Tree-Structured Multiple Description Coding

JEAN CARDINAL
Université Libre de Bruxelles, CP 212, B-1050 Brussels, Belgium

Received September 6, 2001; Revised April 8, 2002; Accepted May 15, 2002

Abstract. Multiple description coding aims at transmitting two mutually refinable descriptions of a signal on two different communication channels. It is successfully applied to transmission of sound and images on diversity systems or packet-switched networks. We present novel algorithms for the design of multiply descriptive tree-structured vector quantizers with linear encoding complexity that improve on our previously published methods. These algorithms allow for both the rates and distortions constraints in a natural way and are shown to yield quantizers that are competitive with full-search codes in terms of rate-distortion performance. We also present a generalization of the scheme to more than two channels.

Keywords: source coding, multiple description coding, vector quantization, tree-structured vector quantization

1. Introduction

The study of signal transmission in distributed environments naturally leads to practical implementations of schemes known from network information theory (see for instance Thomas and Cover [1], Ch. 14). One of these schemes, Multiple description coding (MDC, [2]) was introduced in 1979 by Gersho, Witsenhausen, Wolf, Wyner, Ziv and Ozarow. In this scheme, a source is encoded by a single encoder on two channels with possibly different capacities. A receiver can either receive the information from both channels or only from one of them, and should be able to provide a good reconstruction of the input in all cases.

We can formulate the problem of MDC as that of minimizing the central distortion, when the two descriptions are received, with constraints on the side distortions, when only one of the description is available. If the constraints are loose, the two descriptions tend to become independent, and the redundancy between them is low. If the constraints are tight it is likely that the descriptions become similar, the mutual information increases and there is less interest in using two descriptions instead of one. Between these two extremes lies a range of tradeoffs. Lower bounds on the achievable rate-distortion tuples for various input sources are identified in several publications. General bounds are given in El-Gamal and Cover [2], generalizing the single-description rate-distortion bound. The special case of an iid Gaussian source has been extensively studied by Ozarow [3].

Although MDC has been studied since the early eighties, it is not until recently that it has been applied to practical data transmission. In diversity systems, for instance, multiple channels are used to overcome the impairments of a single channel. We can also identify the two descriptions as two packets sent on a best-effort network. For practical applications it is useful to generalize the scheme to m descriptions. In [4], a m-description generalization of the El-Gamal and Cover bounds are given.

In a multible description vector quantizer (MDVQ, [5–8]), an input vector \( X \in \mathbb{R}^k \) is encoded by \( \alpha \) with the codebook indices \( i_1 \) and \( i_2 \), satisfying the entropy constraints \( R_1 = H(i_1) \leq R_1^* \) and \( R_2 = H(i_2) \leq R_2^* \). These indices are entropy-coded by \( \gamma_1 \) and \( \gamma_2 \). The three reception scenarios correspond to three different decoders \( \beta_1, \beta_2 \) and \( \beta_0 \). \( \beta_1 \) and \( \beta_2 \) are the side decoders, and \( \beta_0 \) is the central decoder. The corresponding output distortions are denoted by \( D_1, D_2 \) and \( D_0 \) respectively,
with \( D_j = E_X d(X, y_j) \) for a certain distortion measure \( d \) (see Fig. 1).

In full-search MDVQ, the pair \((i_1, i_2)\) of indices is chosen to minimize a weighted sum of distortions. That is,
\[
\alpha(X) = \arg \min_{i_1, i_2} d(X, \beta_0(i_1, i_2)) + \mu_{D_1} d(X, \beta_1(i_1)) \\
+ \mu_{D_2} d(X, \beta_2(i_2)),
\]
where the Lagrangian multipliers \( \mu_{D_1} \) and \( \mu_{D_2} \) can be chosen, for instance, according to channel failure probabilities. It can also be interpreted as a Lagrangian relaxation of the constrained problem described above. It can be shown that for each pair \((\mu_{D_1}, \mu_{D_2})\) there exists a pair of constraints on the side distortions such that the solution of the relaxation is also the solution of the corresponding constrained problem [9]. The three codebooks \( C_0 = \{\beta_0(i_1, i_2)\} \) and \( C_j = \{\beta_j(i_j)\}, j = 1, 2 \) are usually optimized using a generalization of the Generalized Lloyd Algorithm (GLA, [10]), leading to a locally optimal solution.

The proposed scheme, called tree-structured multiple description vector quantizers (TS-MDVQ), integrates the multiple description capability in a tree-structured vector quantizer (TSVQ).

In a TSVQ [11–14], the codebook is organized in a binary tree, each node of which contains a codevector. The encoder simply selects at each node the son node whose codevector is the closest to \( X \). The actual codebook is the set of codevectors contained in the leaf nodes. TSVQ encoding complexity is linear in the bitrate, as opposed to full-search quantizers whose encoding complexity is exponential. It also provides successive refinement: if the index is suitably encoded, one can obtain a good reconstruction of the input by considering only the first bits. This bitrate scalability is important in many multimedia applications. On the other hand the performance of TSVQ is a little worse than that of full-search VQ since the chosen codevector is not necessarily the closest one. An illustration of this method is given in Fig. 2.

Numerous TSVQ design methods have been designed in the last decade. In the simplest one [11], the tree is designed one level at a time by iteratively splitting all the leaf nodes. In the Generalized BFOS algorithm described by Chou [12], a large initial tree is optimally pruned using a marginal return analysis. A simpler and still efficient TSVQ design method was described by Riskin [13]. It recursively splits leaf nodes until the total rate, defined as the leaf entropy, is equal to the target rate. At each iteration, the chosen split node maximizes the marginal return \(-\Delta D/\Delta R\) between the distortion decrease and the rate increase.