AN EFFICIENT MMSE DECODER FOR DIFFERENTIAL SOURCE CODERS WITH ADAPTIVE QUANTIZER: A JOINT SOURCE CHANNEL CODING PERSPECTIVE

Behzad Asadi and Farshad Lahouti

Wireless Multimedia Communications Lab,
School of ECE, College of Engineering, University of Tehran
b.asadi@ece.ut.ac.ir, lahouti@ut.ac.ir

ABSTRACT

A minimum mean square error (MMSE) source decoder for the reconstruction of a signal, encoded using a differential source code with an adaptive quantizer, and transmitted over a noisy channel is proposed. One group of the previous related works on MMSE-based joint source channel decoding aim to reconstruct the codec parameter rather than the original signal, which is suboptimal. The other group that aim to reconstruct the original signal have not considered the adaptive differential encoder. The proposed solution results in a two stage MMSE decoder. The decoder benefits from residual redundancy at the output of the source encoder modeled by a γ-order Markov model. The results demonstrate the effectiveness of the proposed scheme.

Index Terms— JSCC (joint source channel coding), MMSE, Markov chain, adaptive differential encoder.

1. INTRODUCTION

Using the residual redundancy in the output of the source encoder during the decoding process is a well-known class of joint source channel coding techniques [1]. This redundancy is due to the constraints on delay and computational complexity of the source encoder.

The performance of both source decoder and channel decoder can be improved using this redundancy. The redundancy is used to improve the source decoder performance in e.g., [1-6] and to improve the channel decoder performance in e.g., [7]. A conventional approach to capture the residual redundancy is through the use of a Markov model.

When the performance measure is the SNR of the reconstructed signal, the optimal decoder is designed for minimum mean square error (MMSE) reconstruction of the original signal (input to the encoder). One group of the previous related works on MMSE based joint source channel decoding aim to reconstruct the codec parameter rather than the original signal, which is suboptimal. The other previous related works that aim to reconstruct the original signal have not considered the adaptive differential encoder. In [3] and [4] MMSE-based decoders have been designed to efficiently reconstruct the original signal of a differential pulse code modulation (DPCM) encoder. In [2], a MMSE-based decoder has been designed to reconstruct the codec parameters of an adaptive differential pulse code modulation (ADPCM) encoder. As stated, this solution is suboptimal in the MMSE sense.

In this paper, we consider the problem of efficient reconstruction of a signal encoded with a differential source code with backward adaptive quantizer, and transmitted over a noisy channel, which has not been previously investigated. The proposed MMSE decoder results in a two stage MMSE decoder reconstructing the quantizer step size and the original signal, respectively. We utilize the proposed MMSE decoder for the continuously variable slope delta modulation (CVSD) codec due to its application in the Bluetooth standard. The results and comparisons to prior art illustrate the effectiveness of the proposed MMSE decoder.

2. SYSTEM MODEL

The block diagram of the system under consideration is presented in Figure 1. We assume that the channel is memoryless without feedback and inter symbol interference (ISI). The block diagram of the adaptive differential encoder with a backward adaptive quantizer is shown in Figure 2.

Figure 1. System model

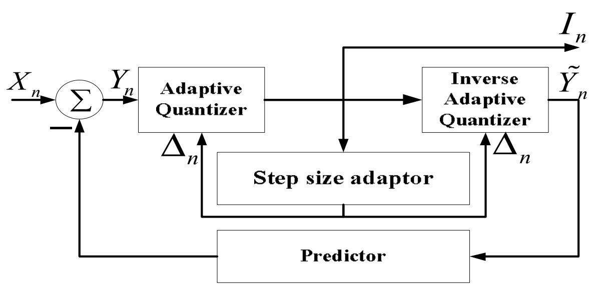

Figure 2. Conventional structure of adaptive differential encoder with backward adaptive quantizer

The notations used are as follows. Capital letters e.g. I, represent random variables and small letters e.g. i represent a random variable realization. $\underline{I}_{n_2}^{n_1}$ represents the sequence of random variables $(I_{n_1}, I_{n_1+1}, ..., I_{n_2})$ and $\underline{I}_n^1$ is equivalently represented by $\underline{I}_n$.

