Source Coding:

Part I of Fundamentals of Source and Video Coding

Source Coding:
Part I of Fundamentals of Source and Video Coding

Thomas Wiegand

Berlin Institute of Technology and Fraunhofer Institute for Telecommunications — Heinrich Hertz Institute
Germany
thomas.wiegand@tu-berlin.de

Heiko Schwarz

Fraunhofer Institute for Telecommunications — Heinrich Hertz Institute
Germany
heiko.schwarz@hhi.fraunhofer.de

the essence of knowledge

Boston – Delft

Foundations and Trends® in Signal Processing

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
USA
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is T. Wiegand and H. Schwarz, Source Coding: Part I of Fundamentals of Source and Video Coding, Foundations and Trends® in Signal Processing, vol 4, nos 1–2, pp 1–222, 2010

ISBN: 978-1-60198-408-1

Foundations and Trends® in Signal Processing
Volume 4 Issues 1–2, 2010

Editorial Scope

Foundations and Trends® in Signal Processing will publish survey and tutorial articles on the foundations, algorithms, methods, and applications of signal processing including the following topics:

- Adaptive signal processing
- Audio signal processing
- Biological and biomedical signal processing
- Complexity in signal processing
- Digital and multirate signal processing
- Distributed and network signal processing
- Image and video processing
- Linear and nonlinear filtering
- Multidimensional signal processing
- Multimodal signal processing
- Multiresolution signal processing
- Nonlinear signal processing
- Randomized algorithms in signal processing
- Sensor and multiple source signal processing, source separation
- Signal decompositions, subband and transform methods, sparse representations
- Signal processing for communications
- Signal processing for security and forensic analysis, biometric signal processing
- Signal quantization, sampling, analog-to-digital conversion, coding and compression
- Signal reconstruction, digital-to-analog conversion, enhancement, decoding and inverse problems
- Speech/audio/image/video compression
- Speech and spoken language processing
- Statistical/machine learning
- Statistical signal processing
- classification and detection
- estimation and regression
- tree-structured methods

Information for Librarians

Foundations and Trends® in Signal Processing, 2010, Volume 4, 4 issues. ISSN paper version 1932-8346. ISSN online version 1932-8354. Also available as a combined paper and online subscription.

Foundations and Trends® in
Signal Processing
Vol. 4, Nos. 1–2 (2010) 1–222

DOI: 10.1561/2000000010

Source Coding: Part I of Fundamentals of Source and Video Coding

Thomas Wiegand[1] **and Heiko Schwarz**[2]

[1] *Berlin Institute of Technology and Fraunhofer Institute for Telecommunications — Heinrich Hertz Institute, Germany, thomas.wiegand@tu-berlin.de*

[2] *Fraunhofer Institute for Telecommunications — Heinrich Hertz Institute, Germany, heiko.schwarz@hhi.fraunhofer.de*

Abstract

Digital media technologies have become an integral part of the way we create, communicate, and consume information. At the core of these technologies are source coding methods that are described in this monograph. Based on the fundamentals of information and rate distortion theory, the most relevant techniques used in source coding algorithms are described: entropy coding, quantization as well as predictive and transform coding. The emphasis is put onto algorithms that are also used in video coding, which will be explained in the other part of this two-part monograph.

Contents

1

Introduction

The advances in source coding technology along with the rapid developments and improvements of network infrastructures, storage capacity, and computing power are enabling an increasing number of multimedia applications. In this monograph, we will describe and analyze fundamental source coding techniques that are found in a variety of multimedia applications, with the emphasis on algorithms that are used in video coding applications. The present first part of the monograph concentrates on the description of fundamental source coding techniques, while the second part describes their application in modern video coding.

The block structure for a typical transmission scenario is illustrated in Figure 1.1. The *source* generates a signal s. The *source encoder* maps the signal s into the bitstream b. The bitstream is transmitted over the *error control channel* and the received bitstream b' is processed by the *source decoder* that reconstructs the decoded signal s' and delivers it to the *sink* which is typically a human observer. This monograph focuses on the source encoder and decoder parts, which is together called a source codec.

