Rate Distortion Bounds for Voice and Video

Jerry D. Gibson
University of California, Santa Barbara
gibson@ece.ucsb.edu

Jing Hu
University of California, Santa Barbara
jinghu@ece.ucsb.edu

now
the essence of knowledge
Boston — Delft

Foundations and Trends® in Communications and Information Theory

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
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

J. D. Gibson and J. Hu. *Rate Distortion Bounds for Voice and Video.* Foundations and Trends® in Communications and Information Theory, vol. 10, no. 4, pp. 379–514, 2013.

This Foundations and Trends® issue was typeset in LaTeX using a class file designed by Neal Parikh. Printed on acid-free paper.

ISBN: 978-1-60198-779-2

Foundations and Trends® in Communications and Information Theory
Volume 10, Issue 4, 2013

Editorial Scope

Topics

Foundations and Trends® in Communications and Information Theory publishes survey and tutorial articles in the following topics:

- Coded modulation
- Coding theory and practice
- Communication complexity
- Communication system design
- Cryptology and data security
- Data compression
- Data networks
- Demodulation and Equalization
- Denoising
- Detection and estimation
- Information theory and statistics
- Information theory and computer science
- Joint source/channel coding
- Modulation and signal design
- Multiuser detection
- Multiuser information theory
- Optical communication channels
- Pattern recognition and learning
- Quantization
- Quantum information processing
- Rate-distortion theory
- Shannon theory
- Signal processing for communications
- Source coding
- Storage and recording codes
- Speech and Image Compression
- Wireless Communications

Information for Librarians

Foundations and Trends® in Communications and Information Theory, 2013, Volume 10, 4 issues. ISSN paper version 1567-2190. ISSN online version 1567-2328. Also available as a combined paper and online subscription.

Foundations and Trends® in Communications and
Information Theory
Vol. 10, No. 4 (2013) 379–514

DOI: 10.1561/0100000061

Rate Distortion Bounds for Voice and Video

Jerry D. Gibson
University of California, Santa Barbara
gibson@ece.ucsb.edu

Jing Hu
University of California, Santa Barbara
jinghu@ece.ucsb.edu

Contents

Abstract

Numerous voice, still image, audio, and video compression standards have been developed over the last 25 years, and significant advances in the state of the art have been achieved. However, in the more than 50 years since Shannon's seminal 1959 paper, no rate distortion bounds for voice and video have been forthcoming. In this volume, we present the first rate distortion bounds for voice and video that actually lower bound the operational rate distortion performance of the best-performing voice and video codecs. The bounds indicate that improvements in rate distortion performance of approximately 50% over the best-performing voice and video codecs are possible. Research directions to improve the new bounds are discussed.

J. D. Gibson and J. Hu. *Rate Distortion Bounds for Voice and Video.* Foundations and Trends® in Communications and Information Theory, vol. 10, no. 4, pp. 379–514, 2013.

DOI: 10.1561/0100000061.

1

Introduction

Numerous voice, still image, audio, and video compression standards have been developed over the last 25 years, and significant advances in the state of the art have been achieved. There are several reasons for researchers and standards bodies to consider developing new voice or video codecs. One motivation might be a new application that has different constraints than those imposed on prior codecs. For example, a new application might require better quality, lower complexity, a different transmitted bit rate, or improved robustness to channel impairments. A second motivation might be that the input source changes, namely a different resolution for video, a requirement for 3D video, or a different bandwidth and sampling rate for audio. A third motivation might be that a particular codec is relatively old and that there is the possibility of improving performance, perhaps by increasing complexity because of advances due to Moore's Law.

In each of these cases, it would seem natural to ask what is the best possible performance theoretically achievable by a new codec? Or, alternatively, given the operational rate distortion performance of a particular codec, how close is the operational rate distortion performance to the optimal performance theoretically achievable?