In Figure 2 the difference between encoder input, X_n, and its estimate is the prediction residue, Y_n, that is quantized to $\tilde{Y}_n$. This results in the quantized value of X_n, denoted by $\tilde{X}_n$. In Figure 2, adaptive quantizer outputs index symbol, I_n, based on Y_n and the step size at time instant n, Δ_n. The inverse adaptive quantizer outputs, $\tilde{Y}_n$, based on I_n and Δ_n as follows

$$\tilde{Y}_n = f(I_n, \Delta_n). \quad (1)$$

The model in (1) applies to a number of quantization schemes of interest, such as uniform quantizers and select non-uniform quantizers. In a moving-average (MA) predictor, the prediction function is based on $\tilde{Y}$ and in an autoregressive (AR) predictor $\tilde{X}$ are used for prediction. Alternatively, an autoregressive moving-average (ARMA) predictor uses both sets of data. Similarly, in the step size adaptor block, Δ_n may be adapted based on I and/or previous Δ.

3. PROPOSED MMSE DECODER

We use the MMSE criterion for efficient reconstruction of X_n transmitted over a noisy channel, as follows

$$Min\, E\{(X_n - \hat{X}_n)^2 \mid \underline{J}_{n+\delta}\}, \quad (2)$$

where $\hat{X}_n$ is the reconstructed signal at the receiver and is given by

$$\hat{X}_n = E\{X_n \mid \underline{J}_{n+\delta}\}, \quad (3)$$

$\underline{J}_{n+\delta}$ is the received sequence and $\delta \geq 0$ is the delay allowed in the decoding process.

We consider the class of noise robust algorithms for the adaptation of quantizer step size presented in [8], as follows

$$\Delta_n = \beta.\Delta_{n-1} + w(\underline{I}_{n-a}^{n-b}) \quad \beta < 1. \quad (4)$$

To develop the proposed MMSE decoder, we consider a first order AR predictor. We have

$$X_n = Y_n + A.\tilde{X}_{n-1}, \quad (5)$$

and

$$\tilde{X}_n = \tilde{Y}_n + A.\tilde{X}_{n-1}. \quad (6)$$

This model is particularly attractive since the corresponding solutions will contain the solutions of the cases with an MA predictor, and can be easily extended to the cases with a higher order AR predictor.

As demonstrated below, the proposed solution for calculation of (3) results in a two stage MMSE decoder as shown in Figure 3.

The first MMSE decoder uses the received sequence, $\underline{J}_{n+\delta}$, for estimation of the step size at time instant n-k, and benefits from $\delta+k$ delay in this process.

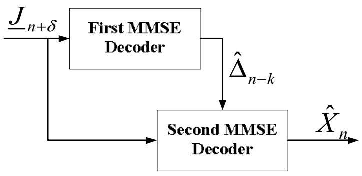

Figure. 3. Structure of the proposed MMSE decoder

The second MMSE decoder reconstructs $\hat{X}_n$ using the received sequence, $\underline{J}_{n+\delta}$, and the estimate of the qunatizer step size, $\hat{\Delta}_{n-k}$, received from the first MMSE decoder. This solution is obtained as elaborated below.

From (5) and a recursive use of equation (6), we have

$$X_n = Y_n + A\tilde{X}_{n-1} = Y_n + \sum_{j=1}^{\mu_1} A^j \tilde{Y}_{n-j} + A^{\mu_1+1}\tilde{X}_{n-\mu_1-1}. \quad (7)$$

From (3) and (7), assuming that μ_1 indicates the effective memory of the differential encoder [3], $\hat{X}_n$ is given by

$$\hat{X}_n = A^{\mu_1+1}.\hat{X}_{n-\mu_1-1} + E\left\{\sum_{j=0}^{\mu_1} A^j.\tilde{Y}_{n-j} \mid \underline{J}_{n+\delta}\right\}, \quad (8)$$

which, corresponds to finalizing the values of the prediction residues beyond effective memory [3]. In fact, increasing the value of parameter μ_1 results in a value for $\hat{X}_n$ that is closer to the optimal.

