Bounding the Rate Distortion Performance of Voice and Video Codecs

Jerry D. Gibson
Department of Electrical and Computer Engineering
University of California, Santa Barbara
Santa Barbara, CA 93106

Abstract

Bounding the rate distortion performance of voice and video codecs has been an elusive goal for over 50 years. The challenge has been to find accurate, but tractable, source models, and meaningful, but tractable, fidelity criteria, and then deriving provably optimal bounds. We address these two voice and video rate distortion problems within the general approach of developing new composite source models each for voice and video, combining these models with meaningful distortion measures, and utilizing classical results on conditional rate distortion theory to obtain the bounds. For video, we develop a new five parameter spatial correlation model coupled with a temporal correlation component conditioned on different spatial textures. For voice, we utilize a five mode composite source model that includes an autoregressive model for voiced speech and a memoryless source model for unvoiced speech. The conditional rate distortion theory results employ the mean squared error (MSE) distortion measure, and we map the MSE into a perceptual distortion measure widely used for voice codec evaluation to obtain meaningful results. We show that for video, our results lower bound the performance of the H.264 and recent HEVC video coding standards, and that for voice, our results lower bound the performance of the best known voice codecs, including the AMR-NB and AMR-WB standards. Comparisons to the bounds indicate where existing voice and video codecs might be improved. Directions for improving the rate distortion performance bounds are presented.

Biographical Sketch: Jerry D. Gibson is Professor and Chair of Electrical and Computer Engineering at the University of California, Santa Barbara. He has been an Associate Editor of the *IEEE Transactions on Communications* and the *IEEE Transactions on Information Theory*. He was President of the IEEE Information Theory Society in 1996, and he has served on the Board of Governors of the IT Society and the Communications Society. He was an IEEE Communications Society Distinguished Lecturer for 2007-2008, a member of the IEEE Awards Committee (2008-2010), and a member of the IEEE Medal of Honor Committee (2009-2010). He is an IEEE Fellow, and he has received The Fredrick Emmons Terman Award (1990), the 1993 IEEE Signal Processing Society Senior Paper Award, the 2009 IEEE Technical Committee on Wireless Communications Recognition Award, and the 2010 Best Paper Award from the *IEEE Transactions on Multimedia*.