Delay-Constrained Coding and Transmission of Continuous-Amplitude Sources

Norbert Görtz
Institute for Communications Engineering
Munich University of Technology
80290 Munich, Germany
norbert.goertz@ei.tum.de

Abstract— **Delay-constrained coding and transmission of continuous-amplitude source signals over noisy channels is considered. The associated optimization problem for the encoder/decoder-design is formulated and it is shown, that its direct solution is intractable in most relevant cases. Therefore, some approaches for the practical system design are discussed.**

I. Introduction and System Model

We will consider the basic model of a communication system that is depicted in Fig. 1. The goal is to transmit the input

Fig. 1. System model

source signal, the N-dimensional source vector X, to a destination, where we want to obtain a reproduction $\tilde{X}$ with the highest quality possible. The maximum block length N is limited due to delay constraints. The source vector components x_l, $l = 0, 1, ..., N-1$, have continuous amplitudes (they may, e.g., be samples of a speech signal), and we will use the mean squared error

$$d(x, \tilde{x}) = \frac{1}{N} \sum_{l=0}^{N-1} \left(x_l - \tilde{x}_l\right)^2 \qquad (1)$$

as a distance measure. In what follows, we will use capital letters for random variables (excluding the system constants N, K and the system distortion D) and small letters for their realizations. We will assume that the vector dimensions N and K are fixed and known, and our goal is to find an encoder-decoder pair that minimizes the expected system distortion

$$D \doteq E\{d(X, \tilde{X}\} \; . \qquad (2)$$

The probability density function $p(X = x)$ of the input vector is assumed to be known.

A. Channel

The transmission is carried out over a discrete-time channel, which is used K times to transmit the N source vector components; thus, the channel input signal in our model is the K-dimensional vector Z. Where it is necessary, the input alphabet Ω_Z and the output alphabet $\Omega_{\tilde{Z}}$ of the channel are specified below; however, as an example, we may think of a K-bit channel input and K-dimensional continuous-valued output from an additive white Gaussian noise channel.

We will assume, that the noise on the channel is stationary, memoryless, and independent of all other signals and the components of the system. The conditional probability density function $p(\tilde{Z} = \tilde{z} \mid Z = z)$, which describes the "probability" of the output vector $\tilde{z}$, given the input vector z, is assumed to be known.

B. Encoder

The encoder is a device that maps the input source-signal vector to the input of the channel. The encoder is described by the mapping

$$Z = \varepsilon(X) \; , \qquad (3)$$

with $X \in \Omega_X = I\!R^N$ and $Z \in \Omega_Z$. The encoder output may have to fulfill a power constraint, for instance

$$\frac{1}{K} \, E\{\|Z\|^2\} \leq P_{max} \; , \qquad (4)$$

i.e., on average, the transmission power may be P_{max} for each channel-use.

C. Decoder

The decoder is a device that maps the output $\tilde{Z}$ of the channel to the decoder output signal, which should be a good estimate of what has been transmitted. The decoder mapping is

$$\tilde{X} = \varrho(\tilde{Z}) \; . \qquad (5)$$

with $\tilde{X} \in \Omega_X = I\!R^N$ and $\tilde{Z} \in \Omega_{\tilde{Z}}$. Since we want to obtain a good reproduction $\tilde{X}$ of the input X in the mean-square-sense, the output alphabet of the decoder equals that of the encoder input. The alphabet $\Omega_{\tilde{Z}}$ of the channel output, however, may be different from Ω_Z.

II. System Distortion

The expected system distortion is given by

$$D(\varepsilon, \varrho) \doteq \int_{\Omega_{\tilde{Z}}} \int_{\Omega_X} d\big(x, \tilde{x} = \varrho(\tilde{z})\big) \cdot p(\tilde{Z} = \tilde{z}, Z = \varepsilon(x)) \, dx d\tilde{z} \; . \qquad (6)$$

Clearly, both the encoder mapping ε and the decoder mapping ϱ are required to compute (6). In what follows, our goal is to choose both mappings such, that the expected distortion D is minimized. For this, it is useful to introduce an alternative representation of D. We may write:

$$p(\tilde{Z} = \tilde{z}, Z = \varepsilon(x)) \; = \; p(Z = \varepsilon(x) \mid \tilde{Z} = \tilde{z}) \cdot p(\tilde{Z} = \tilde{z}) \; . \qquad (7)$$

