A RECURSIVE SIDE-DECODER FOR MULTIPLE DESCRIPTION QUANTIZATION

Pradeepa Yahampath

Department of Electrical and Computer Engineering
University of Manitoba, Winnipeg, Canada
pradeepa@ee.umanitoba.ca

ABSTRACT

Theoretically, the achievable side-distortion of a multiple description (MD) quantizer has a lower bound for a given level of central distortion, but such rate-distortion bounds cannot be achieved in practice. This paper presents a new decoder which exploits the residual redundancy in a sequence of outputs from a sub-optimal MD encoder to reduce side-distortion, without compromising central distortion. A hidden Markov model-based recursive side-decoder is derived for the reconstruction of a source from an incomplete set of descriptions. Simulation results based on encoding a Gauss-Markov source and a simulated speech process are presented to demonstrate the effectiveness of the proposed decoder.

1. INTRODUCTION

Multiple description vector quantization (MDVQ) has received considerable attention due to its potential applications in speech and video coding over packet networks and wireless systems. An M*-channel* MDVQ encoder consists of a VQ encoder (a partition of input vector space into a set of regions) followed by an *index assignment* (IA) which assigns M channel indexes to each encoding region (an ordinary VQ would assign a single index to each region). The MDVQ decoding problem can be viewed as an estimation problem, in which the encoded source vector is estimated based on any sub-set of $m \leq M$ indexes. The fundamental issue in an MDVQ is the trade-off between *central distortion* (average error of a decoder which receives all the descriptions) and *side distortion* (average error of a decoder which receives only a sub-set of descriptions). Rate-distortion theoretic results for 2-channel MDVQ [1] show that, for a given level of central distortion, achievable side-distortion has a lower bound. In an optimal MDVQ, the encoded output on each channel will contain the minimum amount of redundancy required to achieve the lower bound on side-distortion. However, this ideal condition can only be reached in the limit of infinite vector dimension. A finite dimensional (practical) MDVQ will always have some "excess" redundancy on each channel due to the sub-optimality of the encoder. The objective of the work presented in this paper is to reduce the side-distortion of an MDVQ, without penalizing the central distortion, by exploiting at the decoder the *residual redundancy* in the output of a suboptimal encoder. Previously, the use of residual redundancy for joint source-channel coding has been studied by many authors. In [2], Sayood and Borkenhagen (who coined the term *residual redundancy*) used the correlation in the output of a differential pulse code modulation (DPCM) encoder for decoding of images in the presence of channel noise. In [3], the residual redundancy of a VQ encoder is used for decoding a Gauss-Markov source over a discrete memoryless channel (DMC). In these work and in a number of other work, the residual redundancy of the source encoder is implicitly used as a form of error control coding.

This work has been supported by NSERC of Canada.

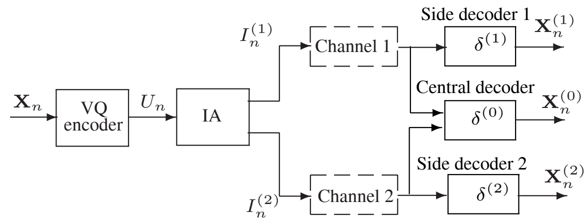

Fig. 1. *A 2-channel multiple description vector quantizer.*

In this paper, it is shown that the residual redundancy of an MDVQ encoder, resulting from sub-optimal encoding of a correlated source, can be used at the decoder to reduce the source reconstruction error when only partial information is received as a result of channel erasures (*e.g.* packet losses). Residual redundancy is characterized by modeling the MDVQ encoder output as a first-order Markov process. The decoding problem is then shown to be related to an estimation problem in hidden Markov models (HMM) [4]. Based on this observation, a recursive decoder is presented, which is well suited for packet-oriented transmission. For simplicity, only 2 channel MDVQ is considered in this paper. However, the decoding method proposed here readily generalizes to M-channel MDVQ.

