

# Optimal Sequential Vector Quantization of Markov Sources

Vivek S. Borkar<sup>1 2</sup>  
 Dept. of Computer Science and  
 Automation  
 Indian Institute of Science  
 Bangalore 560012, India  
 Email: borkar@csa.iisc.ernet.in

Sanjoy K. Mitter<sup>2</sup>  
 Laboratory for Information and  
 Decision Systems  
 M.I.T.  
 Cambridge, MA, 02139  
 Email: mitter@lids.mit.edu

Sekhar C. Tatikonda<sup>2</sup>  
 Laboratory for Information and  
 Decision Systems  
 M.I.T.  
 Cambridge, MA, 02139  
 Email: tatikond@mit.edu

**Abstract** — The problem of optimal sequential vector quantization of Markov sources is cast as a stochastic control problem with partial observations and constraints, leading to useful existence results for optimal codes and their characterizations.

## I. INTRODUCTION

Traditional rate distortion theory resides in the “asymptopia” of arbitrarily long block codes. Nevertheless, there are several situations, notably in control applications, where long block codes imply long delays. Hence there is a need for “real time” rate distortion theory. In this paper, we consider sequential vector quantization followed by coding itself as an optimal control process. An important insight gained is the key role played by the process of conditional laws of the source given its encodings. This process is given by an appropriate recursive nonlinear filter. The optimization problem then translates into an appropriate “certainty equivalent” control problem for the latter process.

## II. SEQUENTIAL VECTOR QUANTIZATION

We assume a stationary Markov source  $X_n, n \geq 0$ , with transition kernel  $p(x, dy)$ , which is to be mapped (vector quantized) into a finite subset  $A$  of its range  $S \subset \mathbb{R}^d$ . The latter is then encoded into a prescribed discrete codebook  $\Sigma$ . Letting  $l_A(\cdot), i_A(\cdot)$ , denote the respective maps, the encoding of  $X_{n+1}$  is  $q_n = i_{Q_n} \circ l_{Q_n}(X_{n+1})$ , where  $Q_n$  is a finite subset of  $S$  chosen based on  $q_i, i \leq n$ . (We impose certain mild technical restrictions on  $i_A$ , see [1] for details.) The problem is to choose  $Q_n$  at time  $n, n \geq 0$ , based on  $q_m, m \leq n$ , so as to minimize the average entropy rate

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{m=0}^{n-1} E[H(q_{m+1}|q_i, i \leq m)]$$

subject to a constraint on the average distortion

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{m=0}^{n-1} E[\|X_m - i_{Q_{m-1}}^{-1}(q_m)\|^2] \leq D,$$

for a prescribed  $D > 0$ . This is the “constrained” version of the sequential rate distortion problem. We also consider the “weighted” version wherein we seek to minimize, for a given  $\lambda > 0$ ,

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{m=0}^{n-1} E[H(q_{m+1}|q_i, i \leq m) + \lambda \|X_m - i_{Q_{m-1}}^{-1}(q_m)\|^2].$$

<sup>1</sup>This work was supported a Homi Bhabha Fellowship.

<sup>2</sup>This research supported by U.S. Army Grant PAAL03-92-G-0115, Center for Intelligent Control Systems.

## III. REDUCTION TO CONTROL PROBLEM

Let  $\pi_n$  denote the conditional law of  $X_n$  given  $q_m, m \leq n$ . Then one verifies that

$$E[H(q_{m+1}|q_i, i \leq m)] = E[k(\pi_n, Q_n)]$$

where

$$k(\pi, A) = - \sum_a h_a(\pi, A) \log h_a(\pi, A),$$

$$h_a(\pi, A) = \int \pi(dx) \int p(x, dy) I_{\{i_A \circ l_A(y)=a\}},$$

( $I$  denotes the indicator function) and

$$E[\|X_n - i_{Q_{m-1}}^{-1}(q_m)\|^2] = E[g(\pi_n, Q_n)]$$

where

$$g(\pi, A) = \int \pi(dx) \int p(x, dy) \|y - l_A(y)\|^2.$$

The problem then reduces to that of optimally controlling the nonlinear filter

$$\pi_{n+1}(dy) = \frac{\int I_{\{i_{Q_n} \circ l_{Q_n}(y)=q_{n+1}\}} p(x, dy) \pi_n(dx)}{\int \int I_{\{i_{Q_n} \circ l_{Q_n}(z)=q_{n+1}\}} p(x, dz) \pi_n(dx)}$$

so as to minimize either

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{m=0}^{n-1} E[k(\pi_m, Q_m)]$$

subject to

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{m=0}^{n-1} E[g(\pi_m, Q_m)] \leq D,$$

or that of minimizing

$$\limsup_{n \rightarrow \infty} \frac{1}{n} \sum_{m=0}^{n-1} E[k(\pi_m, Q_m) + \lambda g(\pi_m, Q_m)].$$

These can be analyzed by stochastic control techniques to yield the existence of a (possibly randomized) optimal vector quantization-encoder scheme and in the weighted case, a dynamic programming characterization of the optimal scheme. Further research includes using approximate dynamic programming techniques for developing these coding schemes and applications to control problems under communication constraints [2].

## REFERENCES

- [1] V.S. Borkar, S. K. Mitter, S. Tatikonda, “Optimal Sequential Vector Quantization of Markov Sources,” MIT LIDS report LIDS-P-2395, August 1997, submitted.
- [2] V.S. Borkar, S. K. Mitter, S. Tatikonda, “Markov Control Problems Under Communication Constraints,” submitted.