The error characteristic of the digital channel can be controlled by the *channel encoder*, which adds redundancy to the bits at the source

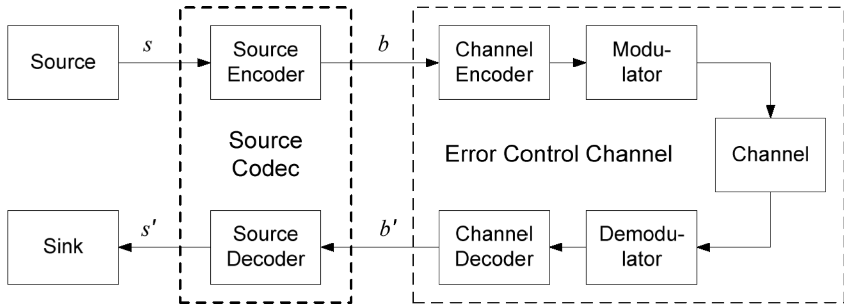

Fig. 1.1 Typical structure of a transmission system.

encoder output b. The *modulator* maps the channel encoder output to an analog signal, which is suitable for transmission over a physical *channel*. The *demodulator* interprets the received analog signal as a digital signal, which is fed into the *channel decoder*. The channel decoder processes the digital signal and produces the received bitstream b', which may be identical to b even in the presence of channel noise. The sequence of the five components, channel encoder, modulator, channel, demodulator, and channel decoder, are lumped into one box, which is called the error control channel. According to Shannon's basic work [63, 64] that also laid the ground to the subject of this text, by introducing redundancy at the channel encoder and by introducing delay, the amount of transmission errors can be controlled.

1.1 The Communication Problem

The basic communication problem may be posed as conveying source data with the highest fidelity possible without exceeding an available bit rate, or it may be posed as conveying the source data using the lowest bit rate possible while maintaining a specified reproduction fidelity [63]. In either case, a fundamental trade-off is made between bit rate and signal fidelity. The ability of a source coding system to suitably choose this trade-off is referred to as its *coding efficiency* or *rate distortion performance*. Source codecs are thus primarily characterized in terms of:

- *throughput of the channel*: a characteristic influenced by the transmission channel bit rate and the amount of protocol and error-correction coding overhead incurred by the transmission system; and

- *distortion of the decoded signal*: primarily induced by the source codec and by channel errors introduced in the path to the source decoder.

However, in practical transmission systems, the following additional issues must be considered:

- *delay*: a characteristic specifying the start-up latency and end-to-end delay. The delay is influenced by many parameters, including the processing and buffering delay, structural delays of source and channel codecs, and the speed at which data are conveyed through the transmission channel;
- *complexity*: a characteristic specifying the computational complexity, the memory capacity, and memory access requirements. It includes the complexity of the source codec, protocol stacks, and network.

The practical source coding design problem can be stated as follows:

> *Given a maximum allowed delay and a maximum allowed complexity, achieve an optimal trade-off between bit rate and distortion for the range of network environments envisioned in the scope of the applications.*

1.2 Scope and Overview of the Text

This monograph provides a description of the fundamentals of source and video coding. It is aimed at aiding students and engineers to investigate the subject. When we felt that a result is of fundamental importance to the video codec design problem, we chose to deal with it in greater depth. However, we make no attempt to exhaustive coverage of the subject, since it is too broad and too deep to fit the compact presentation format that is chosen here (and our time limit to write this text). We will also not be able to cover all the possible applications of video coding. Instead our focus is on the source coding fundamentals of video coding. This means that we will leave out a number of areas including implementation aspects of video coding and the whole subject of video transmission and error-robust coding.

The monograph is divided into two parts. In the first part, the fundamentals of source coding are introduced, while the second part explains their application to modern video coding.

