A low complexity soft decoder for error concealment in the presence of noise and packet loss

Farshad Lahouti, Amir K. Khandani

Abstract— **Exploiting the residual redundancy in a source coder output stream during the decoding process has been proven to be a bandwidth efficient way to combat the noisy channel degradations. In this paper, several schemes are presented which exploit different levels of residual redundancy for improved signal reconstruction in presence of noise and/or packet loss. A packetization strategy, based on the concepts of multiple description coding, is proposed which is matched to the presented error concealment units. For decoders that exploit the residual redundancy, extensive complexity has been a serious concern, especially as the quantizer bitrate increases. In this work, a method is presented to construct reduced complexity algorithms. The proposed methodology is based on the classification of the signal domain and efficient approximation of the residual redundancy or the a priori transition probabilities. We employ the proposed schemes for soft reconstruction of speech spectrum, in GSM Adaptive Multi-Rate vocoder, transmitted over a channel disturbed with noise and/or packet loss. Numerical results are provided which demonstrate the effectiveness of the proposed solutions.**

I. Introduction

Joint source channel (JSC) coding techniques that exploit the residual redundancy in the source coder output stream have found increasing attention in recent years, e.g., [1]-[7]. One of the reasons for such interest is the fact that, they provide improved signal protection against channel errors, with no additional bandwidth requirement, using only the redundancy left due to suboptimal source coding. For design of source decoders, such JSC-based techniques replace the conventionally heuristic approaches of error concealment with a formulation within the formal framework of estimation theory. There are two challenges, however, in using these techniques to combat the effect of noise and packet loss in typical communications applications. First, although they provide effective solutions, but in general, the decoding complexity is rather high for practical applications. Specifically, the complexity grows exponentially with quantizer bitrate. Secondly, the focus of the prior art in this area have so far been on the design of source decoders to combat bit-level channel degradations. In some important applications, however, the transmitted signal is exposed to both noise and packet loss. Examples of particular interest include the communications over wireless packet networks; and the scenario when the communication involves both a wireless and a (packet-based) wireline link. Solutions to these two challenges are proposed in this article. A more comprehensive version of this work is available in [7]. An earlier version appeared in [6].

The authors are affiliated with the Coding & Signal Transmission Lab., Dept. of E&CE, University of Waterloo, Waterloo, ON, N2L 3G1, Canada, Email: (farshad, khandani)@cst.uwaterloo.ca. Tel.: (519) 888-4567 Ext. 5324

II. MMSE-based Source Decoding

Consider the scenario of a source coder composed of η vector quantizers. At time instant n, corresponding to the input samples $[\mathbf{Y}_n^{(1)}, \mathbf{Y}_n^{(2)}, \ldots, \mathbf{Y}_n^{(\eta)}]$, the source coder produces η symbols $[I_1^{(1)}, I_2^{(3)}, \ldots, I_n^{(\eta)}]$, where $I_n^{(k)} \in \mathcal{J}^{(k)}$ and $\mathcal{J}^{(k)}$ is a set of $M^{(k)}$ codewords of quantizer k, $0 < k \leq \eta$. In general, there is a dependency (residual redundancy) both in the sequence of source coder symbols in time (interframe) and between the symbols at each time instant or frame (intraframe). For the input symbol I_n, the channel output is denoted by J_n. We assume two types of channel degradations (i) a random independent loss at the packet level with probability $Packet\,Loss\,Rate(PLR)$ and (ii) a random memoryless bit-level channel disturbance. In the followings, for the development of the source decoders, we assume that the probability distribution of $P(J_n|I_n)$ is available. If a bit is located in a packet that is lost during transmission, then this probability is equal to 0.5; otherwise it reflects the effect of channel degradations at the bit level.

Exploiting the encoder intraframe and interframe dependencies using a first-order Markov model, the decoder MS1 is given by

$$\hat{\mathbf{y}}_n^{(k)} = \sum_{I_n^{(k)} \in \mathcal{J}^{(k)}} E[\mathbf{Y}_n^{(k)}|I_n^{(k)}] P(I_n^{(k)}|\underline{\mathbf{J}}_{n+\delta}),\ 0 < k \leq \eta, \quad (1)$$

where $\underline{\mathbf{J}}_{n+\delta} = [\underline{J}_{n+\delta}^{(1)}, \ldots, \underline{J}_{n+\delta}^{(\eta)}]$, $\underline{J}_{n+\delta}^{(k)} = [J_1^{(k)}, \ldots, J_{n+\delta}^{(k)}]$, $E[\mathbf{Y}|I]$ represents the codewords, and δ is the delay allowed in the decoding process. The symbol a posteriori probability $P(I_n|\underline{\mathbf{J}}_{n+\delta})$ in equation (1) is given by

$$P(I_n^{(k)}|\underline{\mathbf{J}}_{n+\delta}) \approx C.P_{fwd}(I_n^{(k)}) \,.\, P_{bwd}(I_n^{(k)}) \,.\, P_{rgt}(I_n^{(k)}) \,.\, P_{lft}(I_n^{(k)}), \quad (2)$$

which is composed of four terms, namely the forward and backward recursive terms, and the left and right recursive terms. These recursions, which are similar to those of the BCJR algorithm, both in time and within a frame, rely on the channel-related probabilities and the a priori information in the form of encoder symbol transition probabilities. In this manuscript, C is a term that normalizes the probabilities to one. We have

$$P_{fwd}(I_n^{(k)}) = P(I_n^{(k)}|\underline{J}_n^{(k)}) = C.P(J_n^{(k)}|I_n^{(k)}) \,. \sum_{I_{n-1}^{(k)} \in \mathcal{J}^{(k)}} P(I_n^{(k)}|I_{n-1}^{(k)}) P_{fwd}(I_{n-1}^{(k)}) \quad (3)$$

$$P_{bwd}(I_n^{(k)}) = P(\underline{J}_{n+\delta}^{n+1\,(k)}|I_n^{(k)}) = \sum_{I_{n+1}^{(k)} \in \mathcal{J}^{(k)}} P(J_{n+1}^{(k)}|I_{n+1}^{(k)}) . P(I_{n+1}^{(k)}|I_n^{(k)}) P_{bwd}(I_{n+1}^{(k)}) \quad (4)$$