To answer this question, one natural place to look in order to characterize the best possible performance of any lossy source codec would appear to be rate distortion theory. In particular, it would be of great utility if the host of existing rate distortion theory results could be applied to bounding the performance of practical codecs or if new rate distortion bounds for such practical sources and their attendant perceptual distortion measures could be obtained. However, no such applications of existing rate distortion theory results, nor any appropriate new results, have been forthcoming. While there are many reasons for this lack of progress, one main reason is that such an effort is not easy – in fact, it is particularly difficult.

The particular challenges involved were anticipated by experts in Information Theory very early. Specifically, Robert Gallager, in his classic text on Information Theory [18], summarizes the challenges at the end of his rate distortion theory chapter where he notes that information theory has been more useful for channel coding than for source coding and that the reason, "... appears to lie in the difficulty of obtaining reasonable probabilistic models and meaningful distortion measures for sources of practical interest." He goes on to say, "... it is not clear at all whether the theoretical approach here will ever be a useful tool in problems such as speech digitization ... " [18].

Finding suitable statistical models for video has been considered a very difficult topic as well. In 1998, almost 40 years after Shannon's landmark paper developing rate distortion theory [76], Ortega and Ramchandran wrote, "'Unfortunately, to derive bounds one needs to first characterize the sources and this can be problematic for complex sources such as video. Indeed, bounds are likely to be found only for the simpler statistical models"' [67].

Thus, like all rate distortion problems, the two primary challenges are (1) finding good source models for speech and video, and (2) identifying a distortion measure that is perceptually meaningful, yet computationally tractable. There have been only a few prior research efforts in the last 25 years that have attempted to address various aspects of this problem for either speech or video, and broad-based bounds of significance have not been obtained. It is clear, however, that the utility of such bounds would be substantial.

In this volume, we present our recent results on obtaining rate distortion functions for both voice and video sources. For both sources, we overcome past limitations on source modeling by employing composite source models to achieve more accurate modeling of the different voice and video source modes. Although we use composite source models for both voice and video, the treatments of the distortion measure for the two sources are distinctly different. For speech, we devise a mapping technique to extend existing MSE $R(D)$ results to the perceptually meaningful PESQ-MOS distortion measure. For video, no such mappings are developed and the MSE distortion measure, or equivalently peak SNR (PSNR), is used directly to develop our video $R(D)$ bounds. This is because although MSE and PSNR are widely criticized as not having a direct interpretation in terms of reconstructed video quality, PSNR is known to order the performance of codecs in the same class correctly. In fact, since optimizing MSE/PSNR often produces competitive performance in terms of perceptual measures, and its limitations are well known, it is still a dominant performance measure in video codec standardization efforts.

For future progress, as well as for the development of future practical rate distortion results, it is critical to note from the above outline of the approaches used here that there are two key elements in play in order to obtain the rate distortion bounds presented in this volume. These are (1) a grasp and fundamental understanding of key rate distortion theory results, and (2) a deep understanding of the real-world sources and their codec performance evaluation methods. Either one alone is not sufficient. Indeed, the first author has emphasized to his students repeatedly over the past 30 years that in order to utilize significant theoretical results for practical problems, one must also have an understanding of the physical problem being addressed. This combination is not often present, perhaps because, as noted by Berger and Gibson [7], rate distortion theorists and voice and video codec designers are mostly non-intersecting sets of researchers.

We summarize the contents of this volume for each source in the following subsections.

1.1 Rate Distortion Functions for Speech Sources

We develop new rate distortion bounds for narrowband and wideband speech coding based on composite source models for speech and perceptual PESQ-MOS/WPESQ distortion measures. It is shown that these new rate distortion bounds do in fact lower bound the performance of important standardized speech codecs, including, G.726, G.727, AMR-NB, G.729, G.718, G.722, G.722.1, and AMR-WB.

Our approach is to calculate rate distortion bounds for mean squared error (MSE) distortion measures using the classic eigenvalue decomposition and reverse water-filling method for each of the subsource modes of the composite source model, and then use conditional rate distortion theory to calculate the overall rate distortion function for the composite source. While composite source models for speech have been considered previously for obtaining $R(D)$ functions for speech, our method of choosing the subsources based on a knowledge of speech signals and on successful multi-mode voice codecs, as well as the inclusion of diverse subsources in the composite source models, are new.