0-7803-8874-7/05/$20.00 ©2005 IEEE

2. MD QUANTIZATION

We will use upper case letters to denote random variables and lower case letters to denote realizations. A 2-channel MDVQ is shown in Fig. 1. Let the source vector, $\mathbf{X}_n \in \mathbb{R}^k$, be the output of a discrete-time stationary random process, where n denotes the discrete-time. For notational convenience, define $\mathbb{I}_j \triangleq \{1, \ldots, j\}$. An MDVQ encoder can be viewed as a cascade of an ordinary k-dimensional, N-level VQ encoder and an index assignment (IA) [5] which maps each output U_n of the VQ encoder onto an index pair $(I_n^{(1)}, I_n^{(2)})$, where $U_n \in \mathbb{I}_N$, $I_n^{(1)} \in \mathbb{I}_{N_1}$, and $I_n^{(2)} \in \mathbb{I}_{N_2}$, with $N \leq N_1 N_2$. The rate of the MDVQ encoder on the m-th channel is given by $R_m = (1/k)\log_2 N_m$ bits/sample, $m = 1, 2$. Each index is assumed to be transmitted over an independent erasure channel (we assume that channels are memoryless). In order to describe the outputs of the channels, define $\mathbb{I}_{j,\emptyset} \triangleq \{1, \ldots, j, \emptyset\}$, where $\emptyset$ denotes the null-output in the event of an erasure. Then the output of the m-th channel can be described by $\hat{I}_n^{(m)} \in \mathbb{I}_{N_m,\emptyset}$, $m = 1, 2$.

In this paper we will use square error as the distortion measure. Let the decoder input be $V_n \triangleq (\hat{I}_n^{(1)}, \hat{I}_n^{(2)})$. Then the optimal decoder $\delta_n(V_n)$ minimizes the mean square error (MSE) MSE $D_0 = E\|\mathbf{X}_n - \delta(V_n)\|^2$. If we assume that the decoders are memoryless and stationary, the optimal decoder is given by the centroid condition [5]

$$\delta^*(v) = E\{\mathbf{X}_n | V_n = v\} = \sum_{u=1}^{N} \mathbf{g}_u P(U_n = u | V_n = v), \tag{1}$$

where $\mathbf{g}_u = E\{\mathbf{X}_n | u\}$. When channel erasures occur, this decoder optimally estimates the unknown source vector based on partial information about the encoder output. That is, it estimates $\mathbf{X}_n$ based on either $I_n^{(1)}$ or $I_n^{(2)}$. When some information is lost due to channel erasures, one must fully utilize all the received information to estimate the unknown source vector. This means that, if the sequence $\{U_n\}$ is correlated, then for example, the complete sequences $\{I_n^{(1)}\}$ and $\{I_n^{(2)}\}$ will carry information about $\mathbf{X}_n$. Such correlation exists in the VQ encoder output if the source is correlated and the encoder is not optimal for the source. Since all practically designed VQ encoders are typically sub-optimal, this situation can be expected to exist in most cases. The residual correlation in the VQ encoder output can be used by a side-decoder to better estimate the encoded source vector. In the following section such a decoder is presented.

3. RECURSIVE SIDE-DECODER

Consider a packet-based communication system, in which each block of source vectors $(\mathbf{x}_0, \ldots, \mathbf{x}_{L-1})$, is transmitted in a single packet. For convenience, let the time origin $n = 0$ be the start of a packet. Then, a packet-based MDVQ is assumed to operates as follows. Each block of L consecutive outputs from the VQ encoder, $(u_0, \ldots, u_{L-1})$ is assembled into two packets $\pi^{(1)} = (i_0^{(1)}, \ldots, i_{L-1}^{(1)})$ and $\pi^{(2)} = (i_0^{(2)}, \ldots, i_{L-1}^{(2)})$, which are transmitted over two independent channels. Thus, each *source packet* $(\mathbf{x}_0, \ldots, \mathbf{x}_{L-1})$ is encoded into two *channel packets* $\pi^{(1)}$ and $\pi^{(2)}$. Now consider the situation in which, at least one channel packet is received by the decoder. In this case, the input to the decoder is a sequence $(v_0, \ldots, v_{L-1})$, where $v_n = (\hat{i}_n^{(1)}, \hat{i}_n^{(2)})$ and $\hat{i}_n^{(m)} \in \mathbb{I}_{N_m,\emptyset}$, $m = 1, 2$. Denote $(v_0, \ldots, v_{L-1})$ by v_0^{L-1}. Now, given that v_0^{L-1} can be used to decode $\mathbf{X}_n$, we wish to find the decoder $\delta_n(v_0^{L-1})$ which minimizes the MSE

$$E\{\|\mathbf{X}_n - \delta_n(v_0^{L-1})\|^2 | v_0^{L-1}\}. \tag{2}$$

It is known from classical estimation theory that the minimum MSE (MMSE) estimate is given by

$$\begin{aligned} \delta_n^*(v_0^{L-1}) &= E\{\mathbf{X}_n | v_0^{L-1}\} \\ &= \sum_{u_0^{L-1}} E\{\mathbf{X}_n | u_0^{L-1}\} P(u_0^{L-1} | v_0^{L-1}) \end{aligned} \tag{3}$$

