, Vol.IT-15, January 1960, pp122-127.

[56] Shu Lin, Daniel J. Costello, Jr., *Error Control Coding: Fundamentals and Applications*, Englewood Cliffs, N.J.:Prentice Hall, c1983.

[57] Y. Yasuda, K. Kashiki, and Y. Hirata, 'High Rate Punctured Convolutional codes for Soft Decision Viterbi Decoding', IEEE Trans. Commun., Vol. COM-32, March, 1984, pp315-319.

[58] A.J. Viterbi, 'Convolutional Codes and their Performance in Communication Systems', IEEE Trans. Commun. Technol., Vol COM-19, pp751-772, Oct. 1971.