According to (1), we can write (8) as follows

$$\hat{X}_n = A^{\mu_1+1}.\hat{X}_{n-\mu_1-1} + E\left\{\sum_{j=0}^{\mu_1} f(\Delta_{n-j}, I_{n-j}) \mid \underline{J}_{n+\delta}\right\}. \quad (9)$$

By recursive replacement of Δ_{n-1} in (4), it is straightforward to see that Δ_n can be described as a function of $\underline{I}_{n-a}$, thus, we can compute (9) as follows

$$\hat{X}_n = A^{\mu_1+1}.\hat{X}_{n-\mu_1-1} + \sum_{j=0}^{\mu_1} \sum_{\underline{I}_{n-j}} f(\underline{I}_{n-j}) p(\underline{I}_{n-j} \mid \underline{J}_{n+\delta}). \quad (10)$$

According to (10), increasing the value of n will engage a larger sequence of encoder output symbols (I) in the MMSE decoding. This results in a growing complexity to achieve a value for $\hat{X}_n$ that is closer to the optimal. To achieve a manageable complexity, we propose presenting Δ_{n-j} in (9) using (4) as follows

$$\Delta_{n-j} = \beta^{k-j}.\Delta_{n-k} + \sum_{i=j}^{k-1} \beta^{i-j}.w(\underline{I}_{n-a-i}^{n-b-i}) \quad k \geq j \quad (11)$$

and finalizing the value of Δ_{n-k} to its decoded value $\hat{\Delta}_{n-k}$.

We have

$$\hat{X}_n = A^{\mu_1+1}.\hat{X}_{n-\mu_1-1} + \sum_{j=0}^{\mu_1} \sum_{\underline{I}_{n-j}^{n-b-k+1}} f(\hat{\Delta}_{n-k}, \underline{I}_{n-j}^{n-b-k+1}).p(\underline{I}_{n-j}^{n-b-k+1} \mid \underline{J}_{n+\delta}) \quad (12)$$

in which, $k \geq \mu_1$. Equation (12) indicates a two stage MMSE decoder, where in the first stage Δ_{n-k} is decoded and subsequently the original signal is reconstructed. It is noteworthy that increasing the parameter k results in a value for $\hat{X}_n$ that is closer to the optimal at the cost of increased complexity. Although the algorithm in (12) indicates an approximate MMSE solution, our results validates its effectiveness.

We propose computing the value of $\hat{\Delta}_{n-k}$ in (12) based on the following MMSE decoding rule

$$\hat{\Delta}_{n-k} = E\{\Delta_{n-k} \mid \underline{J}_{n+\delta}\}. \qquad (13)$$

By recursive application of equation (4), we have

$$\Delta_{n-k} = \beta^{\mu_2+1}.\Delta_{n-k-\mu_2-1} + \sum_{i=0}^{\mu_2} \beta^i.w(\underline{I}_{n-k-a-i}^{n-k-b-i}), \qquad (14)$$

and similar to the approach taken in (8), $\hat{\Delta}_{n-k}$ is computed from (13) and (14) as follows

$$\hat{\Delta}_{n-k} = \beta^{\mu_2+1}.\hat{\Delta}_{n-k-\mu_2-1} + \sum_{i=0}^{\mu_2} \sum_{\underline{I}_{n-k-a-i}^{n-k-b-i}} \beta^i.w(\underline{I}_{n-k-a-i}^{n-k-b-i}).p(\underline{I}_{n-k-a-i}^{n-k-b-i} \mid \underline{J}_{n+\delta}) \qquad (15)$$

indicating an effective memory length of μ_2 for the effect of encoder output symbols, I, on the value of Δ.