In order to develop $R(D)$ bounds for speech in terms of a meaningful distortion measure that still allows a tractable mathematical calculation of the bounds required a new innovation as well. Mapping functions are developed to map rate distortion curves based on MSE to rate distortion curves subject to the perceptually meaningful distortion measures PESQ-MOS and WPESQ. These final rate distortion curves are then compared to the performance of the best known standardized speech codecs based on the code-excited linear prediction paradigm.

In addition to the striking result that these new bounds do in fact lower bound the best known narrowband and wideband standardized speech codecs, the bounds are revealing in that performance comparisons show that current linear predictive codecs do a relatively good job of coding voiced speech, but are much less effective for other subsources, such as unvoiced speech, Onset, and Hangover modes. Equally important is that the procedure used in developing our bounds can easily be reproduced by other researchers, and thus other, perhaps more refined, rate distortion curves can be generated. For example, one could utilize a different composite source model with the known MSE rate distortion theory results outlined here, and then employ our mapping functions to determine new bounds for the utterances considered in this paper.

1.2 Rate Distortion Functions for Video Sources

For the video source we address the difficult task of modeling the correlation in pixel values by first proposing a new spatial correlation model for two close pixels in one frame of digitized natural video sequences that is conditional on the local texture. This new spatial correlation model is dependent upon five parameters whose optimal values are calculated for a specific image or specific video frames. The new spatial correlation model is simple, but it performs very well, as strong agreement is discovered between the approximate correlation coefficients and the correlation coefficients calculated by the new correlation model, with a mean absolute error (MAE) usually smaller than 5%.

Further, we extend the correlation coefficient modeling from pixels within one video frame to pixels that are located in nearby video frames. We show that for two pixels located in nearby video frames, their spatial correlation and their temporal correlation are approximately independent. Therefore the correlation coefficient of two pixels in two nearby video frames, denoted by ρ, can be modeled as the product of ρ_s, the texture dependent spatial correlation coefficient of these two pixels, as if they were in the same frame, and ρ_t, a variable to quantify the temporal correlation between these two video frames. ρ_t does not depend on the textures of the blocks the two pixels are located in and is a function of the indices of the two frames.

With the new block-based local-texture-dependent correlation model, we first study the marginal rate distortion functions of the different local textures. These marginal rate distortion functions are shown to be quite distinct from each other. Classical results in information theory are utilized to derive the conditional rate distortion function when the universal side information of local textures is available at both the encoder and the decoder. We demonstrate that by involving this side information, the lowest rate that is theoretically achievable in *intra-frame* video compression can be as much as 1 bit per pixel lower than that without the side information; and the lowest rate that is theoretically achievable in *inter-frame* video compression can be as much as 0.7 bit per pixel lower than that without the side information. The rate distortion bounds with local texture information taken into account while making no assumptions on coding, are shown indeed to be valid lower bounds with respect to the operational rate distortion curves of both *intra-frame* and *inter-frame* coding in Advanced Video Coding (AVC/H.264) and in the newly standardized High Efficiency Video Coding (HEVC/H.265).

The incorporation of the new correlation model into existing operational models of practical image and video compression systems is also promising. We demonstrate this by studying the common "blocking" scheme used in most video compression standards [32, 33, 34, 35], which divides a video frame into 16×16 macroblocks (MB) or smaller blocks before processing. With the block based nature of the new correlation model, we study the penalty paid in average rate when the correlation among the neighboring MBs or blocks is disregarded completely or is incorporated partially through predictive coding. A constrained rate distortion bound is calculated for the scenario when the texture information is coded losslessly and optimal predictive coding is employed. This lower bound is shown to be reasonably tight with respect to the operational rate distortion curves of intra-frame coding in AVC/H.264. Furthermore, it is near linear in terms of average bit rate per pixel versus PSNR of a video frame and can easily be utilized in future video codec designs.

