Antonio Ortega, *University of Southern California*
and Kannan Ramchandran, *University of Illinois at Urbana-Champaign*

From Rate-Distortion Theory to Commercial Image and Video Compression Technology

The current explosion in digital multimedia- and Internet-driven technology, particularly that involving digital visual media such as images and video, has been catalyzed in no small measure by recent advances in commercial compression technology. To the purist, any connection between this and the theoretical efficacies of Shannon's groundbreaking work on rate-distortion theory may appear to be tenuous at best. A less arrogant and broader outlook, however, reveals the unmistakable stamp that the teachings of Shannon (even as we celebrate the 50th anniversary of his pioneering work [1]) have had on the philosophy, design, and evolution of modern-day image and video compression technology. An example of this is the evolution of current entropy coders such as Huffman coders that form the core of baseline image and video compression standards like JPEG and MPEG, or even better, adaptive arithmetic coders that form an integral part of current state-of-the-art image coding algorithms.

This is not to imply that commercial image and video compression standards such as JPEG and MPEG, which "rule the world" in the wake of today's digital multimedia revolution, follow directly from the "random coding arguments" that form the backbone of so many asymptotic proofs in Shannon's rate-distortion theory. Nor is the claim that modern-day algorithms for real-time multimedia applications, which are victims of stringent delay and complexity concerns, have much in common with error exponents in information theory. What is true, however, is that one can reflect the philosophy of Shannon's rate-distortion theory to the practical framework of "operational" rate-distortion theory imposed by practical considerations. The fundamental principle of the "fidelity versus bit rate" trade-off applies not only to Shannon theory but also to operational frameworks, even those imposed by very practical considerations, such as commercial standards! For example, the use of Lagrangian-based optimization finds its way into "operational" rate-distortion theory as well as its theoretical counterpart.

However, we want to stress that the tools of the trade are not the same for the two frameworks. As we will see in this special issue, due to the extra constraints (based on structure, delay, complexity, etc.) imposed by the "operational" version of rate-distortion theory, one has to resort to a number of optimization tools that are not commonly used in the "information-theoretic" version. One fundamental feature of a typical operational framework, especially as it applies to commercial standards like JPEG and MPEG, is that it is a discrete optimization problem. The framework is typically equipped with a multitude of system parameters to be tweaked, and the goal of operational rate-distortion theory is to tweak these parameters (i) optimally (i.e., in the fidelity-versus-bit rate trade-off sense), and (ii) efficiently (i.e., using fast optimization algorithms rather than with brute-force search).

➤ **An unwritten goal of this special issue is to serve as the middle ground between "irrelevant theory" and "ad hoc tweaking."**

This special issue is dedicated to surveying and describing, in tutorial form, a good cross-section of the large body of work that has been done recently related to rate-distortion-based methods for image and video compression. An unwritten goal of this special issue is to serve as the middle ground between "irrelevant theory" and "ad hoc tweaking," which are common stereotypes in the extended community involving both information theorists and practitioners attending standards meetings. The "irrelevant" stereotype on information theorists is associated with the rather simplistic input modeling assumptions typically used (e.g., i.i.d. Gaussian), and due to the asymptotic and existential-only nature of most of the results. On the other hand, the "ad hoc" stereotype is typically levied on industrial practitioners due to their perceived knob-tweaking-happy mentality that is associated with a lack of scientific or academic rigor.

While these stereotypes do have some truth to them, they attest more to the lack of communication between those two communities than to potential shortcomings in the respective areas. As an example, recent results in the information-theory literature based on sophisticated universal coding techniques, such as the "context tree weighting" concept of Willems et al. that received a Best Paper award, hold a lot of promise but have yet to be fully incorporated into state-of-the-art coding algorithms. Likewise, as shown by the article by Sullivan and Wiegand in this issue and several references mentioned therein, there is a great deal of awareness in the industry that significant gains can be achieved through both creative and systematic coder optimization, where, given the dimension of the parameter space, a solid formalization is of the outmost importance. If more of this work does not come to light, it may be due in part to the intellectual-property issues accompanying these types of developments.