The exact implementation of this decoder is infeasible as the sum is taken over all possible N^L encoder output sequences of length L (typically, packet length $L >> 1$). In general, one can approximate $E\{\mathbf{X}_n | u_0^{L-1}\} \simeq E\{\mathbf{X}_n | u_{n-n_1}^{n+n_2}\}$, for some sufficiently large positive integers n_1, n_2. However, this would result in an increase of decoder complexity that is exponential in n_1 and n_2, for a marginal decrease in MSE. Hence we assume in this paper that $n_1 = n_2 = 0$, that is $E\{\mathbf{X}_n | u_0^{L-1}\} \simeq E\{\mathbf{X}_n | u_n\}$. Then, we obtain the following generalization of the memoryless decoder in (1).

$$\tilde{\delta}_n^*(v_0^{L-1}) = \sum_{u_n} \mathbf{g}_{u_n} P(u_n | v_0^{L-1}) \tag{4}$$

It should be clear that the MSE of this decoder cannot be worse than that of (1). In fact, the simulation results presented in this paper show that it can be substantially less in some cases of practical interest.

Let us observe that, if the VQ encoder output $\{U_n\}$ is uncorrelated, $P(u_n | v_0^{L-1}) = P(u_n | v_n)$ and (4) is identical to (1). However, since a VQ encoder does not completely eliminate the correlation in the source, the resulting residual correlation will be reflected in the posterior probability distribution $P(u_n | v_0^{L-1})$. In this paper we will assume that the encoder residual redundancy can be represented by a first-order Markov process, *i.e.* $P(u_n | u_{n-1}, \ldots) = P(u_n | u_{n-1})$. This assumption is reasonable if the source is approximately Markov and the encoded output is a good approximation for the source. Regardless, with $\{V_n\}$ being an incomplete observation of $\{U_n\}$, one can then view the encoder output as a HMM at the decoder and use the *forward-backward* algorithm [4] to conveniently compute the posterior probabilities in (4). This is the basis of the decoder presented below.

First consider that

$$P(u_n|\mathbf{v}_0^{L-1}) = \frac{P(u_n, \mathbf{v}_0^{L-1})}{P(\mathbf{v}_0^{L-1})} = \frac{P(\mathbf{v}_{n+1}^{L-1}|u_n)P(u_n, \mathbf{v}_0^n)}{P(\mathbf{v}_0^{L-1})} \tag{5}$$

where $P(\mathbf{v}_0^{L-1}) = \sum_{u_n} P(\mathbf{v}_{n+1}^{L-1}|u_n)P(u_n, \mathbf{v}_0^n)$ and we use the fact that given the current state, the future outputs of a Markov process are independent of the previous outputs. Now, using the same notation as in [4], let

$$\begin{aligned} \alpha_n(u_n) &\triangleq P(u_n, \mathbf{v}_0^n) \\ \beta_n(u_n) &\triangleq P(\mathbf{v}_{n+1}^{L-1}|u_n). \end{aligned} \tag{6}$$

Then, (5) can be expressed as

$$P(u_n|\mathbf{v}_0^{L-1}) = \frac{\alpha_n(u_n)\beta_n(u_n)}{\sum_{u'_n} \alpha_n(u'_n)\beta_n(u'_n)} \tag{7}$$

The quantities $\alpha_n(u_n)$, $u_n \in \mathbb{I}_N$ can be computed using the *forward* recursions

$$\alpha_n(u_n) = P(\mathbf{v}_n|u_n) \sum_{u_{n-1}} P(u_n|u_{n-1})\alpha_{n-1}(u_{n-1}) \tag{8}$$

for $n = 0, \ldots, L-1$, and the quantities $\beta_n(u_n)$, $u_n \in \mathbb{I}_N$ can be computed using the *backward* recursions

$$\beta_n(u_n) = \sum_{u_{n+1}} P(\mathbf{v}_{n+1}|u_{n+1})P(u_{n+1}|u_n)\beta_{n+1}(u_{n+1}) \tag{9}$$

for $n = L-2, \ldots, 0$ (a closer look will reveal that $\beta_{L-1}(u) = 1$ for $u \in \mathbb{I}_N$). If the received packets are assumed to be free of bit errors, $P(\mathbf{v}_n|u_n)$ values are either 0 or 1 and are solely determined by the IA mapping in the MDVQ encoder. In the simulations presented in Sec. 4, we will assume this to be the case.