The a posteriori probabilities, $p(\underline{I}_{n-j}^{n-b-k+1} \mid \underline{J}_{n+\delta})$ and $p(\underline{I}_{n-k-a-i}^{n-k-b-i} \mid \underline{J}_{n+\delta})$ in equation (12) and (15) are calculated as described in the next subsection. The effect of the parameters on the performance is investigated in section 5.

3.1. Computing the a posteriori probabilities

We assume that the output of the source encoder due to the residual redundancy form a γ-order Markov sequence. This may be represented by a trellis structure, where the states at time instant n are defined by

$$S_n = (I_{n-\gamma+1}, I_{n-\gamma+2}, \ldots, I_n), \qquad (16)$$

and we have

$$p(S_n \mid S_{n-1}, S_{n-2}, \ldots, S_1) = p(I_n \mid S_{n-1}, S_{n-2}, \ldots, S_1) = p(I_n \mid S_{n-1}) = p(S_n \mid S_{n-1}) \qquad (17)$$

which indicates a first order Markov sequence for the states. Each branch in the trellis is identified by the pair (S_n, S_{n+1}) indicating the states it connects together and corresponds to one a priori probability given by

$$p(S_{n+1} \mid S_n) = p(I_{n+1} \mid S_n). \qquad (18)$$

The a posteriori probability of state is calculated recursively by the following forward-backward equation as follows

$$p(S_n \mid \underline{J}_{n+\delta}) = \mathrm{C}.p(S_n \mid \underline{J}_n).p(\underline{J}_{n+\delta}^{n+1} \mid S_n) \qquad (19)$$

where C is a factor which normalize the sum of probabilities to one. The term $p(S_n \mid \underline{J}_n)$ is the forward term and is given by

$$p(S_n \mid \underline{J}_n) = \mathrm{C}.p(J_n \mid I_n).\sum_{S_{n-1}} p(S_n \mid S_{n-1})p(S_{n-1} \mid \underline{J}_{n-1}) \qquad (20)$$

The term $p(\underline{J}_{n+\delta}^{n+1} \mid S_n)$ in (19) is the backward term and is calculated recursively by [3]

$$p(\underline{J}_{n+\delta}^{n+1} \mid S_n) = \sum_{I_{n+1}} p(J_{n+1} \mid I_{n+1}).p(I_{n+1} \mid S_n).p(\underline{J}_{n+\delta}^{n+2} \mid S_{n+1}) \qquad (21)$$

and starts from

$$p(\underline{J}_{n+\delta}^{n+\delta} \mid S_{n+\delta-1}) = p(J_{n+\delta} \mid S_{n+\delta-1}) = \sum_{I_{n+\delta}} p(J_{n+\delta} \mid I_{n+\delta}).p(I_{n+\delta} \mid S_{n+\delta-1}) \qquad (22)$$

Then calculate a posteriori probability of $\underline{I}_{n-\eta}^{n-\mu}$ from equation (19)[3].

4. CVSD ENCODER

The CVSD encoder is a delta modulation encoder with an adaptive quantizer [8]. The predictor filter of CVSD is a first order AR predictor, $\tilde{X}_n = \tilde{Y}_n + A.\tilde{X}_{n-1}$. In CVSD a two level quantizer is used. At time instant n, if the prediction residue is greater or less than zero the index symbol, I_n, is respectively set to 1 or -1. In the CVSD encoder the step size Δ_n is updated based on the last three output bits according to the following equation

$$\Delta_n = \beta\Delta_{n-1} + \alpha_n \quad \Delta_{\min} \le \Delta_n \le \Delta_{\max}, \qquad (23)$$

where

$$\alpha_n = \begin{cases} \Delta_0 & \text{if } I_n = I_{n-1} = I_{n-2} \\ 0 & \text{otherwise} \end{cases} \qquad (24)$$

When Y_n is greater than zero $\tilde{Y}_n$ is equal to Δ_n and when Y_n is less than zero $\tilde{Y}_n$ is equal to $-\Delta_n$. Therefore, for the IAQ, we have $\tilde{Y}_n = I_n.\Delta_n = f(\Delta_n, I_n)$.

