

1

Introduction

Although data links are increasing in bandwidth and are becoming faster, speech communication is still the most dominant and common service in telecommunication networks. The fact that commercial and private usage of telephony in its various forms (especially wireless) continues to grow even a century after its first inception is obvious proof of its popularity as a form of communication. This popularity is expected to remain steady for the foreseeable future. The traditional plain analogue system has served telephony systems remarkably well considering its technological simplicity. However, modern information technology requirements have introduced the need for a more robust and flexible alternative to the analogue systems. Although the encoding of speech other than straight conversion to an analogue signal has been studied and employed for decades, it is only in the last 20 to 30 years that it has really taken on significant prominence. This is a direct result of many factors, including the introduction of many new application areas.

The attractions of digitally-encoded speech are obvious. As speech is condensed to a binary sequence, all of the advantages offered by digital systems are available for exploitation. These include the ease of regeneration and signalling, flexibility, security, and integration into the evolving new wireless systems. Although digitally-encoded speech possesses many advantages over its analogue counterpart, it nevertheless requires extra bandwidth for transmission if it is directly applied (without compression). The 64 kb/s Log-PCM and 32 kb/s ADPCM systems which have served the many early generations of digital systems well over the years have therefore been found to be inadequate in terms of spectrum efficiency when applied to the new, bandwidth limited, communication systems, e.g. satellite communications, digital mobile radio systems, and private networks. In these and other systems, the bandwidth and power available is severely restricted, hence signal compression is vital. For digitized speech, the signal compression is achieved via elaborate digital signal processing techniques that are facilitated by the

rapid improvement in digital hardware which has enabled the use of sophisticated digital signal processing techniques that were not feasible before. In response to the requirement for speech compression, feverish research activity has been pursued in all of the main research centres and, as a result, many different strategies have been developed for suitably compressing speech for bandwidth-restricted applications. During the last two decades, these efforts have begun to bear fruit. The use of low bit-rate speech coders has been standardized in many international, continental and national communication systems. In addition, there are a number of private network operators who use low bit-rate speech coders for specific applications.

The speech coding technology has gone through a number of phases starting with the development and deployment of PCM and ADPCM systems. This was followed by the development of good quality medium to low bit-rate coders covering the range from 16 kb/s to 8 kb/s. At the same time, very low bit-rate coders operating at around 2.4 kb/s produced better quality synthetic speech at the expense of higher complexity. The latest trend in speech coding is targeting the range from about 6 kb/s down to 2 kb/s by using speech-specific coders, which rely heavily on the extraction of speech-specific information from the input source. However, as the main applications of the low to very low bit-rate coders are in the area of mobile communication systems, where there may be significant levels of background noise, the accurate determination of the speech parameters becomes more difficult. Therefore the use of active noise suppression as a preprocessor to low bit-rate speech coding is becoming popular.

In addition to the required low bit-rate for spectral efficiency, the cost and power requirements of speech encoder/decoder hardware are very important. In wireless personal communication systems, where hand-held telephones are used, the battery consumption, cost and size of the portable equipment have to be reasonable in order to make the product widely acceptable.

In this book an attempt is made to cover many important aspects of low bit-rate speech coding. In Chapter 2, the background to speech coding, including the existing standards, is discussed. In Chapter 3, after briefly reviewing the sampling theorem, scalar and vector quantization schemes are discussed and formulated. In addition, various quantization types which are used in the remainder of this book are described.

In Chapter 4, speech analysis and modelling tools are described. After discussing the effects of windowing on the short-time Fourier transform of speech, extensive treatment of short-term linear prediction of speech is given. This is then followed by long-term prediction of speech. Finally, pitch detection methods, which are very important in speech vocoders, are discussed.

It is very important that the quantization of the linear prediction coefficients (LPC) of low bit-rate speech coders is performed efficiently both in terms of bit rate and sensitivity to channel errors. Hence, in Chapter 5, efficient quantization schemes of LPC parameters in the form of Line Spectral Frequencies are formulated, tested and compared.

In Chapter 6, more detailed modelling/classification of speech is studied. Various pitch estimation and voiced – unvoiced classification techniques are discussed.

In Chapter 7, after a general discussion of analysis by synthesis LPC coding schemes, code-excited linear prediction (CELP) is discussed in detail.

In Chapter 8, a brief review harmonic coding techniques is given.

In Chapter 9, a novel hybrid coding method, the integration of CELP and harmonic coding to form a multi-modal coder, is described.

Chapters 10 and 11 cover the topics of voice activity detection and speech enhancements methods, respectively.