By insertion of (7) into (6) we obtain

$$D(\varepsilon, \varrho) = \int_{\Omega_{\tilde{Z}}} D_D(\tilde{z}, \varepsilon, \varrho) \; \cdot \; p(\tilde{Z} = \tilde{z}) \, d\tilde{z} \; , \qquad (8)$$

with

$$D_D(\tilde{z}, \varepsilon, \varrho) \doteq \int_{\Omega_X} d\big(x, \varrho(\tilde{z})\big) \; \cdot \; p(Z = \varepsilon(x) \mid \tilde{Z} = \tilde{z}) \, dx \; . \qquad (9)$$

The quantity defined in (9) is the conditional expected distortion, given a particular channel output $\tilde{z}$.

III. Optimal Decoder for a Given Encoder

In a first step of the system optimization, our goal is to find the optimal decoder mapping $\varrho^{\circledast}$ for a given encoder mapping ε. Since the distance measure $d(x,\tilde{x})$ in (6) is, for any combination of x and $\tilde{x}$, non-negative by definition, the distortion $D(\varepsilon,\varrho)$ is minimized, if $D_D(\tilde{z},\varepsilon,\varrho)$ is minimized for each particular channel output vector $\tilde{z}$. Thus, the optimal decoder is given by

$$\varrho^{\circledast}(\tilde{z},\varepsilon) = \arg\min_{\varrho} D_D(\tilde{z},\varepsilon,\varrho)\ . \quad (10)$$

With the mean squared error as a distance measure in the source signal space, the solution of (10) is the well-known mean-square estimator

$$\varrho^{\circledast}(\tilde{z},\varepsilon) = E\{X|\tilde{Z}=\tilde{z}\} = \int_{\Omega_X} x\, p\big(X=x \mid \tilde{Z}=\tilde{z}\big) dx\ . \quad (11)$$

By use of the Bayes rule one can show, that (11) equals

$$\varrho^{\circledast}(\tilde{z},\varepsilon) = A \int_{\Omega_X} x \cdot \underbrace{p\big(\tilde{Z}=\tilde{z} \mid Z=\varepsilon(x)\big)}_{\text{conditional channel pdf}} \cdot \underbrace{p(X=x)}_{\text{source pdf}}\ dx\ . \quad (12)$$

with

$$\frac{1}{A} \doteq p\big(\tilde{Z}=\tilde{z}\big) = \int_{\Omega_X} p\big(\tilde{Z}=\tilde{z} \mid Z=\varepsilon(x)\big) \cdot p(X=x)\ dx\ . \quad (13)$$

For any deterministic choice of the encoder mapping ε, (12) defines the optimal decoder, which minimizes the expected system distortion. The optimality of the decoder mapping, however, does not guarantee a good system performance, because the encoder mapping might be bad a choice.

IV. Optimal Encoder

Now that we found the optimal decoder for any encoder, we will use it and try to optimize the encoder mapping. Since our goal is to minimize the expected system distortion, the general optimization problem is given by

$$\varepsilon^{\circledast} = \arg \min_{\varepsilon: E\{\|\varepsilon(X)\|^2\} \leq P_{max}\cdot K} D(\varepsilon, \varrho^{\circledast}(\varepsilon))\ . \quad (14)$$

In most practically relevant situations, the solution of (14) is intractable, even by numerical methods. In some special cases, however, the problem can be solved.

A. Gaussian Source and Gaussian Channel

Assume that we want to transmit an uncorrelated discrete-time Gaussian source signal with the variance σ_X^2. We may use the channel exactly once per source sample, i.e., $N=K$, and the average input power of the channel (per use) is limited to $\frac{1}{K}E\{\|Z\|^2\} = P_{max}$. The channel adds uncorrelated zero-mean Gaussian noise samples, which have the variance σ_n^2. Fig. 2 shows the optimal system [1]: the input samples—the vector components of X—are individually scaled by $\sqrt{P_{max}}/\sigma_X$ (encoder mapping) and each channel output sample is individually scaled by $\sigma_X/\sqrt{P_{max}}/\big(1+\sigma_n^2/P_{max}\big)$ (decoder mapping). The resulting minimum distortion equals $D=\sigma_X^2/(1+P_{max}/\sigma_n^2)$.