The parameters $\Delta_{\max}$, $\Delta_{\min}$, β, A, Δ_0 are fixed and are determined according to the dynamic range, bandwidth and sampling frequency of the input signal. For a CVSD input signal, whose amplitude is normalized to one and sampled at 8 kHz, the signal is up sampled to 64 kHz and scaled to 32768, and the parameters of the encoder is set as $\Delta_{\max}$=1280, $\Delta_{\min}$=10, β=1023/1024, A=31/32 and Δ_0=10.

5. SIMULATION RESULTS

We have considered separate training and test speech sets using male and female samples with different accents from the TIMIT database.

For measurement of the residual redundancy in the output of the source encoder we have used $R(\gamma)$ given by

$$R(\gamma) = \log_2 2 - H(I_n \mid S_{n-1}). \qquad (25)$$

Using the training set, $R(\gamma)$ for γ=1 to 7 is computed as 0.475, 0.479, 0.497, 0.518, 0.530, 0.541 and 0.546, respectively. The redundancy exploited by means of a Markov model of order four or higher is greater than 50%.

We consider a memory-less AWGN channel, with BPSK modulation and soft outputs (soft bit decoding (SBD) [2]). In this case, the instantaneous error probability for each received bit is given by

$$p_n^{(e)} = 1/(1+\exp|L_c.J_n|) \quad L_c = 4.\sqrt{E_b}/N_0 \qquad (26)$$

Figure 4 presents the performance of the proposed MMSE decoder, the standard CVSD decoder and the method presented in [2]. These results could be elaborated as follows: (i) As expected, increasing the Markov model order improves the performance of the proposed decoder. In particular, the reconstructed signal SNR is improved by 0.8 dB at E_b/N_0=3, when the Markov order is increased from 3 to 5; (ii) For $\gamma = 5$ increasing the parameters from (μ_1, μ_2, k, δ) = (0, 0, 1, 0) to (3, 2, 4, 3) leads to a gain of 0.96 dB at E_b/N_0=3 dB; (iii) The proposed MMSE decoder (with the latest parameter set) achieves a substantial performance improvement. Specifically, a gain of 5.9 dB and 2.03 dB in the reconstructed signal SNR is achieved at E_b/N_0=4 dB (BER = 1.25e-2), in comparison with the CVSD decoder and the method of [2], respectively. This gain exceeds 13 dB for poor channel conditions compared with the CVSD decoder.

In Table I the effect of delay, δ, parameters μ_2 and μ_1 on the performance of the MMSE decoder is presented. It is seen that for γ = 4 and, (μ_1, μ_2, k, δ) = (0, 0, 1, 0), increasing the parameters enhances the performance. Specifically, at E_b/N_0 = 3 dB, 1.16, 0.11 and 0.35 dB gain in the reconstructed signal SNR is achieved, when parameters δ , μ_2, and μ_1 is increased to 3.

Our simulations and analysis indicated a reasonable complexity for the proposed algorithm. Specifically, although the decoding complexity grows exponentially with Markov model order γ [3], the small decoding complexity is due to two facts: (i) the encoder output is binary, i.e., $I_n \in$ {-1,1}, and (ii) small values of γ provides almost all of the available performance gain. For example, at γ=5, there are only 32 states at each stage of the trellis.

6. CONCLUSION

An efficient MMSE decoder for a differential encoder with an adaptive quantizer is presented. The proposed MMSE decoder is designed to reconstruct efficiently the original input signal of the encoder, leading to a two stage decoding scheme. This decoder benefits from residual redundancy in the output of the source encoder captured by a γ-order Markov model. As confirmed by the results, the presented algorithms provide effective solutions for efficient decoding of *adaptive* differential encoders over noisy channels.

