Chapter 1

Introduction

The main goals of this book are to provide a theoretical framework for joint source-channel coding and to improve communication systems for the transmission of continuous-amplitude signals such as speech, music, images, or videos. Due to the desired practical implementation of the systems by digital computers the signals are represented in discrete-time, which is always possible by appropriate low-pass-filtering and subsequent sampling.[1]

Practical systems will always have constraints on the delay between the input of a signal sample and its output at the receiver. The tolerable delay is determined by the application and by complexity constraints. As an example, let us consider the transmission of speech in GSM mobile radio: the speech signals are band-limited to less than 4 kHz and sampled with a period of $T_A = 125\ \mu$s (sampling frequency $f_A = 8$ kHz). The delay is critical in such an application since it strongly affects the quality of the conversation between two users. Therefore, it is limited to 20 ms, i.e., the maximum[2] number of samples that may be jointly processed is $N = 20\text{ms}/125\mu\text{s} = 160$. In non-conversational applications such as video streaming the delay constraints are much weaker, but still N is limited, possibly more because of memory limitations than by the requirements of the application itself.

For large N, information theory[3] implicitly suggests a way, of how to

[1]Since the word-length is also limited in digital computers, the raw signals after sampling are already quantized, but with high resolution; this allows to regard them as quasi-continuous.

[2]We have neglected in this discussion that in GSM the channel-coded bits of the source encoder are interleaved across two blocks, i.e., the truly occurring systematic delay is larger than 20 ms.

[3]In Appendix A, we summarize some results of information theory that are upper bounds for the best possible system performance, if there is no constraint on the block-length N. As these bounds are usually weak in our case of interest (we have to match

design a communication system. This approach is usually called the "separation principle." Roughly speaking it states that, for infinitely large block-lengths N, the transmitter can be separated into source and channel coding without loss in performance. The source encoder produces a sequence of independent and uniformly distributed bits (at a bit rate that equals the capacity of the channel) and the channel code protects the bits against the channel noise. Error correction is perfectly possible (in theory) if the code rate of the channel code (the number of data bits divided by the number of channel code bits) is not larger than the channel capacity. Except for the data bits, no additional communication between the source and channel encoders and decoders is required, as both components work perfectly.[4] Thus, the design of a communication system is greatly simplified, as, e.g., the design of a channel code is not influenced at all by the type of data at the input of the source encoder.

With the invention of turbo codes [Berrou and Glavieux (1996); Hagenauer *et al.* (1996)] and the developments based thereon, some promises of information theory have more or less become practice. For systems with large block-lengths, turbo codes are able to reduce the residual bit error rate after channel decoding down to 10^{-5} and less, at a channel quality that is close to the theoretical limits for error-free transmission.

If, however, the block lengths are moderate or even short (as, e.g., in mobile telephony), both the best known source and channel coding schemes don't even work "almost" perfectly. Hence, there are strong potential gains in transmission quality if the imperfectness of each system component is exploited by all others; that is why we deal with joint source-channel coding.

In Chapter 2 we give an overview over joint source-channel coding and summarize some approaches, especially for the decoder side. In Chapter 3 we state the theory of (near-)optimum joint source-channel *d*ecoding and an approximation called iterative source-channel decoding. We show that, in principle, it is possible find the optimal receiver for any given transmitter; the main issue in joint *de*coding is complexity reduction. In contrast to that, the general solution for the best transmitter for a given source and channel under given delay constraints is unknown. It is, however, crucial to deal with the encoder design, as even a perfect receiver for a badly chosen

delay constraints), we deliberately put this chapter into the Appendix. Analytical results for the best possible performances of delay-constrained systems are, unfortunately, widely unknown, so often the asymptotic information-theoretic results are the best we have for a comparison.

[4]One could also argue that this is the only "static" information the system components need to know about each other.

transmitter will lead to an overall bad system performance. Since a general solution is hard to obtain, we deal in the Chapters 4–7 with some special cases in which, due to additional practical constraints, "good" transmitter designs can be found.