Surprisingly, this system is optimal for any choice of N, even for $N\rightarrow\infty$. No amount of source and channel coding of long blocks (large N) could improve the system [1]. If, however, any of the requirements on the source and the channel is not fulfilled, then the system is no longer optimal and complex coding schemes with long block lengths are required.

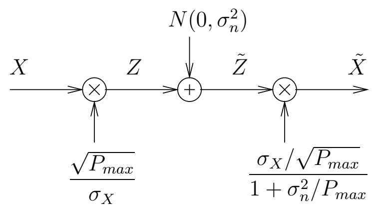

Fig. 2. Optimal encoder-decoder pair for the transmission of a Gaussian source over a Gaussian channel with average input power constraint.

B. Binary Symmetric Channel

If we assume a binary symmetric channel we have N continuous-amplitude source vector components that have to be mapped to K bits by the encoder; the set of possible channel-input values is given by $\Omega_Z=\{0,1\}^K$. Hence, the number of possible channel inputs is limited to $M=2^K$. This means that the source signal has to be quantized, i.e., the source signal space Ω_X has to be partitioned into M disjoint subsets $\Omega_X^{(j)}$, $j=0,1,...,M-1$. If the partition regions $\Omega_X^{(j)}$ are known, which we will assume for the moment, the encoder mapping is simply given by

$$z=\varepsilon(x)=j \quad \text{if} \quad x\in\Omega_X^{(j)}\quad . \quad (15)$$

The binary representation of the number j of the subset that contains the current source vector is transmitted over the channel. At its output, a K-dimensional bitvector is received, i.e., the channel-output alphabet equals the input alphabet, but due to "channel-noise" some bits may have been flipped.

The application of (12) leads to the optimal decoder

$$\tilde{x}^{\circledast}(k) = \varrho^{\circledast}(k,\varepsilon) = \frac{\sum_{j=0}^{M-1} P\big(\tilde{Z}=k \mid Z=j\big)\cdot P(Z=j)\cdot y_j}{\sum_{j\prime=0}^{M-1} P\big(\tilde{Z}=k \mid Z=j\prime\big)\cdot P(Z=j\prime)}\ , \quad (16)$$

with $k=0,1,...,M-1$,

$$P(Z=j) = \int_{\Omega_X^{(j)}} p(X=x)\ dx \quad (17)$$

and

$$y_j = \int_{\Omega_X^{(j)}} x\cdot p(X=x)\ dx \Big/ P(Z=j)\ . \quad (18)$$

Since the optimal decoder produces only a limited number of outputs $\tilde{x}^{\circledast}(k)$, the optimal decoding results (16) may be precomputed and stored in a table—the so-called codebook; thus, if some k appears at the channel output it can be decoded by a simple table-lookup.

The implementation of the encoder (15) is a problem in practice because the mathematical descriptions of the partitions $\Omega_X^{(j)}$ are usually hard to handle. This problem is solved as follows: the expected system distortion can be written as

$$D(\varepsilon,\varrho^{\circledast}) = \int_{\Omega_X} D_E(x,\varepsilon,\varrho^{\circledast})\ \cdot\ p(X=x)\ dx\ , \quad (19)$$

with

$$D_E(x,\varepsilon,\varrho^{\circledast}) \doteq \sum_{k=0}^{M-1} d\big(x,\varrho^{\circledast}(k,\varepsilon)\big)\ \cdot\ P\big(\tilde{Z}=k \mid Z=\varepsilon(x)\big)\ . \quad (20)$$

Since $D_E(x,\varepsilon,\varrho^{\circledast})$ is positive for every x, we will minimize $D(\varepsilon,\varrho^{\circledast})$ if we minimize $D_E(x,\varepsilon,\varrho^{\circledast})$. Thus, encoding can also be carried out by

$$\varepsilon(x) = \arg \min_{l=0,1,\ldots,M-1} \underbrace{\sum_{k=0}^{M-1} d\big(x, \varrho^{\circledast}(k,\varepsilon)\big) \cdot P(\tilde{Z}=k \mid Z=l)}_{\doteq D_E(x,\varepsilon,\varrho^{\circledast},l)} \; . \tag{21}$$