A more positive twist can be put on this matter by noting that both groups have contributed heavily to the science and engineering of signal processing and communications—one only has to look at the evolution of modem technology to assess the tremendous influence of information theory, while the multimedia and Internet-based products that hit the market or spawn startup enterprises on a daily basis testify to the influence that standards have had on technology. One of the goals of this special issue is to act as a bridge between these communities by revealing that "knob-tweaking" can be more of a science than an art form: myriad optimization tools will be revealed in tutorial-style description. Another goal of this special issue is to inspire the need for theoretical analysis based on more complex statistical source models that more accurately reflect "real" images and video signals, as well as to inspire the establishment of fundamental performance bounds subject to delay and/or complexity constraints.

In order to provide a historical perspective of the evolution of rate-distortion-based methods in image and video compression, let us point out some of the key developments that underpin the work described in this special issue. To the above-mentioned work by Shannon [1], we can add the development of dynamic programming by Bellman [2] and of the discrete version of Lagrangian optimization by Everett [3] in the late 1950s and early 1960s, respectively. Optimization techniques of this type were used in the early work on vector quantization by Gray, Gersho, and others [4]. The earliest examples of applications of optimization to standards-based coders can be found in the work of Shoham and Gersho [5] and Wu and Gersho [6].

There are four articles in this special issue. The first article by Ortega and Ramchandran, "Rate-Distortion Methods for Image and Video Compression," provides both a tutorial and a survey of some recent work on rate-distortion-based methods in image and video compression. This article overviews the relationship between techniques used in practice and relevant information-theoretic results. It then takes a look at some of the typical problems encountered in image and video compression and transmission and describes the toolbox needed to solve these problems efficiently: the most popular tools are based on Lagrangian optimization and dynamic programming. It then surveys some of the recent work in the field.

The article by Effros, "Optimal Modeling for Complex System Design," focuses on the basics of coder design. In particular, it describes (in tutorial form) the powerful tools of the trade, such as the ubiquitous Lloyd algorithm in its many powerful variant forms, which allows the optimization of the encoder and decoder designs. This article also shows how the lessons of information theory can be applied to construct simple yet powerful models that are useful in many applications, not just compression. It relates how rigorous optimization techniques can be used on rather complex models, not just simple ones as typically done, so as to more accurately capture complex sources such as images. This is done through the powerful use of two-stage universal coding techniques.

The article by Sullivan and Wiegand, "Rate-Distortion Optimization for Video Compression," provides concrete evidence of the power of efficient model selection in practical video coding by showing results in the popularly deployed H.263 coder. This article provides an overview of the plethora of encoding modes available in state-of-the-art coders and provides a comprehensive guide to the approaches used by practitioners to optimize the mode selection. The underlying philosophy is based on using specific modes for parts of the video sequence when this is "worthwhile" in a rate-distortion sense. Further, the article illustrates how this philosophy can result in significant improvements over simpler approaches where there is less flexibility in the mode selection.

The article by Schuster, Melnikov, and Katsaggelos, "Operationally Optimal Vertex-based Shape Coding," provides a more focused look at one particular application of rate-distortion techniques, namely shape coding. The article presents a complete description of vertex-based shape coding and demonstrates how rate-distortion optimization tools can be applied to achieve the best trade-off between the rate required to code a contour and how well this contour approximates the original one. This issue is likely to be an important one as object-based, rather than frame-based, coding becomes commonplace.

—**Antonio Ortega**
Integrated Media Systems Center
Dept. of Electrical Eng.-Systems,
University of Southern California

—**Kannan Ramchandran**
Beckman Institute
Electrical and Computer
Engineering Dept.
University of Illinois at
Urbana-Champaign

References

1. C.E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. Journal*, vol. 27, pp. 379-423, 1948.

2. R. Bellman, *Dynamic Programming*, Princeton University Press, 1957.

3. H. Everett, "Generalized Lagrange multiplier method for solving problems of optimum allocation of resources," *Operations Research*, vol. 11, pp. 399-417, 1963.

4. P.A. Chou, T. Lookabaugh, and R.M. Gray, "Optimal pruning with applications to tree-structured source coding and modeling," *IEEE Trans. on Inform. Theory*, vol. 35, pp. 299-315, Mar. 1989.

5. Y. Shoham and A. Gersho, "Efficient bit allocation for an arbitrary set of quantizers," *IEEE Trans. ASSP*, vol. 36, pp. 1445-1453, Sep. 1988.

6. S.-W. Wu and A. Gersho, "Rate-constrained optimal block-adaptive coding for digital tape recording of HDTV" *IEEE Trans. on Circuits and Sys. for Video Tech.*, vol. 1, pp. 100-112, Mar. 1991.