An important issue related to forward recursion is the initialization. At the beginning of a packet period, $\alpha_{-1}(u)$, $u \in \mathbb{I}_N$ have to be initialized. In decoding a sequence of packets, the initialization is necessary only at the beginning of the sequence, provided that at least one channel packet is received for every packet period. On the other hand when both channel packets get lost in a packet period, the decoder has to be arbitrarily initialized for the following packet period. The frequency of this happening would increase with the increasing packet loss probability. In such situations, one solution would be to include at the beginning of each channel packet, the value of u_0.

4. SIMULATION RESULTS AND DISCUSSION

In this section, the performance of the proposed residual redundancy-based decoder is evaluated and compared using simulations. Two different signal sources have been used in the simulations: (i) the first-order Gauss-Markov (GM) process, given by $X_n = 0.9X_{n-1} + W_n$, where W_n is an iid

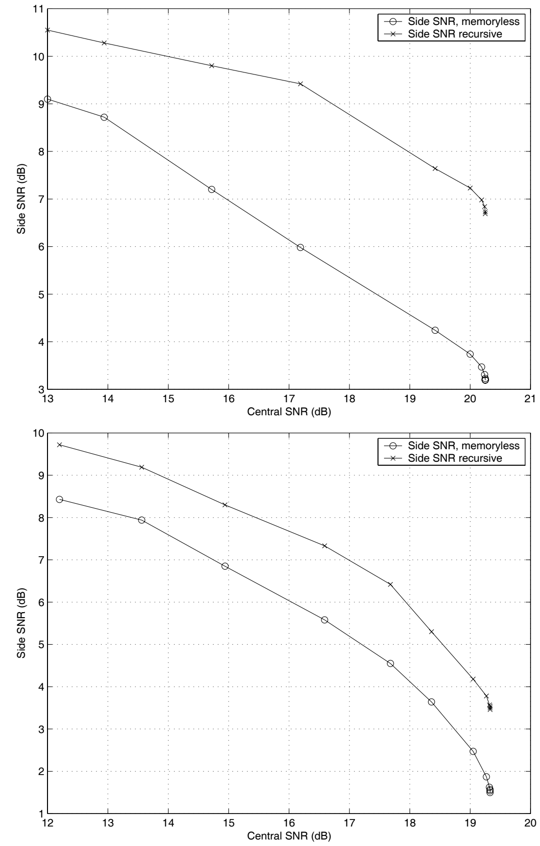

Fig. 2. *The average side SNR as a function of central SNR for 2-channel MDSQ of the GM source (top) and the simulated speech process (bottom). The points on the curves correspond to packet-loss probabilities 0.0001 (right-hand end), 0.0002, 0.0005, 0.001, 0.002, 0.005, 0.01, 0.02, 0.05, 0.1, 0.2, 0.5 (left-hand end).*

Gaussian process and (ii) a simulated speech process based on the third-order Markov process [6]:

$$X_n = 1.748X_{n-1} - 1.222X_{n-2} + 0.301X_{n-3} + Z_n, \tag{10}$$

where Z_n is an iid Laplace process. The simulated speech process in (10) has a correlation coefficient of 0.864 [6]. Only 2-channel systems ($M = 2$) with equal transmission rates ($R_1 = R_2$) and identical packet loss probabilities (P_L) have been considered. In all simulations, a packet length of $L = 128$ source vectors has been used.

In the first set of experiments, the performance of the memoryless decoder in (1) and the residual redundancy based recursive decoder in (4) (referred to hereafter as the *recursive decoder*) is compared by considering 2-channel, 2 bits per sample per channel MD scalar quantization (MDSQ). In the following, the performance of a decoder is measured by its output *signal to noise ratio* (SNR). The average side-SNR as a function of central SNR at different packet loss probabilities is plotted in Fig. 2 for both types of decoders (recall that both decoders yield the same central distortion when the same encoder is used). For a given source and a given value of P_L, the MDSQ encoder was designed by iterative codebook optimization as described in [5], with an initial IA obtained by simulated annealing as described in [7]. The substantial improvement in side SNR due to the utilization of residual redundancy in decoding is quite clear in both cases. For example, an improvement of about 3.5 dB is obtained for the GM process at $P_L = 0.01$ (central SNR about 20 dB). This improvement tends to decrease at higher loss probabilities, since under such conditions, the two descriptions produced by the optimal MD encoder are very similar and little information is lost due to the loss of a description.