This simply means that we compute the distortion D_E for all possible channel-inputs $Z=l$ and we select the "index" l with the lowest value of D_E for the transmission. This way, we can also give an alternative description of the partition regions of the source signal space, that are implicitly used:

$$\Omega_X^{(j)} = \{x \in \Omega_X : j = \arg \min_{l=0,1,\ldots,M-1} D_E(x,\varepsilon,\varrho^{\circledast},l)\} \; . \tag{22}$$

Up to now, the partition regions $\Omega_X^{(j)}$, and thus the encoder mapping, have been assumed to be known. Now we will describe how to optimize them by an iterative procedure. Initially, let us assume that we have a large set of training source vectors available, and that we randomly select an initial codebook $\tilde{x}(k)$, $k=0,1,...,M-1$ (the initial codebook is needed, because we don't know the partition regions $\Omega_X^{(j)}$ and, thus, we cannot compute (16)). In the first step, (21) is used to encode all training vectors. In the second step, (17) and (18) are approximately computed by averaging over the training vectors that have been assigned to the partitions in the first step and a new "codebook" $\tilde{x}(k)$, $k=0,1,...,M-1$, is computed by (16). Step 1 and step 2 are iteratively repeated until the total distortion does not decrease further; this way, a local minimum of the expected distortion is found.

The scheme described above is called Channel-Optimized Vector Quantization (COVQ); details are given in [2]. It should be noticed that, due to the complexity of encoding and the memory requirements for the codebook, this scheme is only applicable, if the number K of bits is small, e.g., $K < 10$. Therefore, COVQ cannot be directly used, e.g,. for speech, audio and image coding, because at a "typical" bit rate of 1 bit per sample the maximum block-length would only be 10 samples: unfortunately, source coding at that rate with sufficient quality is only possible with block lengths of a few hundred samples. Nevertheless, COVQ is a very interesting concept, because it is the only non-trivial scheme, where the encoder mapping is *not* divided into a cascade of partial mappings.

V. Practical Approaches to Source-Channel Coding

A. Separation of Source and Channel Coding

The foundation for the structure of today's systems is Information Theory [3]: roughly speaking, it guarantees that by use of an infinitely large block length N one may replace the encoder and the decoder mappings in Fig. 1 by cascades of mappings with binary interfaces, at which independent and uniformly distributed bits are exchanged—without any loss in performance compared with the best possible direct mapping. The basic notion is to apply channel coding to achieve error-free bits at a bit rate that equals the channel capacity. Then, a source code is applied, to reduce the number of bits required to represent the source signal to the amount that is available on the channel; clearly, this reduction is only possible at the price of some distortion that is imposed on the source signal. It should be noticed, that the problem stated in Section IV-A, with its simple solution depicted in Fig. 2, could also be solved by the separation into source and channel coding—at the price of infinite block length and complexity. Thus, the blindfold adoption of asymptotic results from Information Theory may turn out to be a bad choice in practice. Nevertheless, with some modification discussed below, source and channel coding are separate processing steps in all practical system. The reason is that the separation allows to construct partial mappings, e.g., channel encoders/decoders, that, compared with asymptotic results, sometimes may have only moderate performance,[1] but the complexity is tolerable.

If the separation principle is applied for limited block lengths (usually caused by delay-constraints) the following problems occur: (i) it is impossible to realize channel codes that have a residual bit error rate of zero, (ii) practical source encoders don't produce independent and uniformly distributed bits; moreover, the distortion at the source decoder output depends on which bit is in error, and (iii) the source encoder output bits are not independent and uniformly distributed, i.e., the coded bits contain residual redundancies.

B. Unequal Error Protection and Error Concealment

Fig. 3 shows some system modifications that account for the issues stated above. Implicitly, a channel with digital input

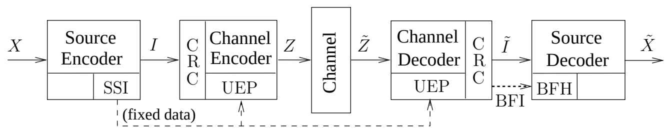

Fig. 3. Unequal error protection and error concealment

(e.g., by application of a digital modulation scheme) is assumed, because the channel codeword Z—a bit vector—is directly fed into the channel. The output bits of the source encoder are not equally sensitive to bit errors, i.e., some of the data bits cause stronger quality degradations than others if they are in error but are used for decoding anyhow. The source significance information (SSI) that can be derived from that is used by the channel-coding scheme to apply stronger error protection to the more sensitive bits, e.g., by puncturing of those output bits of a low-rate channel code that protect the less sensitive bits.