1.3 Conclusion

In this volume, we present the first rate distortion bounds for voice and video that actually lower bound the operational rate distortion performance of the best-performing voice and video codecs. Members of the Panel on "New Perspectives on Information Theory" held at the IEEE Information Theory Workshop at Paraty, Brazil, on October

20, 2011, repeatedly expressed their concern about the gap between lossy compression theory and practice [82]. The new rate distortion bounds presented here, for the first time, make the gap specific for voice and video, and as discussed later, aid in pointing the way forward to improving the performance of practical voice and video codecs.

References

[1] 3GPP. Mandatory Speech Codec speech processing functions; Adaptive Multi-Rate (AMR) speech codec; Transcoding functions. TS 26.090, 3rd Generation Partnership Project (3GPP), Mar. 2011.

[2] 3GPP. Speech codec speech processing functions; Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; Transcoding functions. TS 26.190, 3rd Generation Partnership Project (3GPP), Mar. 2011.

[3] B. S. Atal and S. L. Hanauer. Speech analysis and synthesis by linear prediction of the speech wave. *Journal of the Acoustic Society of America*, 50:637–655, 1971.

[4] B. S. Atal and M. R. Schroeder. Adaptive predictive coding of speech signals. *The Bell System technical journal*, pages 1973–1986, 1970.

[5] B. Bessette, et al. The adaptive multirate wideband speech codec (AMR-WB). *IEEE Trans. on Speech and Audio Processing*, 10:620–636, Nov. 2002.

[6] T. Berger. *Rate Distortion Theory*. Prentice-Hall, 1971.

[7] T. Berger and J. D. Gibson. Lossy Source Coding. *IEEE Trans. on Information Theory*, 44(6):2693 –2723, Oct. 1998.

[8] M.J. Carter. *Soource coding of composite sources*. PhD thesis, The University of Michigan, 1984.

[9] Tihao Chiang and Ya-Qin Zhang. A new rate control scheme using quadratic rate distortion model. *IEEE Transactions on Circuits and Systems for Video Technology*, 7(1):246–251, Feb. 1997.

[10] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley-Interscience, New York, Aug. 1991.

[11] T. M. Cover and J. A. Thomas. *Elements of information theory*. Wiley-Interscience, 1991.

[12] Richard V Cox, Simao Ferraz De Campos Neto, Claude Lamblin, and Mostafa Hashem Sherif. Itu-t coders for wideband, superwideband, and fullband speech communication [series editorial]. *Communications Magazine, IEEE*, 47(10):106–109, 2009.

[13] D.-K. Kwon, M.-Y. Shen and C.-C. J. Kuo. Rate control for H.264 video with enhanced rate and distortion models. *IEEE Transactions on Circuits and Systems for Video Technology*, 17(5):517–529, May 2007.

[14] W. Daumer. Subjective Evaluation of Several Efficient Speech Coders. *IEEE Trans. on Communications*, 30(4):655 – 662, Apr. 1982.

[15] L. D. Davisson. Rate-distortion theory and application. *Proceedings of the IEEE*, 60(7):800 – 808, July 1972.

[16] A. De and P. Kabal. Rate-distortion function for speech coding based on perceptual distortion measure. *IEEE Global Telecommunications Conference*, pages 452–456, Orlando, Dec. 1992.

[17] R.J. Fontana. *A class of composite sources and their erfodic and information theoretic properties*. PhD thesis, Standford University, 1978.

[18] R. G. Gallager. *Information Theory and Reliable Communication*. John Wiley & Sons, Inc., New York, NY, 1968.

[19] S.A. Garde. *Communication of composite sources*. PhD thesis, Univeristy of California, Berkeley, 1980.

[20] J. D. Gibson. Speech Coding Methods, Standards, and Applications. *IEEE Circuits and Systems Magazine*, 5(4):30 – 49, Fouth Quarter 2005.

[21] J. D. Gibson, J. Hu, and P. Ramadas. New Rate Distortion Bounds for Speech Coding Based on Composite Source Models. *Information Theory and Applications Workshop (ITA)*, UCSD, Jan. 31 - Feb. 5, 2010.