Source Coding Fundamentals. In the present first part, we describe basic source coding techniques that are also found in video codecs. In order to keep the presentation simple, we focus on the description for one-dimensional discrete-time signals. The extension of source coding techniques to two-dimensional signals, such as video pictures, will be highlighted in the second part of the text in the context of video coding. Section 2 gives a brief overview of the concepts of probability, random variables, and random processes, which build the basis for the descriptions in the following sections. In Section 3, we explain the fundamentals of lossless source coding and present lossless techniques that are found in the video coding area in some detail. The following sections deal with the topic of lossy compression. Section 4 summarizes important results of rate distortion theory, which builds the mathematical basis for analyzing the performance of lossy coding techniques. Section 5 treats the important subject of quantization, which can be considered as the basic tool for choosing a trade-off between transmission bit rate and signal fidelity. Due to its importance in video coding, we will mainly concentrate on the description of scalar quantization. But we also briefly introduce vector quantization in order to show the structural limitations of scalar quantization and motivate the later discussed techniques of predictive coding and transform coding. Section 6 covers the subject of prediction and predictive coding. These concepts are found in several components of video codecs. Well-known examples are the motion-compensated prediction using previously coded pictures, the intra prediction using already coded samples inside a picture, and the prediction of motion parameters. In Section 7, we explain the technique of transform coding, which is used in most video codecs for efficiently representing prediction error signals.

Application to Video Coding. The second part of the monograph will describe the application of the fundamental source coding techniques to video coding. We will discuss the basic structure and the basic concepts that are used in video coding and highlight their application in modern video coding standards. Additionally, we will consider advanced encoder optimization techniques that are relevant for achieving a high coding efficiency. The effectiveness of various design aspects will be demonstrated based on experimental results.

1.3 The Source Coding Principle

The present first part of the monograph describes the fundamental concepts of source coding. We explain various known source coding principles and demonstrate their efficiency based on one-dimensional model sources. For additional information on information theoretical aspects of source coding the reader is referred to the excellent monographs in [4, 11, 22]. For the overall subject of source coding including algorithmic design questions, we recommend the two fundamental texts by Gersho and Gray [16] and Jayant and Noll [40].

The primary task of a source codec is to represent a signal with the minimum number of (binary) symbols without exceeding an "acceptable level of distortion", which is determined by the application. Two types of source coding techniques are typically named:

- *Lossless coding*: describes coding algorithms that allow the exact reconstruction of the original source data from the compressed data. Lossless coding can provide a reduction in bit rate compared to the original data, when the original signal contains dependencies or statistical properties that can be exploited for data compaction. It is also referred to as *noiseless coding* or *entropy coding*. Lossless coding can only be employed for discrete-amplitude and discrete-time signals. A well-known use for this type of compression for picture and video signals is JPEG-LS [35].
- *Lossy coding*: describes coding algorithms that are characterized by an irreversible loss of information. Only an approximation of the original source data can be reconstructed from

the compressed data. Lossy coding is the primary coding type for the compression of speech, audio, picture, and video signals, where an exact reconstruction of the source data is not required. The practically relevant bit rate reduction that can be achieved with lossy source coding techniques is typically more than an order of magnitude larger than that for lossless source coding techniques. Well known examples for the application of lossy coding techniques are JPEG [33] for still picture coding, and H.262/MPEG-2 Video [34] and H.264/AVC [38] for video coding.

Section 2 briefly reviews the concepts of probability, random variables, and random processes. Lossless source coding will be described in Section 3. Sections 5–7 give an introduction to the lossy coding techniques that are found in modern video coding applications. In Section 4, we provide some important results of rate distortion theory, which will be used for discussing the efficiency of the presented lossy coding techniques.

References

[1] N. M. Abramson, *Information Theory and Coding*. New York, NY, USA: McGraw-Hill, 1963.

[2] N. Ahmed, T. Natarajan, and K. R. Rao, "Discrete cosine transform," *IEEE Transactions on Computers*, vol. 23, no. 1, pp. 90–93, 1974.

[3] S. Arimoto, "An algorithm for calculating the capacity of an arbitrary discrete memoryless channel," *IEEE Transactions on Information Theory*, vol. 18, pp. 14–20, January 1972.

[4] T. Berger, *Rate Distortion Theory*. NJ, USA: Prentice-Hall, Englewood Cliffs, 1971.