The residual redundancies in the source-encoder output bits can be exploited to conceal bit-errors (bad frame handling (BFH)), e.g., by repetition of the old bits from the previous block; clearly, this only makes sense, if the bits are correlated in time. For the initiation of the error concealment, an error detection is required, which is usually realized by means of an error-detecting channel code (Cyclic Redundancy Check (CRC)). The strongly sensitive bits of the source encoder are CRC-encoded (Class-1-bits) and afterwards the CRC codeword is channel encoded. Thus, in combination with UEP, the bits of the source encoder are divided into several classes which are protected by individually adjustable amounts of channel-code redundancy.

Unequal error protection and error concealment are very effective and are frequently used, e.g., in every mobile radio standard.

C. Channel Decoding aided by Residual Redundancies

The basic scheme is depicted in Fig. 4. The idea is to exploit the redundancies as a-priori information in the decision procedure for the data bits in channel decoding to reduce the

[1] It should be mentioned, that, in the "waterfall-region", Turbo Codes [4] and the following developments brought the performance of channel coding schemes with fairly long block-lengths very close to the information theoretical bounds.

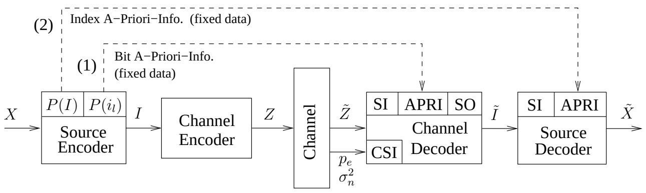

Fig. 4. (1) Channel decoding aided by residual redundancies and (2) estimation-based source decoding

residual bit error rate. This concept can be implemented very efficient for the decoding of convolutional codes; moreover, the idea can also be extended to soft-output decoders (APRI-SOVA, [5]), which supply reliability information—and not only hard decisions—for the data bits. These reliabilities can be exploited by estimation-based source decoders described below. For a correct weighting of the a-priori information and the soft channel-outputs, a channel state information (CSI) must be available.

A drawback is that the redundancies are usually exploited on bit-basis[2], although the source encoder often emits correlated *indices* that consist of several bits; thus, a large part of the redundancies is removed if the marginal distributions for bits are used as a-priori information.

D. Estimation-Based Source Decoding

In conventional systems, simple procedures, e.g., quantizer table-lookups, are frequently applied for source decoding. If, however, source encoding is not perfect and residual redundancies—e.g., correlations in time—are left in the data bits, this a-priori knowledge may be exploited for better source decoding by performing (optimal) estimations of the source-encoder input [7–9]; such a system is also depicted in Fig. 4.

Conceptually, the idea is similar to the one described in the previous section, but now it is possible to exploit the full correlations on index-basis. The soft-in/soft-out (SISO) channel decoder acts as a device that improves the reliability of the "virtual" channel for the source decoder. As in the previous section, bit-based a-priori information may be used to aid channel decoding, but it is not clear, how the a-priori information shall be optimally divided between the channel decoder and the source decoder.

E. Iterative Source-Channel Decoding

In [10–12] iterative source-channel decoding is introduced. Although the sketch of the system in Fig. 5 is similar to the systems in Fig. 4, the theoretical concept is different: the transmitter is interpreted as a serially concatenated channel-coding scheme; the constituent "codes" are the implicit residual redundancies within the source encoder indexes and the explicit redundancy of the channel code. As in all iterative decoding schemes [13], decoders for both constituent "codes" must be available that are able to exchange extrinsic information on the data bits within the iterations. While such decoders for convolutional channel codes are well-known from literature [14], they can also be formulated on the basis of estimation-based decoding for the source-coding-part of the system. It can be shown [10], that the iterative decoding scheme is a well-defined approximation of the optimal joint decoder, which would be prohibitively complex.