7. ACKNOWLEDGMENT

This work has been supported in part by the Iran Telecommunications Research Center. The authors would like to thank A. K. Khandani for helpful discussions on the topic.

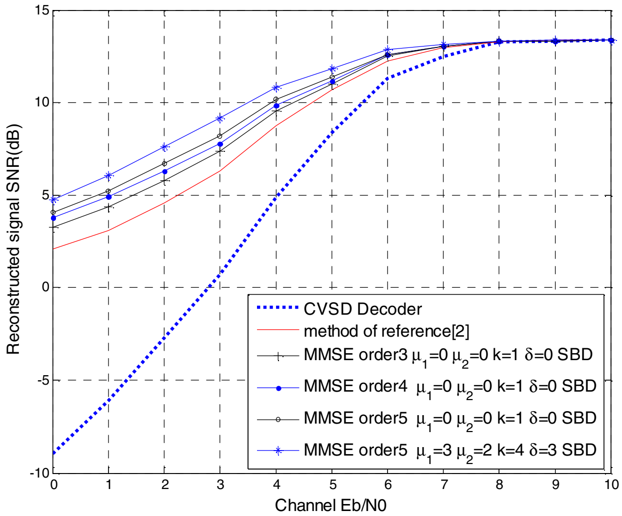

Figure. 4. Performance of CVSD Decoder, method of [2] and proposed MMSE Decoder for order 3, 4 and 5

Table I. The effect of increasing δ , μ_2 and μ_1 in the reconstructed SNR (dB)

$\frac{E_b}{N_0}$ (dB)	SBD γ=4 μ_1=0 μ_2=0 k=1 δ=0	SBD γ=4 μ_1=0 μ_2=0 k=1 δ=3	SBD γ=4 μ_1=0 μ_2=3 k=1 δ=0	SBD γ=4 μ_1=3 μ_2=0 k=4 δ=0
1	4.90	5.84	5.19	5.30
2	6.29	7.33	6.53	6.72
3	7.80	8.96	7.91	8.15
4	9.83	10.58	9.94	10.04
5	11.17	11.82	11.25	11.39

8. REFERENCES

[1] K. Sayood, J. C. Borkenhagen, “Use of residual redundancy in the design of joint source/channel coders,” *IEEE Trans. Commun*, Vol. 39, pp. 838-845, June 1991.

[2] T. Fingscheidt and P. Vary, “Softbit speech decoding: A new approach to error concealment,” *IEEE Trans. Speech and Audio Processing*, Vol. 9, No. 3, pp. 240-251, Mar. 2001.

[3] F. Lahouti and A. K. Khandani, “Reconstruction of predictively encoded signals over noisy channels using a sequence MMSE decoder,” *IEEE Trans. Commun*, Vol. 52, No. 8, pp. 1292-1301, Aug. 2004.

[4] D. J. Miller, E. S. G. Carotti, Y. W. Wang, J. C. De Martin, “Joint source-channel decoding of predictively and nonpredictively encoded sources: A two-stage estimation approach,” *IEEE Trans. Commun*, Vol. 52, No. 9, Sept. 2004.

[5] M. Adrat, J. Spittka, S. Heinen, and P. Vary, “Error concealment by near optimum MMSE estimation of source codec parameters,” *in Proc. IEEE Workshop on speech coding*, pp. 84-86, 2000.

[6] N. Phamdo and N. Farvardin, “Optimal detection of discrete Markov sources over discrete memoryless channels–Application to combined-source channel coding,” *IEEE Trans. Inform. Theory*, vol. 40, pp. 186-193, Jan. 1994.

[7] G. C. Zhu and F. Alajaji, “Joint source-channel turbo decoding for binary markov sources,” *IEEE Trans. wireless commun*, Vol. 5, No. 5, May 2006.

[8] N. S. Jayant, P. Noll, *digital coding of waveforms principle and application to speech and video*, Prentice Hall, New Jersey, 1984.