[5] J. Binia, M. Zakai, and J. Ziv, "On the ϵ-entropy and the rate-distortion function of certain non-gaussian processes," *IEEE Transactions on Information Theory*, vol. 20, pp. 514–524, July 1974.

[6] R. E. Blahut, "Computation of channel capacity and rate-distortion functions," *IEEE Transactions on Information Theory*, vol. 18, pp. 460–473, April 1972.

[7] M. Burrows and D. Wheeler, *A block-sorting lossless data compression algorithm*. CA, USA: Research Report 124, Digital Equipment Corporation, Palo Alto, May 1994.

[8] P.-C. Chang and R. M. Gray, "Gradient algorithms for designing predictive vector quantizers," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 34, no. 4, pp. 679–690, August 1986.

[9] P. A. Chou, T. Lookabaugh, and R. M. Gray, "Entropy-constrained vector quantization," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 37, no. 1, pp. 31–42, January 1989.

[10] R. J. Clarke, *Transform Coding of Images*. Orlando, FL: Academic Press, 1985.

[11] T. M. Cover and J. A. Thomas, *Elements of Information Theory.* Hoboken, NJ, USA: John Wiley and Sons, 2nd Edition, 2006.

[12] R. D. Dony and S. Haykin, "Optimally adaptive transform coding," *IEEE Transactions on Image Processing*, vol. 4, no. 10, pp. 1358–1370, October 1995.

[13] M. Effros, H. Feng, and K. Zeger, "Suboptimality of the Karhunen-Loève transform for transform coding," *IEEE Transactions on Information Theory*, vol. 50, no. 8, pp. 1605–1619, August 2004.

[14] R. G. Gallager, *Information Theory and Reliable Communication.* New York, USA: John Wiley & Sons, 1968.

[15] R. G. Gallager, "Variations on a theme by huffman," *IEEE Transactions on Information Theory*, vol. 24, no. 6, pp. 668–674, November 1978.

[16] A. Gersho and R. M. Gray, *Vector Quantization and Signal Compression.* Boston, Dordrecht, London: Kluwer Academic Publishers, 1992.

[17] H. Gish and J. N. Pierce, "Asymptotically efficient quantizing," *IEEE Transactions on Information Theory*, vol. 14, pp. 676–683, September 1968.

[18] G. H. Golub and H. A. van der Vorst, "Eigenvalue computation in the 20th century," *Journal of Computational and Applied Mathematics*, vol. 123, pp. 35–65, 2000.

[19] V. K. Goyal, "High-rate transform coding: How high is high, and does it matter?," in *Proceedings of the IEEE International Symposium on Information Theory*, Sorento, Italy, June 2000.

[20] V. K. Goyal, "Theoretical foundations of transform coding," *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 9–21, September 2001.

[21] V. K. Goyal, J. Zhuang, and M. Vetterli, "Transform coding with backward adaptive updates," *IEEE Transactions on Information Theory*, vol. 46, no. 4, pp. 1623–1633, July 2000.

[22] R. M. Gray, *Source Coding Theory.* Norwell, MA, USA: Kluwer Academic Publishers, 1990.

[23] R. M. Gray, "Toeplitz and circulant matrices: A review," *Foundations and Trends in Communication and Information Theory*, vol. 2, no. 3, pp. 155–329, 2005.

[24] R. M. Gray, *Linear Predictive Coding and the Internet Protocol.* Boston-Delft: Now Publishers Inc, 2010.

[25] R. M. Gray and L. D. Davisson, *Random Processes: A Mathematical Approach for Engineers.* Englewood Cliffs, NJ, USA: Prentice Hall, 1985.

[26] R. M. Gray and L. D. Davisson, *An Introduction to Statistical Signal Processing.* Cambridge University Press, 2004.

[27] R. M. Gray and A. H. Gray, "Asymptotically optimal quantizers," *IEEE Transactions on Information Theory*, vol. 23, pp. 143–144, January 1977.

[28] R. M. Gray and D. L. Neuhoff, "Quantization," *IEEE Transactions on Information Theory*, vol. 44, no. 6, pp. 2325–2383, October 1998.

