Joint Source-Channel Coding of LSP Parameters for Bursty Channels

José L. Pérez-Córdoba, Antonio M. Peinado, Angel M. Gómez, Antonio J. Rubio

Dpto. Teoría de la Señal, Telemática y Comunicaciones.
Universidad de Granada. 18071 Granada (Spain)
jlpc/amp/amgg/rubio@ugr.es

Abstract

This work presents a joint source-channel technique based on Channel Optimized Vector Quantization (COVQ) for transmission over bursty channels applied to LSP parameters coding. The bursty channel is modeled as a Finite State Channel (FSC) with two states. We call Bursty COVQ (BCOVQ) to the resulting quantization technique. The case in which channel state information is only available at the receiver is considered. The application of BCOVQ to LSP (Line Spectrum Pair) parameters coding in wideband speech coding is presented. Our experimental results show that BCOVQ maintains a good performance through all tested channel conditions, while other reference techniques (traditional COVQ, Split Vector Quantization and Split Multi-Stage Vector Quantization) are more dependent on these conditions.

1. Introduction

Channel errors are usually avoided through the use of channel coding. However, this method yields an increase in the required bit rate. Joint source-channel coding techniques can be applied to avoid this increase while maintaining the error protection capacity. One of these techniques, Channel Optimized Vector Quantization (COVQ) [1] has arisen an special interest during the last years. COVQ codebooks are optimized for a channel condition, so a channel mismatch condition appears when these codebooks are used under different channel conditions resulting in a performance degradation.

In some real situations (for example, wireless communications or IP networks) errors tend to appear concentrated into bursts. In such conditions, it is expected that COVQ can not work adequately since it has to work in conditions for which it is not designed.

One approach to combat errors due to burst is to transmit more than one description of the source over different channels ([2]) but this implies an increase on the gross bit-rate. Another possibility is the approach presented in reference [3], where a study of optimal scalar quantization for Finite State Channels (FSC) is discussed. Bursty channels are treated as a special case of a FSC with two states: a "good" state, which is the normal mode of operation, and a "bad" state, in which the channel enters occasionally. This way, the occurrence of a burst only implies the use of a different pair of encoder-decoder.

This paper presents a generalization of [3] to the case of vector quantization (VQ). We develop a similar formulation and present optimal conditions for the encoder and decoder which gives rise to a design algorithm. We call Bursty COVQ (BCOVQ) the resulting technique. The performance of BCOVQ is measured over an application such as speech coding. Specifically, we applied BCOVQ to LSP parameters [4] coding when transmission is over a bursty channel. For comparison purposes, we also study the performance of Split VQ [5] and Split Multi-Stage Vector Quantization (MSVQ) [6].

This paper is organized as follows. In Section 2, the fundamentals of the BCOVQ technique are given. We present the optimization problem and obtain optimum expressions. Section 3 presents the application of BCOVQ to LSP parameters coding. The performance results of the studied techniques are reported in Section 4. Finally, Section 5 contains a summary of this paper.

2. COVQ for bursty Channels

To study the BCOVQ technique, the system model to optimize is the one represented in Figure 1. We consider a real-valued independent and identically distributed (i.i.d.) source \( X = \{X_n\}_n \), with probability density function (pdf) \( p(x) \). The source is to be encoded by means of a vector quantizer (VQ) whose output is transmitted over a waveform channel. We consider an \( M \)-level VQ system with vectors of dimension \( N \).

As Figure 1 shows, the coding system consists of two encoder functions, a signal selection module and two decoder functions. The encoder is described in terms of a partition of \( \mathbb{R}^N \). Since the transmitter has state information, it will use a different partition for each state of the channel. Let \( S_l = \{S_{l,1}, S_{l,2}, \ldots, S_{l,M}\}, l = 1, 2 \) be the two partitions of the space. The encoding functions \( (\gamma_i, l, i = 1, 2) \) are described according to

\[
\gamma_i(x) = i \quad \text{if } x \in S_{l,i}, \quad i \in I, \quad l = 1, 2
\]

where \( x \) is a typical source output vector and \( I = \{1, 2, \ldots, M\} \). The signal selection module maps an index \( i \) into a signal \( s \) that is transmitted over the channel.

