Predictive Vector Quantizer Design by Deterministic Annealing

Hosam Khalil and Kenneth Rose
Department of Electrical and
Computer Engineering
University of California
Santa Barbara, CA, 93106
e-mail: {hosam,rose}@ece.ucsb.edu

Predictive vector quantizers (PVQ) are useful for the compression of sources with memory and other correlated data by incorporating memory into the quantization process [1][2]. A new approach is proposed for the design of PVQ, which is inherently probabilistic and is based on ideas from information theory and analogies to statistical physics. By formulating both the prediction and quantization in probabilistic terms, a joint optimization procedure is made possible, which directly minimizes the expected distortion. The approach resolves two longstanding fundamental shortcomings of standard PVQ design.

The first shortcoming is due to the piecewise constant nature of the quantizer function; it is difficult to optimize the predictor with respect to the overall reconstruction error. The second complication is due to the PVQ prediction loop, which has a detrimental impact on the convergence and the stability of the design procedure.

We propose a new PVQ design approach, DA-ACL, that embeds our recent asymptotically closed-loop approach (ACL) [3] within a Deterministic Annealing (DA) framework. The overall DA-ACL method profits from its two main components in a complementary way. DA eliminates the first design shortcoming by offering two benefits: Its probabilistic framework replaces hard quantization with a differentiable expected cost function that can be jointly optimized for the predictor and quantizer parameters; and its annealing schedule allows avoiding many poor local optima. ACL is used to overcome the second difficulty and offers the means for stable quantizer design as it provides an open-loop design platform, yet allows the PVQ design algorithm to asymptotically converge to the objective closed-loop performance.

DA has been successfully applied in several applications including pattern recognition and signal compression [4]. It is motivated by the observation of annealing processes in physical chemistry. Certain chemical systems can be driven to their low energy states by annealing, which is a gradual reduction of temperature, spending a long time in the vicinity of phase transition points. Analogously, we randomize the encoding rule of the predictive quantizer system and seek to minimize the expected distortion cost subject to a specified level of randomness measured by the Shannon entropy. This problem can be formulated as the minimization of a Lagrangian functional that is analogous to the Helmholtz free energy of chemical systems, where the degree of randomness is determined by the Lagrange multiplier (temperature T of the configuration):

$$F = D - TH \tag{1}$$

where D is distortion, and H is entropy.

We start at a high degree of randomness, where the entropy is maximized. Here, the globally optimal configuration requires that all code vectors be coincident at the centroid of the source distribution; no initialization of quantizer or predictor is necessary. We then track the minimum at successively lower levels of entropy, by re-calculating the optimum locations of the reproduction points and the encoding probabilities at each stage.

At each consecutive temperature, the predictor is optimized for the fixed quantizer, and the quantizer is optimized for the fixed predictor, iteratively, until convergence. For the process to converge, it is necessary to base prediction, in open-loop, on a fixed sequence of vectors. We therefore utilize the ACL algorithm where our prediction is based on the fixed reconstructed vectors of the *previous iteration*. By basing prediction on an "older" version of reconstructed vectors, the prediction residuals are in effect calculated in open-loop and we can thus circumvent the otherwise destabilizing effects of the feedback of the closed-loop system. As a direct consequence of this effectively open-loop approach, an optimization procedure is possible for jointly optimizing both quantizer and predictor which is monotone decreasing in the cost. Once such quantizer and predictor have converged, an improved set of reconstructed vectors is produced (also in open-loop, using ACL), the temperature is reduced, and a new iteration is begun.

As the temperature is reduced, the association probabilities become increasingly discriminatory. At the limit of zero randomness, they become "hard" (0 or 1 probabilities), the algorithm directly minimizes the expected reconstruction distortion, and the optimal deterministic PVQ is obtained.

Substantial performance gains over traditional methods have been achieved in the simulations. For example, first-order Gauss-Markov sequences have yielded gains of up to 2 dB. Similar gains were obtained for line spectral frequency parameter sequences for speech coding applications.

[1]This work was supported in part by the NSF under grant no. MIP-9707764, EIA-9986057, the University of California MICRO Program, Conexant Systems, Inc., Dolby Laboratories, Inc., Lucent Technologies, Inc., MedioStream, Inc., and Qualcomm, Inc.

References

[1] V. Cuperman and A. Gersho, "Vector predictive coding of speech at 16 kbits/s," *IEEE Trans. Communications*, vol. COM-33, no. 7, pp. 685-696, July 1985.

[2] P. C. Chang and R. M. Gray, "Gradient algorithms for designing predictive vector quantizers," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. ASSP-34, pp. 679-690, Aug. 1986.

[3] H. Khalil and K. Rose, "Robust Predictive Vector Quantizer Design," *Data Compression Conf.*, Snowbird, UT, pp. 33-42, Mar. 2001.

[4] K. Rose, E. Gurewitz, and G. C. Fox, "Vector quantization by deterministic annealing," *IEEE Trans. Information Theory*, vol. 38, no. 4, July 1992.

0-7803-7123-2/01/$10.00 ©2001 IEEE