[29] U. Grenander and G. Szegö, *Toeplitz Forms and Their Applications.* Berkeley and Los Angeles, USA: University of California Press, 1958.

[30] D. A. Huffman, "A method for the construction of minimum redundancy codes," in *Proceddings IRE*, pp. 1098–1101, September 1952.

[31] ISO/IEC, "Coding of audio-visual objects — part 2: Visual," ISO/IEC 14496-2, April 1999.

[32] ITU-T, "Video codec for audiovisual services at $p \times 64$ kbit/s," ITU-T Rec. H.261, March 1993.
[33] ITU-T and ISO/IEC, "Digital compression and coding of continuous-tone still images," ITU-T Rec. T.81 and ISO/IEC 10918-1 (JPEG), September 1992.
[34] ITU-T and ISO/IEC, "Generic coding of moving pictures and associated audio information — part 2: Video," ITU-T Rec. H.262 and ISO/IEC 13818-2, November 1994.
[35] ITU-T and ISO/IEC, "Lossless and near-lossless compression of continuous-tone still images," ITU-T Rec. T.87 and ISO/IEC 14495-1 (JPEG-LS), June 1998.
[36] ITU-T and ISO/IEC, "JPEG 2000 image coding system — core coding system," ITU-T Rec. T.800 and ISO/IEC 15444-1 (JPEG 2000), 2002.
[37] ITU-T and ISO/IEC, "JPEG XR image coding system — image coding specification," ITU-T Rec. T.832 and ISO/IEC 29199-2 (JPEG XR), 2009.
[38] ITU-T and ISO/IEC, "Advanced video coding for generic audiovisual services," ITU-T Rec. H.264 and ISO/IEC 14496-10 (MPEG-4 AVC), March 2010.
[39] C. G. J. Jacobi, "Über ein leichtes Verfahren, die in der Theorie der Säcularströmungen vorkommenden Gleichungen numerisch aufzulösen," *Journal für reine und angewandte Mathematik*, vol. 30, pp. 51–94, 1846.
[40] N. S. Jayant and P. Noll, *Digital Coding of Waveforms.* Englewood Cliffs, NJ, USA: Prentice-Hall, 1994.
[41] A. N. Kolmogorov, *Grundbegriffe der Wahrscheinlichkeitsrechnung.* Springer, Berlin, Germany, 1933. An English translation by N. Morrison appeared under the title Foundations of the Theory of Probability (Chelsea, New York) in 1950, with a second edition in 1956.
[42] Y. Linde, A. Buzo, and R. M. Gray, "An algorithm for vector quantizer design," *IEEE Transactions on Communications*, vol. 28, no. 1, pp. 84–95, January 1980.
[43] T. Linder and R. Zamir, "On the asymptotic tightness of the Shannon lower bound," *IEEE Transactions on Information Theory*, vol. 40, no. 6, pp. 2026–2031, November 1994.
[44] Y. N. Linkov, "Evaluation of Epsilon entropy of random variables for small esilon," *Problems of Information Transmission*, vol. 1, pp. 12–18, 1965.
[45] S. P. Lloyd, "Least squares quantization in PCM," *IEEE Transactions on Information Theory*, vol. 28, pp. 127–135, Unpublished Bell Laboratories Technical Note, 1957, March 1982.
[46] T. D. Lookabaugh and R. M. Gray, "High-resolution quantization theory and the vector quantizer advantage," *IEEE Transactions on Information Theory*, vol. 35, no. 5, pp. 1020–1033, September 1989.
[47] J. Makhoul, "Linear prediction: A tutorial review," *Proceedings of the IEEE*, vol. 63, no. 4, pp. 561–580, April 1975.
[48] J. Makhoul, S. Roucos, and H. Gish, "Vector quantization in speech coding," *Proceedings of the IEEE*, vol. 73, no. 11, pp. 1551–1587, November 1985.
[49] H. S. Malvar, *Signal Processing with Lapped Transforms.* Norwood, MA, USA: Artech House, 1992.
[50] D. Marpe, H. Schwarz, and T. Wiegand, "Context-adaptive binary arithmetic coding for H.264/AVC," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, no. 7, pp. 620–636, July 2003.