We consider that the channel is an Additive White Gaussian Noise (AWGN) Channel with two states. Let \( t_{ji} (r_{ji}) \) be the probability that the state \( s_j \) is perceived as \( s_j \) at the transmitter (receiver), \( j, l = 1, 2 \). \( t_{ji} \) and \( r_{ji} \) form the elements of the

![Figure 1: Block diagram of the transmission system.](image-url)
channel state transition matrices $T$ and $R$, respectively (Figure 1).

Restricting our study to the case in which the decoder performs hard-decision decoding, the signal selection module and the channel block can be approximated by a discrete channel (with two states) with transition probabilities $P_{x_i}(k|i)$ that is the probability of receiving channel output $k$ when the input is $i$ and the channel is in the state $s_j$, $j = 1, 2$; $i, k \in T = \{1, 2, \ldots, M\}$.

The decoder also has access to channel state information. For example, the presence of a burst can be detected through a CRC check. Therefore, the decoder will use two different decoding functions, $\beta_j$, $j = 1, 2$ based on available channel state information. $\beta_j$ makes an estimate $\hat{x}$ of the transmitted source vector based on the received vector (channel output) $r$. Actually, the decoder $\beta_j$ makes an estimate $\hat{r}_j$ of the transmitted index $i$ (hard-decision decoder). Given $\hat{r}_j$, the estimate $\hat{x}$ is selected from a finite reproduction alphabet (codebook) $C_j = \{C_{j,1}, C_{j,2}, \ldots, C_{j,M}\}$ ($j = 1, 2$) that describes the decoder through

$$\beta_j(\hat{r}_j) = \beta_j(\hat{r}(r)) = c_{\hat{r},j} \in \mathbb{R}^N, \; \bar{i} \in T, \; j = 1, 2 \tag{2}$$

Our goal is to design the quantizers in such way to minimize the distortion

$$D_1(S_1, C_1) = E \left[ \|x - \hat{x}\|^2 \right] \text{ channel is in state } 1 \tag{3}$$

subject to the constraint

$$D_2(S_2, C_2) = E \left[ \|x - \hat{x}\|^2 \right] \text{ channel is in state } 2 \leq D \tag{4}$$

where $D$ is the maximum allowed distortion when the channel is in state 2.

### 2.1. System Optimization

The usual approach to resolve this problem is to convert this constrained optimization problem to an equivalent unconstrained optimization problem by minimizing the Lagrangian

$$L = D_1 + \lambda(D_2 - D) \tag{5}$$

where $\lambda$ is the Lagrange multiplier.

For the case of VQ, the equivalent expression for the Lagrangian given in [3], Equ. (3), is

$$L = \sum_{j=1}^{M} \lambda_j \sum_{m=1}^{2} \left( \sum_{t_{m,j}} \sum_{i=1}^{M} P_{x_i}(i) \right) \int_{\hat{r}_{m,j}} p(x) \|x - \beta_m(i)\|^2 dx - \lambda D \tag{6}$$

where $\lambda_1 = 1$, $\lambda_2 = \lambda \geq 0$ and $p(x) = \prod_{i=1}^{N} p(x_i)$ is the N-dimensional source pdf. As in references [7] and [3], we fix the Lagrange multiplier $\lambda$. We also consider one of the special cases studied in [7]. It is the case where the channel state information is only available at the receiver, which is specified by the following values of the channel state transition matrices $T$ and $R$:

$$T = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}, \quad R = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

These particular values involve that the quantization regions at the transmitter are the same ($S_1 = S_2$). At the receiver, two different codebooks are maintained ($C_1$ and $C_2$).