[22] J. D. Gibson and Y.-Y. Li. Rate Distortion Performance Bounds for Wideband Speech. *Information Theory and Applications Workshop*, San Diego, CA, Feb. 5-10, 2012.

[23] Jerry D. Gibson, Toby Berger, Tom Lookabaugh, Dave Lindbergh, and Richard L. Baker. *Digital compression for multimedia: principles and standards*. Morgan Kaufmann Publishers Inc., San Francisco, CA, 1998.

[24] Bernd Girod. The efficiency of motion-compensating prediction for hybrid coding of video sequences. *IEEE Journal on selected areas in communications*, SAC-5(7):1140–1154, Aug. 1987.

[25] Bernd Girod. Motion-compensating prediction with fractional-pel accuracy. *IEEE Transactions on Communications*, 41:604–612, Apr. 1993.

[26] R. M. Gray. Information rates of autoregressive processes. *IEEE Trans. on Information Theory*, 16(4):412 – 421, Jul. 1970.

[27] R. M. Gray. A new class of lower bounds to information rates of stationary sources via conditional rate-distortion functions. *IEEE Trans. on Information Theory*, IT-19(4):480–489, July 1973.

[28] R. M. Gray. A new class of lower bounds to information rates of stationary sources via conditional rate-distortion functions. *IEEE Tran. Inform. Theory*, IT-19(4):480–489, Jul. 1973.

[29] H. Brehm and K. Trottler. Rate distortion functions for speech-model signals. *Signal Processing III: Theories and Applications*, pages 353–356, EURASIP, 1986.

[30] Ali Habibi and Paul A. Wintz. Image coding by linear transformation and block quantization. *IEEE Transactions on Communication Technology*, Com-19(1):50–62, Feb. 1971.

[31] J. Hu and J. D. Gibson. New rate distortion bounds for natural videos based on a texture dependent correlation model in the spatial-temporal domain. *the 46th Annual Allerton Conference on Communication, Controls, and Computing*, Sept. 2008.

[32] ISO/IEC 13818-1:2000. Information technology – generic coding of moving pictures and associated audio information: Systems. 2000.

[33] ISO/IEC 14496-1:2001. Information technology – coding of audio-visual objects – part 1: Systems. 2001.

[34] ITU Recommendations. Video coding for low bit rate communication. *ITU-T rec. H.263*, Jan. 2005.

[35] ITU-T and ISO/IEC JTC 1. Advanced video coding for generic audio-visual services. 2003.

[36] ITU-T and ISO/IEC JTC 1. H.265 : High efficiency video coding. *http://www.itu.int/rec/T-REC-H.265-201304-I*, Apr. 2013.

[37] ITU-T Recommendation G.191. Software tools for speech and audio coding standardization. Mar. 2010.

[38] ITU-T Recommendation G.718. Frame error robust narrow-band and wideband embedded variable bit-rate coding of speech and audio from 8-32 kbit/s. June 2008.

[39] ITU-T Recommendation G.719, Series G. Transmission systems and media, digital systems and networks, Digital terminal equipments-Coding of analogue signals by pulse code modulation. June 2008.

[40] ITU-T Recommendation G.722. 7 kHz Audio-Coding within 64 kbits/s . Nov. 1988.

[41] ITU-T Recommendation G.722.1. Low-complexity coding at 24 and 32 kbit/s for hands-free operation in systems with low frame loss . May 2005.

[42] ITU-T Recommendation G.722.2. Wideband coding of speech at around 16 kbit/s using Adaptive Multi-Rate Wideband (AMR-WB). July 2003.

[43] ITU-T Recommendation G.726. 40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM) . Dec. 1990.

[44] ITU-T Recommendation G.727. 5-, 4-, 3- and 2-bit/sample embedded Adaptive Differential Pulse Code Modulation (ADPCM). Dec. 1990.

[45] ITU-T Recommendation G.729. Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction (CS-ACELP) . Jan. 2007.

[46] ITU-T Recommendation P.56. Objective measurement of active speech level. Mar. 1993.

[47] ITU-T Recommendation P.830. Subjective performance assessment of telephone-band and wideband digital codecs. Feb. 1996.

