First-order multiplexed source codes for error-resilient entropy coding

Hervé Jégou
IRISA/Univ. of Rennes I
e-mail: `Herve.Jegou@irisa.fr`

Christine Guillemot
IRISA/INRIA
e-mail: `Christine.Guillemot@irisa.fr`

Abstract — **This paper describes a new class of variable length codes (VLCs) that allow to exploit first-order source statistics while still being resilient to transmission errors. This paper extends the work of [1] to take into account the source conditional probabilities. Theoretical performances in terms of compression efficiency and error resilience are analyzed.**

I. Introduction

Entropy Coding produces variable length codewords which are very sensitive to channel noise: synchronization losses can occur at the receiver, leading to dramatic error rates. Multiplexed codes, introduced in [1] and referred to as *stationary* multiplexed codes, exploit the fact that compression systems of real signals generate sources of information with different levels of priority. In this paper, we consider two sources: the *high* priority source $\mathbf{S}_H$ and the *low* priority source $\mathbf{S}_L$.

The codes in [1] have been shown to warranty the synchronization for $\mathbf{S}_H$. The risk of "de-synchronization" is confined to the representation of $\mathbf{S}_L$. Unequal error protection is intrinsically provided by the source codes, without considering the use of channel codes. However, stationary multiplexed codes only allow to reach (asymptotically) the stationary entropy of the source. This paper extends these codes to exploit the first-order statistics of the source $\mathbf{S}_H$.

II. Stationary Multiplexed Codes

The high priority source $\mathbf{S}_H = (S_1, \dots S_t, \dots S_{K_H})$ takes its values in a finite alphabet $\mathcal{A} = \{a_1, \dots a_i, \dots a_\Omega\}$ and is assumed to be a first-order Markov process. The stationary and conditional probabilities of the source $\mathbf{S}_H$ are respectively denoted $\mu_i = \mathbb{P}(S_t = a_i)$ and $\nu_{i,i'} = \mathbb{P}(S_t = a_i | S_{t-1} = a_{i'})$, and are assumed to be greater than 0. We also consider a source $\mathbf{S}_L$ of low priority information. This source is assumed to be pre-encoded (e.g. using an arithmetic code).

A *stationary* multiplexed code is constructed by partitioning the space $\mathcal{X} = \{0,1\}^c$ of codewords of length c into Ω subsets $\mathcal{C}_i$ of cardinal $|\mathcal{C}_i|$, called equivalence classes. Each equivalence class $\mathcal{C}_i$ is associated to a symbol a_i of the alphabet $\mathcal{A}$. A codeword $x \in \mathcal{C}_i$ is associated to a symbol a_i of the alphabet $\mathcal{A}$ and to an index value q_i with $0 \le q_i \le |\mathcal{C}_i| - 1$. Thus, a multiplexed code is defined by the bijection

$$\begin{array}{lcl} \bigcup_{a_i \in \mathcal{A}} (\{a_i\} \times [\,0 \,..\, |\mathcal{C}_i| - 1\,]) & \to & \mathcal{X} \\ (\ a_i,\, q_i\ ) & \mapsto & x. \end{array} \tag{1}$$

Hence, each fixed length codeword x represents jointly a symbol a_i of the high priority source and an index value q_i. Since the value q_i is the realization of a $|\mathcal{C}_i|$-valued variable, the amount of data (of $\mathbf{S}_L$) that can be described by this index value is $\log_2(|\mathcal{C}_i|)$ bits. The mean description length (mdl) of the high priority source $\mathbf{S}_H$ is subsequently given by $\sum_{a_i \in \mathcal{A}} \mu_i \,(c - log_2(|\mathcal{C}_i|))$. This quantity is equal to the source stationary entropy iff $\forall i \in [1..\Omega],\ c - \log_2(|\mathcal{C}_i|) = -\log_2(\mu_i)$, i.e. iff $|\mathcal{C}_i| = \mu_i\,|\mathcal{X}|$.

III. First-order multiplexed codes

Higher compression efficiency can be obtained by adapting the codes to first-order source statistics. First-order multiplexed codes can thus be constructed for the conditional probability distribution $(\nu_{i,i'})_{(i,i') \in \mathcal{A}^2}$ of $\mathbf{S}_H$. The partition of $\mathcal{X}$ then becomes conditioned by the realization of the previous symbol $S_{t-1} = a_{i'}$ in the sequence to be encoded. Let $\mathcal{C}_i^{i'}$ denote the equivalence class associated to the symbol a_i when the previous symbol realization is $S_{t-1} = a_{i'}$. The set of codes, for all conditional symbol values from $\mathcal{A}$, defines a so-called *first-order multiplexed code* $\mathcal{C}^*$. The mdl of the code $\mathcal{C}^*$ (for the high priority source $\mathbf{S}_H$) is given by

$$\mathrm{mdl}(\mathcal{C}^*) = -\sum_{(a_i, a_i') \in \mathcal{A}^2} \mu_{i'}\, \nu_{i,i'} \log_2 \left(\frac{|\mathcal{C}_i^{i'}|}{|\mathcal{X}|} \right). \tag{2}$$

which is equal to the first-order entropy of the source iff

$$\forall (a_i,\, a_{i'}) \in \mathcal{A}^2,\ |\mathcal{C}_i^{i'}| = \nu_{i,i'}\, |\mathcal{X}|. \tag{3}$$

The higher compression efficiency may not result in a shorter sequence for $\mathbf{S}_H$, as in classical VLCs, but rather in a higher multiplexing capacity to be used for describing the source $\mathbf{S}_L$.

IV. Error resilience analysis

We consider a global transmission chain formed by the source, the multiplexed source coder, the transmission channel and the decoder. This chain is a Markov process represented by the pair of variables $(S_t, \hat{S}_t)$, where S_t and $\hat{S}_t$ respectively denote the symbol emitted by the source and the symbol reconstructed by the decoder. The transition probability of this Markov process is given by

$$\begin{aligned} &\mathbb{P}(\hat{S}_t = a_j; S_t = a_i | \hat{S}_{t-1} = a_{j'}; S_{t-1} = a_{i'}) = \\ &\qquad \frac{\nu_{i,i'}}{|\mathcal{C}_i^{i'}|} \sum_{(X_t, Y_t) \,\in\, \mathcal{C}_i^{i'} \times \mathcal{C}_j^{j}} R(Y_t, X_t), \end{aligned} \tag{4}$$

where $R(Y_t, X_t)$ represents the channel model, i.e., the probability of receiving the codeword Y_t if the codeword X_t has been emitted. According to the Perron-Frobenius theorem, the stationary probability distribution $\mathbb{P}(S_t, \hat{S}_t)$ is given by the normalized eigenvector associated to the eigenvalue 1 of the transition matrix obtained in Eqn. 4. The SER and MSE performance can then be expressed analytically in terms of the probabilities $\mathbb{P}(S_t, \hat{S}_t)$. These expressions provide optimization criteria for different index assignments (IA) strategies. The IA can be carried out using, e.g., a simulated annealing algorithm [2]. For the SER criteria, we observed that the IA optimization reveals some codewords that act as hard synchronization points, at no cost in compression efficiency.

References

[1] H. Jégou and C. Guillemot, "Source multiplexed codes for error-prone channels," *Proc. IEEE ICC'03*, May, Anchorage.

[2] A.A. El Gamal, L.A. Hemachandra, I. Shperling and V.K. Wei, "Using simulated annealing to design good codes," *Trans. Inform. Theory*, vol. 33, p.116-123, Jan. 1987.