From expression (6) and the particular system conditions, the minimization procedure results in the following optimum expressions (which are a generalization to VQ of those given in [3]):

1) For fixed codebooks $C_1$ and $C_2$, the optimum partition $S^* = S_1^* = S_2^* = \{S_1, S_2, \ldots, S_M\}$ is given by

$$S^*_{si} = \left\{ x : 2a_i x - b_i \geq 2a_i x - b_i', \forall i' \neq i, i', i \in I \right\} \tag{7}$$

where parameters $a_i$ and $b_i$ (for $i \in I$) are given by

$$a_i = \frac{1}{2} \sum_{j=1}^{M} P_{x_j}(i) \beta_j(i) + \frac{\lambda}{2} \sum_{j=1}^{M} P_{x_j}(i) \beta_j(i) \tag{8}$$

$$b_i = \frac{1}{2} \sum_{j=1}^{M} P_{x_j}(i) \beta_j(i)^2 + \frac{\lambda}{2} \sum_{j=1}^{M} P_{x_j}(i) \beta_j(i)^2 \tag{9}$$

2) Similarly, for a fixed partition, optimal codebooks $C_{j}^* = \{c_{j,1}^*, c_{j,2}^*, \ldots, c_{j,M}^*\}$ with $j = 1, 2$, are given by

$$c_{j,i}^* = \frac{\sum_{j=1}^{M} P_{x_j}(i) \int_{S_{si}} p(x) dx}{\sum_{j=1}^{M} P_{x_j}(i) \int_{S_{si}} p(x) dx} \tag{10}$$

The successive application of (7) and (8) results in a sequence of encoder-decoder pairs which converges to a local minimum as the algorithm proposed in [3] does.

### 3. BCOVQ FOR LSP PARAMETERS

In this Section we describe how the BCOVQ technique is applied to LSP parameters coding. In order to apply BCOVQ for LSP parameters coding, these are obtained by performing an LP analysis similar to that performed in the AMR-WB standard codec [6]. The LP analysis is carried out producing a vector of 16 coefficients.

The LP coefficients are quantified using the LSP representation. A first order MA prediction is applied and the residual is vector quantized using split BCOVQ. That is, the residual vector is split into 8 subvectors of dimension 2. These are quantized with BCOVQ with 5, 7, 8, 7, 6, 5, 4 and 4 bits respectively (Figure 3). In this paper, in that table, column marked as "VQ" describes the spectral analysis applied for BCOVQ, COVQ and Split VQ techniques and column marked as "MSVQ" describes the spectral analysis for the Split MSVQ technique.
We have used 960 speech files from the TIMIT database for training the quantization codebooks and 192 files out of training also from the TIMIT database to measure the performance of the simulated coders. For BCOVQ codebook design we have considered a bursty channel with two states: a “good” state S1 with a Channel Signal to Noise Ratio (CSNR) per bit of 6 dB and a “bad” state S2 with a CSNR per bit of -3 dB. For COVQ codebook design we have considered two CSNR per bit (6 and -3 dB) and the optimization was performed considering an slow-fading Rayleigh Channel [8], channel model that is used to characterize the fades of a radio channel.

4. RESULTS AND DISCUSSION

In this section results on the performance of the considered LSP quantization techniques are reported. Two distortion measures are used as performance measure, average spectral distortion (SD) and signal to noise ratio (SNR (dB) = 10 log_{10} \sigma_s^2/\sigma_e^2). As it is discussed in [9], average SD can be a misleading measure for certain channel models due to the small differences in obtained values of this distortion measure for the different experiments, meanwhile, subjective tests show a better quality in synthetic speech of those techniques that performs joint source-channel optimization. This is the reason why we also use the SNR. Average SD allows us to compare obtained performance results with those reported in other papers, while SNR gives a better reference to compare the implemented experiments.

Five different experiments are considered. Figure 2 and Table 2 show results for the average SD and Figure 3 and Table 3 show results for the SNR. The S-MSVQ experiment carries out a Split MSVQ of LSP parameters, in the same way as in the AMR-WB codec. SVQ denotes a Split VQ of the residual LSP vectors. BCOVQ-S2(-3)S1(6) experiment represents the application of Split BCOVQ to LSP quantization when the quantization codebooks are trained at the CSNR of the states S1 and S2 (6 and -3 dB, respectively). Finally, COVQ-(X) experiment represents the application of Split COVQ to LSP quantization when quantization codebooks are trained at a CSNR of X dB.