[48] ITU-T Recommendation P.862. Perceptual Evaluation of Speech Quality (PESQ), an objective method for end-to-end Speech Quality Assessment of Narrow-band telephone networks and Speech Codecs. Feb. 2001.

[49] ITU-T Recommendation P.862. Perceptual Evaluation of Speech Quality (PESQ), an objective method for end-to-end Speech Quality Assessment of Narrow-band telephone networks and Speech Codecs. Feb. 2001.

[50] ITU-T Recommendation P.862.2. Wideband extension to Recommendation P.862 for the assessment of wideband telephone networks and speech codecs. Nov. 2007.

[51] ITU-T Recommendation P.862.3. Application guide for objective quality measurement based on Recommendations P.862, P.862.1 and P.862.2. Nov. 2007.

[52] J. B. O'neal Jr. and T. Raj Natarajan. Coding isotropic images. *IEEE Transactions on Information Theory*, IT-23(6):697–707, Nov. 1977.

[53] N. S. Jayant and P. Noll. *Digital Coding of Waveforms: Principles and Applications to Speech and Video.* Prentice Hall, Mar. 1984.

[54] K. Järvinen, I. Bouazizi, L. Laaksonen, P. Ojala, and A. Rämö. Media coding for the next generation mobile system lte. *Computer Communications*, 33(16):1916 – 1927, 2010.

[55] H. Kalveram and P. Meissner. Itakura-saito clustering and rate distortion functions for a composite source model of speech. *Signal Processing*, 18(2):195 – 216, 1989.

[56] H. Kalveram and P. Meissner. Rate Distortion Bounds for Speech Waveforms based on Itakura-Saito-Segmentation. *Signal Processing IV: Theories and Applications*, EURASIP, 1988.

[57] W Bastiaan Kleijn and Kuldip K Paliwal. *Speech coding and synthesis.* Elsevier Science Inc., 1995.

[58] A Kolmogorov. On the shannon theory of information transmission in the case of continuous signals. *IEEE Transactions on Information Theory*, 2(4):102–108, 1956.

[59] A. M. Kondoz. *Digital Speech: Coding for Low Bit Rate Communication Systems.* John Wiley & Sons, West Sussex, England, 2004.

[60] Ahmet M Kondoz. *Digital speech: coding for low bit rate communication systems.* Wiley. com, 2005.

[61] Hung-Ju Lee, Tihao Chiang, and Ya-Qin Zhang. Scalable rate control for MPEG-4 video. *IEEE Transactions on Circuits and Systems for Video Technology*, 10(6):878–894, Sep. 2000.

[62] Q. Li and M. van der Schaar. Providing adaptive QoS to layered video over wireless local area networks through real-time retry limit adaptation. *IEEE Transactions on Multimedia*, 6(2):278–290, Apr. 2004.

[63] A. K. Luthra, G. J. Sullivan, and T. Wiegand. Introduction to the special issue on the H.264/AVC video coding standard. *IEEE Transactions on Circuits and Systems for Video Technology*, 13(7), Jul. 2003.

[64] et al M. Jelinek. On the architecture of the cdma2000ő variable-rate multimode wideband (VMR-WB) speech coding standard. *Proc. ICASSP*, pages I–281–I–284, 2004.

[65] Siwei Ma, Wen Gao, and Yan Lu. Rate control on JVT standard. *Joint Video Team (JVT) of ISO/IEC MPEG & ITU-T VCEG, JVT-D030*, Jul. 2002.

[66] Jens-Rainer Ohm and Gary J. Sullivan. High efficiency video coding: The next frontier in video compression. *IEEE Signal Processing Magazine*, 30(1):152–158, Jan. 2013.

[67] A. Ortega and K. Ramchandran. Rate-distortion methods for image and video compression. *IEEE Signal Processing Magazine*, 15(6):23Ű50, Nov. 1998.

[68] P. Cummiskey, N. S. Jayant, and J. L. Flanagan. Adaptive Quantization in Differential PCM Coding of Speech. *The Bell System technical journal*, 52:1105–1118, Sept. 1973.