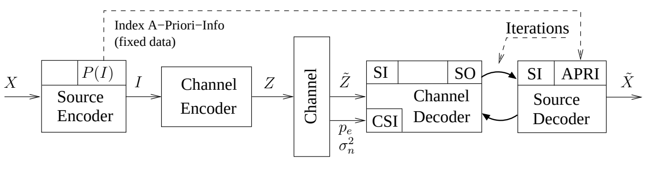

Fig. 5. Iterative source-channel decoding

The system concepts stated in Section V were given in chronological order and, at the same time, in order of increasing performance; some of them, however, may be combined as, for instance, unequal error protection and all other schemes.

VI. Conclusions

The problem of designing encoder-decoder pairs for the transmission of continuous-amplitude source signals with a delay-constraint was considered; the optimal solution turned out to be intractable (analytically and numerically). Therefore, the encoders and the decoders are split into cascades of partial mappings, which can be implemented with acceptable complexity; the concept is based on asymptotic results from information theory. Some modifications that partially rejoin the separated source and channel coders are necessary, however, to compensate for the imperfectness of the system components due to limited block lengths. Some practical approaches for joint source-channel coding were given; the more recent concepts concentrate on the close-to-optimum design of the decoder.

[2] In [6] a channel-decoding scheme for binary convolutional codes is given that is able to optimally exploit index-based a-priori information, if all the bits of an index are adjacently mapped into the input of the channel encoder. However, the method may not be applicable because usually a bit-reordering is employed at the channel-encoder input to account for different sensitivities of the bits.

References

[1] T. Berger and J. D. Gibson, "Lossy source coding," *IEEE Transactions on Information Theory*, vol. 44, pp. 2693–2723, Oct. 1998.

[2] N. Farvardin and V. Vaishampayan, "On the performance and complexity of channel-optimized vector quantizers," *IEEE Transactions on Information Theory*, vol. 37, pp. 155–160, Jan. 1991.

[3] R. J. McEliece, *The Theory of Information and Coding*. Cambridge: Cambridge University Press, second ed., 2002.

[4] C. Berrou and A. Glavieux, "Near optimum error correcting coding and decoding: Turbo-codes," *IEEE Transactions on Communications*, vol. 44, pp. 1261–1271, Oct. 1996.

[5] J. Hagenauer, "Source-controlled channel decoding," *IEEE Transactions on Communications*, vol. 43, pp. 2449–2457, Sept. 1995.

[6] S. Heinen, A. Geiler, and P. Vary, "MAP channel decoding by exploiting multilevel source a priori knowledge," in *Proceedings of the ITG-Fachtagung "Codierung für Quelle, Kanal und Übertragung"*, pp. 89–94, Mar. 1998.

[7] N. Phamdo and N. Farvardin, "Optimal detection of discrete Markov sources over discrete memoryless channels — applications to combined source-channel coding," *IEEE Transactions on Information Theory*, vol. 40, pp. 186–193, Jan. 1994.

[8] T. Fingscheidt and P. Vary, "Robust speech decoding: A universal approach to bit error concealment," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 3, pp. 1667–1670, Apr. 1997.

[9] M. Skoglund, "Soft decoding for vector quantization over noisy channels with memory," *IEEE Transactions on Information Theory*, vol. 45, pp. 1293–1307, May 1999.

[10] N. Görtz, "On the iterative approximation of optimal joint source-channel decoding," *IEEE Journal on Selected Areas in Communications*, vol. 19, pp. 1662–1670, Sept. 2001.

[11] N. Görtz, "A generalized framework for iterative source-channel decoding," *Annals of Telecommunications, Special issue on "Turbo Codes: a widespreading Technique"*, pp. 435–446, Jul./Aug. 2001.

[12] R. Perkert, M. Kaindl, and T. Hindelang, "Iterative source and channel decoding for GSM," in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, vol. 4, pp. 2649–2652, May 2001.

[13] J. Hagenauer, E. Offer, and L. Papke, "Iterative decoding of binary block and convolutional codes," *IEEE Transactions on Information Theory*, vol. 42, pp. 429–445, Mar. 1996.

[14] L. R. Bahl, J. Cocke, F. Jelinek, and J. Raviv, "Optimal decoding of linear codes for minimizing symbol error rate," *IEEE Transactions on Information Theory*, vol. IT-20, pp. 284–287, Mar. 1974.