Performance results are obtained simulating a bursty channel as it is done in [10]. In that work a procedure to obtain the equivalent burst length for a given CSNR is developed. The burst lengths used in the present work are obtained for several CSNR per bit (6, 3, 0 and -3 dB). A CSNR per bit of 12 dB represents a noiseless channel condition.

From Figures 2 and 3, it is clear that joint source-channel
Table 3: SNR for different CSNR per bit. Comparison between coders.

<table>
<thead>
<tr>
<th>CSNR per bit (dB)</th>
<th>N</th>
<th>6</th>
<th>3</th>
<th>0</th>
<th>-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>S-MSVQ</td>
<td>43.3</td>
<td>37.0</td>
<td>29.4</td>
<td>25.1</td>
<td>21.3</td>
</tr>
<tr>
<td>SVQ</td>
<td>44.5</td>
<td>39.3</td>
<td>31.8</td>
<td>27.2</td>
<td>23.6</td>
</tr>
<tr>
<td>BCOVQ-S2(-3)S1(6)</td>
<td>40.9</td>
<td>39.2</td>
<td>35.0</td>
<td>30.4</td>
<td>26.7</td>
</tr>
<tr>
<td>COVQ-(6)</td>
<td>41.6</td>
<td>40.3</td>
<td>34.0</td>
<td>29.1</td>
<td>25.2</td>
</tr>
<tr>
<td>COVQ(-3)</td>
<td>38.4</td>
<td>38.1</td>
<td>35.0</td>
<td>30.8</td>
<td>27.2</td>
</tr>
</tbody>
</table>

Table 3: SNR for different CSNR per bit. Comparison between coders.

coding techniques get better performance results under noisy channel conditions than Split VQ or Split MSVQ, as expected. But, focusing on BCOVQ, we observed that for a noiseless channel it can be obtained a good performance although SVQ and S-MSVQ get better results. However, for a very noisy channel, the results obtained with BCOVQ are significantly better than those of SVQ and S-MSVQ for both distortion measures.

Comparing the obtained results for both distortion measures (Tables 2 and 3), it is clear that the SNR shows better the differences between results of the implemented techniques, especially between BCOVQ and COVQ.

With the BCOVQ technique the performance results are close to those obtained with COVQ when the design condition matches the channel condition as Figure 3 shows or Table 3 gives numerically. It can be observed that for a very noisy channel condition (a CSNR of -3 dB) the SNR for BCOVQ-S2(-3)S1(6) is close to the SNR for COVQ-(6). However, at a CSNR of 6 dB, the SNR for BCOVQ-S2(-3)S1(6) is closer to the SNR for COVQ-(6) than the corresponding SNR of COVQ-(6).

Table 4 compares performance results of BCOVQ-S2(-3)S1(6) experiment with those obtained with COVQ when the codebooks are trained at a CSNR which matches the channel condition. It is observed that the differences in SNR are small or even BCOVQ gets better results.

<table>
<thead>
<tr>
<th>CSNR per bit (dB)</th>
<th>N</th>
<th>6</th>
<th>3</th>
<th>0</th>
<th>-3</th>
</tr>
</thead>
<tbody>
<tr>
<td>BCOVQ-S2(-3)S1(6)</td>
<td>40.9</td>
<td>39.2</td>
<td>35.0</td>
<td>30.4</td>
<td>26.7</td>
</tr>
<tr>
<td>COVQ</td>
<td>42.9</td>
<td>40.3</td>
<td>34.4</td>
<td>30.2</td>
<td>27.2</td>
</tr>
</tbody>
</table>

Table 4: Comparison in SNR for BCOVQ and COVQ when the design condition matches the channel condition.

An informal listening test confirms results from the SNR distortion measure. This subjective test shows a better quality in synthetic speech using the BCOVQ technique, compared to the other tested techniques, when a bursty channel model is used.

5. SUMMARY

In this paper, the COVQ technique is generalized to the case of a bursty channel considered as a Finite State Channel (FSC) with two states. The corresponding technique, Burst COVQ (BCOVQ), shows good performance results under noisy and noiseless channel conditions compared with COVQ or techniques that do not consider channel errors. Besides, BCOVQ shows robustness against channel mismatch.