[69] T. N. Pappas and R. J. Safranek. Perceptual criteria for image quality evaluation. *Handbook of Image & Video Processing (A. Bivok eds.), Academic Press*, 2000.

[70] Mark S Pinsker. Mutual information between a pair of stationary gaussian random processes. In *Dokl. Akad. Nauk. USSR*, volume 99, pages 213–216, 1954.

[71] Quackenbush. MPEG unified speech and audio coding. *IEEE Multimedia*, 20(2):72–78, April-June 2013.

[72] R. C. Reininger and J. D. Gibson. Distributions of the two-dimensional DCT coefficients for images. *IEEE Transactions on Communications*, 31:835–839, Jun. 1983.

[73] P. Ramadas and J. D. Gibson. Phonetically Switched Tree coding of speech with a G.727 Code Generator. *the 43rd Annual Asilomar Conference on Signals, Systems, and Computers*, Nov. 1-4, 2009.

[74] Jordi Ribas-Corbera and Shawmin Lei. Rate control in DCT video coding for low-delay communications. *IEEE Transactions on Circuits and Systems for Video Technology*, 9(1):172–185, Feb. 1999.

[75] C. E. Shannon. A mathematical theory of communication. *Bell Sys. Tech. Journal*, 27:379–423, 1948.

[76] C. E. Shannon. Coding Theorems for a Discrete Source with a Fidelity Criterion. *IRE Conv. Rec.*, 7:142–163, 1959.

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

[78] S. R. Smoot and L. A. Rowe. Study of DCT coefficient distributions. *SPIE Symposium on Electronic Imaging, San Jose, CA*, 2657, Jan. 1996.

[79] Klaus Stuhlmuller, Niko Farber, Michael Link, and Bernd Girod. Analysis of video transmission over lossy channels. *IEEE Journal on Selected Areas in Communications*, 18(6), Jun. 2000.

[80] T. Aach, C. Mota, I. Stuke, M. Mühlich, and E. Barth. Analysis of superimposed oriented patterns. *IEEE Transactions on Image Processing*, 15(12):3690–3700, Dec. 2006.

[81] G. Tziritas. Rate distortion theory for image and video coding. *International Conference on Digital Signal Processing, Cyprus*, 1995.

[82] Sergio Verdu, Venkat Anantharam, Giuseppe Caire, Max Costa, Gerhard Kramer, and Raymond Yeung. Panel on new perspectives for information theory. *Information Theory Workshop, Paraty, Brazil*, Oct. 2011.

[83] Koen Vos, Karsten Vandborg Sørensen, Søren Skak Jensen, and Jean-Marc Valin. The opus codec.

[84] W. B. Kleijn and K. K. Paliwal, eds. *Speech Coding and Synthesis*. Elsevier, Amsterdam, Holland, 1995.

[85] M.S. Wallace. Some techniques in universal source coding and coding for composite sources. Master's thesis, University of Illinois at Urbana-Champaign, 1982.

[86] S. Wang and A. Gersho. Improved Phonetically- Segmented Vector Excitation Coding at 3.4 Kb/s. In *Proceedings, IEEE ICASSP*, San Francisco, Mar. 1992.

[87] S. Wang and A. Gersho. Phonetically-based vector excitation coding of speech at 3.6 kbit/s. *Proceedings, IEEE ICASSP*, pages 49–52, Glasgow, May 1989.

[88] Z. Wang, H. R. Sheikh, and A. C. Bovik. Objective video quality assessment. *The Handbook of Video Databases: Design and Applications (B. Furht and O. Marqure, eds.), CRC Press*, pages 1041–1078, Sep. 2003.

[89] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and Ajay Luthra. Overview of the H.264/AVC video coding standard. *IEEE Transactions on Circuits and Systems for Video Technology*, 13:560–576, Jul. 2003.

[90] Y. K. Tu, J.-F. Yang and M.-T. Sun. Rate-distortion modeling for efficient H.264/AVC encoding. *IEEE Transactions on Circuits and Systems for Video Technology*, 17(5):530–543, May 2007.