[51] D. Marpe, H. Schwarz, and T. Wiegand, "Probability interval partitioning entropy codes," in Submitted to *IEEE Transactions on Information Theory*, Available at http://iphome.hhi.de/marpe/download/pipe-subm-ieee10.pdf, 2010.

[52] J. Max, "Quantizing for minimum distortion," *IRE Transactions on Information Theory*, vol. 6, no. 1, pp. 7–12, March 1960.

[53] R. A. McDonald and P. M. Schultheiss, "Information rates of Gaussian signals under criteria constraining the error spectrum," *Proceedings of the IEEE*, vol. 52, pp. 415–416, 1964.

[54] A. Moffat, R. M. Neil, and I. H. Witten, "Arithmetic coding revisited," *ACM Transactions on Information Systems*, vol. 16, no. 3, pp. 256–294, July 1998.

[55] P. F. Panter and W. Dite, "Quantization distortion in pulse code modulation with nonuniform spacing of levels," *Proceedings of IRE*, vol. 39, pp. 44–48, January 1951.

[56] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes.* New York, NY, USA: McGraw-Hill, 2002.

[57] R. Pasco, "Source coding algorithms for fast data compression," Ph.D. dissertation, Stanford University, 1976.

[58] R. L. D. Queiroz and T. D. Tran, "Lapped transforms for image compression," in *The Transform and Data Compression Handbook.* CRC, pp. 197–265, Boca Raton, FL, 2001.

[59] J. Rissanen, "Generalized Kraft inequality and arithmetic coding," *IBM Journal of Research Development*, vol. 20, pp. 198–203, 1976.

[60] A. Said, "Arithmetic coding," in *Lossless Compression Handbook*, (K. Sayood, ed.), San Diego, CA: Academic Press, 2003.

[61] S. A. Savari and R. G. Gallager, "Generalized tunstall codes for soures with memory," *IEEE Transactions on Information Theory*, vol. 43, no. 2, pp. 658–668, March 1997.

[62] K. Sayood, ed., *Lossless Compression Handbook.* San Diego, CA: Academic Press, 2003.

[63] C. E. Shannon, "A mathematical theory of communication," *The Bell System Technical Journal*, vol. 27, no. 3, pp. 2163–2177, July 1948.

[64] C. E. Shannon, "Coding theorems for a discrete source with a fidelity criterion," *IRE National Convention Record, Part 4*, pp. 142–163, 1959.

[65] Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," *IEEE Transactions on Acoustics, Speech and Signal Processing*, vol. 36, pp. 1445–1453, September 1988.

[66] D. S. Taubman and M. M. Marcellin, *JPEG2000: Image Compression Fundamentals, Standards and Practice.* Kluwer Academic Publishers, 2001.

[67] B. P. Tunstall, "Synthesis of noiseless compression codes," Ph.D. dissertation, Georgia Inst. Technol., 1967.

[68] B. E. Usevitch, "A tutorial on mondern lossy wavelet image compression: Foundations of JPEG 2000," *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 22–35, September 2001.

[69] P. P. Vaidyanathan, *The Theory of Linear Prediction.* Morgan & Claypool Publishers, 2008.

[70] M. Vetterli, "Wavelets, approximation, and compression," *IEEE Signal Processing Magazine*, vol. 18, no. 5, pp. 59–73, September 2001.
[71] M. Vetterli and J. Kovacevic, *Wavelets and Subband Coding*. Englewood Cliffs, NJ: Prentice-Hall, 1995.
[72] I. H. Witten, R. M. Neal, and J. G. Cleary, "Arithmetic coding for data compression," *Communications of the ACM*, vol. 30, no. 6, pp. 520–540, June 1987.
[73] J. Ziv and A. Lempel, "A universal algorithm for data compression," *IEEE Transactions on Information Theory*, vol. 23, no. 3, pp. 337–343, May 1977.