Low Complexity Sequential Lossless Coding for Piecewise Stationary Memoryless Sources[1]

Gil I. Shamir[2] and Neri Merhav[3]

Abstract — **A strongly sequential lossless compression scheme with fixed per-letter computational complexity is presented for memoryless sources with abruptly changing statistics. It achieves redundancy of $O(\log N/N)$ when the transitions in the statistics are large and $O\left(\sqrt{\log N/N}\right)$ otherwise.**

I. Introduction

Universal lossless source coding schemes perform better if they take advantage of prior knowledge of the source class. In this paper, we propose a low complexity strongly sequential probability assignment scheme for coding sources from the *simple* non-stationary class of *piecewise stationary memoryless sources* (PSMS's) [1], [2]. The scheme is based on *context tree coding* [3] techniques, and can be used for coding PSMS's when combined with arithmetic coding. The objective is to approach the lower bound on the redundancy R_N [1] of

$$R_N \geq (1-\varepsilon)\left(\frac{r-1}{2}(C+1)+C\right)\frac{\log N}{N}, \tag{1}$$

where N is the sequence length, C is the number of transitions in the statistics, r is the size of the alphabet, and $\varepsilon > 0$ is arbitrarily small for large enough N.

II. Notation and Definitions

A PSMS is characterized by the triplet $\{\Theta, \mathcal{T}, C\}$. The vector Θ is the set of distributions $\{P_0, P_1, \ldots, P_C\}$ of $C+1$ statistically independent stationary segments. The vector $\mathcal{T} \stackrel{\triangle}{=} \{t_1, t_2, \ldots, t_C\}$ represents the *transition path* (TP).

The *pointwise redundancy* of a scheme, that assigns the probability $Q\left(x^N\right)$ to an N-tuple x^N, is defined as

$$R\left(x^N\right) \stackrel{\triangle}{=} \frac{1}{N}\log\frac{P\left(x^N|\Theta,\mathcal{T}\right)}{Q\left(x^N\right)}, \tag{2}$$

where $P\left(x^N|\Theta,\mathcal{T}\right)$ is the true probability of x^N. The redundancy R_N is the mean of $R\left(x^N\right)$.

III. Decision Weighting Scheme

Probability assignment can be obtained by two stage mixture over all possible TP's and values of Θ. Such a mixture can be performed by a *state diagram*, in which state s_n at time n is defined as the time instant of the most recent transition within $1 \leq t \leq n$ [2]. Each state is assigned a weight for x^n, which is updated recursively, and $Q\left(x^n\right)$ is the sum of the weights at time n. A good choice of state transition weights can achieve the PSMS lower bound.

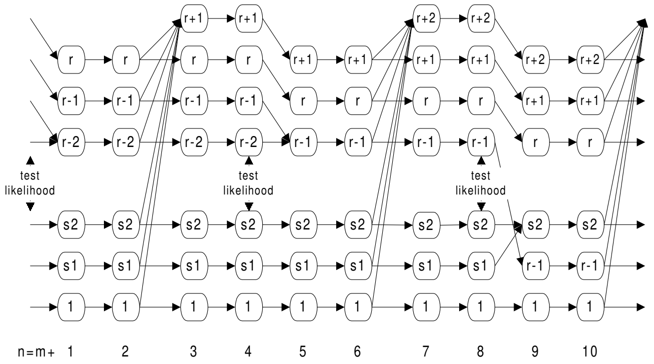

Figure 1: transition diagram for $k=4$ and $S=3$. Block numbers are denoted in state boxes. Diagram starts at time $m+1$, which is the first time point of block $r+1$.

A *reduced state fixed complexity diagram* can be obtained by creating states every block of $k = O(\log N)$ time instants and by eliminating states that represent transitions with low likelihood, keeping only S high likelihood states. A state is created at a block midpoint, is represented by the block index, and represents transitions at any time in the block. Eliminations are performed at block borders. Likelihood is obtained by empirical data in the neighborhood of a block. Figure 1 demonstrates a transition diagram. If a certain distance measure between the probability distributions before and after each transition is sufficiently large, and a fixed bound on the number of transitions is known, it can be shown that for a correct choice of transition weights and likelihood function

$$R_N \leq O(\log N/N). \tag{3}$$

If a TP, that partitions x^N to blocks of length $m_b = \lfloor \alpha b \log b \rfloor$ for $b > 1$ and $m_1 \stackrel{\triangle}{=} 1$, where α is a constant and b is the index of the block, is forced to be weighed in the mixture of TP's, pointwise redundancy $R\left(x^N\right) \leq O\left(\sqrt{\log N/N}\right)$ can be obtained for every x^N.

Acknowledgements

The authors gratefully acknowledge Professor J. Ziv for his helpful suggestions.

References

[1] N. Merhav, "On the minimum description length principle for sources with piecewise constant parameters," *IEEE Trans. Inform. Theory*, Vol. 39, No. 6, pp. 1962-1967, November 1993.

[2] F. M. J. Willems, "Coding for a binary Independent Piecewise-Identically-Distributed source," *IEEE Trans. Inform. Theory*, Vol. 42, No. 6, pp. 2210-2217, November 1996.

[3] F. M. J. Willems, Y. M. Shtarkov and T. J. Tjalkens, "The Context-Tree weighting method: basic properties," *IEEE Trans. Inform. Theory*, Vol. 41, No. 3, pp. 653-664, May 1995.

[1] This work, which partially summarizes the M.Sc. thesis of G. Shamir, was supported by the Israeli Consortium of Ground Stations for Satellite Communications administered by the S. Neaman Institute for Science & Technology and by the Israel Science Foundation administered by the Israeli Academy of Sciences and Humanities.

[2] G. Shamir was with the Dept. EE, Technion, Haifa 32000, Israel. He is now with Dept. EE, University of Notre Dame, Notre Dame, IN 46556, USA. Email: gshamir@nd.edu

[3] N. Merhav is with Dept. EE, Technion, Haifa 32000, Israel. E-mail: merhav@ee.technion.ac.il