In the next set of experiments, the issue of when residual redundancy-based decoding can be expected to be useful over memoryless decoding is addressed. In general, one could expect the residual redundancy-based decoding to be effective when the degree of sub-optimality of the encoder is higher. A particular case of interests is when the encoder is a uniform quantizer. In general, any VQ encoder may be converted into an MDVQ encoder using an optimized IA. This approach allows us to obtain an MD scalar quantizer using a uniform quantizer, *i.e.* an encoder with uniform quantizing intervals. This may be useful, as uniform quantizers are widely used in practice due to their simplicity. However, a uniform quantizer, being optimal only for a uniformly distributed source, may contain a considerable residual redundancy when used with a non-uniform source. It is therefore of interest to investigate if the performance loss resulting from sub-optimality of a uniform quantizer can be compensated for by using the residual redundancy in the decoder. To this end, a uniform MDSQ encoder was designed for a given packet loss probability, by optimizing (using simulated annealing [7]) the IA of a uniform quantizer for a Gaussian source (taken from [8]). However, when the loss probabilities of the channels are identical, the optimal IA is independent of this loss probability [7]. In Fig. 3, we compare the total average SNR of a system which uses this MDSQ encoder (we call it MDSQ1) and that of a system which uses an MDSQ encoder obtained by iterative codebook optimization as in [5] and [7] (we call it MDSQ2). It is interesting to note that, when recursive decoding is used, the average SNR of MDSQ1 is very similar to that of MDSQ2 at loss probabilities above 2%. Since MDSQ1 is obtained without altering the encoder partition, unlike MDSQ2, the central and side SNRs of MDSQ1 are independent of the packet loss probability. Consequently, the central SNR of MDSQ1 is substantially higher than that of MDSQ2 at higher loss probabilities. Furthermore the higher residual redundancy in the MDSQ1 encoder results in a relatively higher side SNR with recursive decoding. Thus, as the packet loss probability is increased, average SNR of MDSQ1 becomes similar to that of MDSQ2.

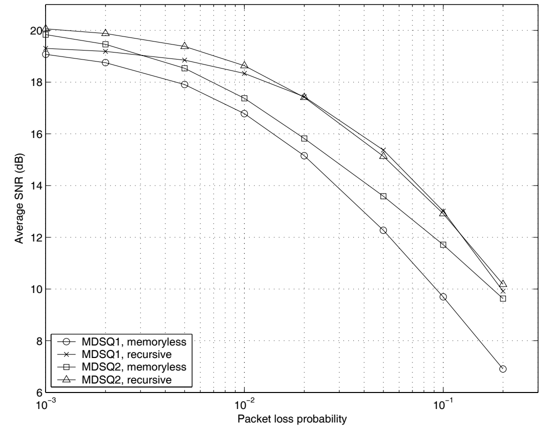

Fig. 3. *The average SNR as a function of packet-loss probability for uniform MDSQ (MDSQ1) and optimal MDSQ (MDSQ2) for the GM source.*

5. CONCLUDING REMARKS

The decoder presented here may be useful in speech and image coding over packet networks, where the encoder cannot usually be optimized to exact source or channel statistics. However, the on-line estimation of encoder Markov model parameters at the decoder is essential when the source statistics are either unknown or time-varying. This issue will be addressed in a future work.

6. REFERENCES

[1] A. A. El Gammal and T. M. Cover, “Achievable rates for multiple descriptions,” *IEEE Trans. Inform. Theory*, vol. IT-28, pp. 851–857, Nov. 1982.

[2] K. Sayood and J. C. Borkenhagen, “Use of residual redundancy in the design of joint source/channel coders,” *IEEE Trans. on Commun.*, vol. 39, pp. 838–846, June 1991.

[3] N. Phamdo and N. Farvardin, “Optimal detection of discrete Markov sources over discrete memoryless channels- applications to combined source-channel coding,” *IEEE Trans. on Inform. Theory*, vol. 40, no. 1, pp. 186–193, January 1994.

[4] L. R. Rabiner and B. H. Huang, “An introduction to hidden Markov models,” *IEEE ASSP Magazine*, pp. 4–16, January 1986.

[5] V. A. Vaishampayan, “Design of multiple description scalar quantizers,” *IEEE Trans. on Inform. Theory*, vol. 39, no. 3, pp. 821–834, May 1993.

[6] J. B. Anderson and J. B. Bodie, “Tree encoding of speech,” *IEEE Trans. Inform. Theory*, vol. 21, no. 4, pp. 379–387, July. 1975.

[7] P. Yahampath, “On index assignment and the design of multiple description quantizers,” in *Proc. ICASSP*, May 2004, pp. 597–600.

[8] J. Max, “Quantizing for minimum distortion,” *IRE Trans. on Inform. Theory*, vol. IT-6, pp. 7–12, March 1960.