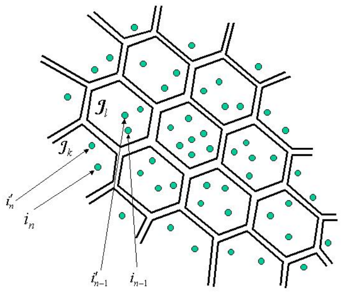

Fig. 1. Classification of the source codewords. See eq. (6) and (7).

The backward recursion starts from $P_{bwd}(I^{(k)}_{n+\delta-1}) = P(\underline{J}^{(k)}_{n+\delta}|I^{(k)}_{n+\delta-1}) = \sum_{I^{(k)}_{n+\delta}\in\mathcal{J}^{(k)}} P(J^{(k)}_{n+\delta}|I^{(k)}_{n+\delta})P(I^{(k)}_{n+\delta}|I^{(k)}_{n+\delta-1})$, and continues backward until $P_{bwd}(I_n)$ is derived. The left and right recursive terms are computed similarly as presented in [5] and enhanced in [7].

We also consider a simplified decoder MS2[6], that only considers the dominant intraframe dependency of symbol k with another symbol k' at the same time instant. The probability required in equation (1) is now given by

$$P(I_n^{(k)}|\underline{\mathbf{J}}_{n+\delta}) \approx P(I_n^{(k)}|\underline{J}^{(k)}_{n+\delta}, J_n^{(k')}) \approx C.P_{fwd}(I_n^{(k)})\,.\,P_{bwd}(I_n^{(k)})\,. \sum_{I_n^{(k')}\in\mathcal{J}^{(k')}} P(J_n^{(k')}|I_n^{(k')})P(I_n^{(k')}|I_n^{(k)}). \quad (5)$$

III. A Low Complexity Solution

As discussed, the MMSE-based decoders, although effective, can be very complex especially at high quantizer bitrates. To provide a solution with a lower complexity, we seek an approximation of the a priori transition probabilities. Consider a vector quantizer with a codebook $\mathcal{C}$ and the corresponding index set $\mathcal{J}$ of M elements. Assume that due to the residual redundancy, the sequence of quantizer output in time forms a first-order Markov model. To derive an approximation to the a priori transition probabilities $P(I_n|I_{n-1})$, we classify the source (and hence the codebook and the index set) to $M', M' \leq M$ classes, i.e., $\mathcal{J} = \{\mathcal{J}_1, \mathcal{J}_2, \ldots, \mathcal{J}_{M'}\}$, in which each class $\mathcal{J}_k$ has m_k members ($\sum_{k=1}^{M'} m_k = M$). We assume that the classification is performed in a way that the probability of transition from a codeword to another codeword only depends on the class they are located in, i.e.,

$$P(I_n = i_n|I_{n-1} = i_{n-1}) = P(I_n = i'_n|I_{n-1} = i'_{n-1}) = P_{kl}, \quad \forall i_n, i'_n \in \mathcal{J}_k,\ \forall i_{n-1}, i'_{n-1} \in \mathcal{J}_l. \quad (6)$$

The relationship between the class and symbol transition probabilities is then given by the following simple derivations:

$$\begin{aligned} P(I_n \in \mathcal{J}_k|I_{n-1} \in \mathcal{J}_l) &= \sum_{I_n\in\mathcal{J}_k} P(I_n|I_{n-1} \in \mathcal{J}_l) \\ &= \sum_{I_n\in\mathcal{J}_k}\sum_{I_{n-1}\in\mathcal{J}_l} P(I_n|I_{n-1})P(I_{n-1}|I_{n-1} \in \mathcal{J}_l) \\ &= m_k\,.\,P(I_n|I_{n-1}), \end{aligned} \quad (7)$$

This indicates that the Markov model of the encoder output sequence is characterized based on the $M' \times M'$ class transition probabilities. This in turn needs much less data and is stored much more efficiently reducing the memory requirement by a factor of $(\frac{M}{M'})^2$. Figure (1) is a representative diagram of such a scenario. The intraframe transition probabilities can be also approximated in a similar fashion.

Using this technique, the presented decoders can be implemented in a much more computationally efficient manner. We have

$$P_{fwd}(I_n) = C.P(J_n|I_n). \sum_{k=1}^{M'} P(I_n|I_{n-1} \in \mathcal{J}_k) \sum_{I_{n-1}\in\mathcal{J}_k} P_{fwd}(I_{n-1}), \quad (8)$$

$$P_{bwd}(I_n) = \sum_{k=1}^{M'} P(I_{n+1}|I_n)P_{bwd}(I_{n+1}). \sum_{I_{n+1}\in\mathcal{J}_k} P(J_{n+1}|I_{n+1}). \quad (9)$$

The right and left recursive terms are given by similar approximate expressions [7]. In the set up of decoder MS1, assuming that the codebook $\mathcal{J}^{(k)}, 0 < k \leq \eta$ is partitioned to $M'^{(k)}$ classes, the computational complexity is now given by

$$CC'_{fwd} = \sum_{k=1}^{\eta} 3M^{(k)} + 2{M'^{(k)}}^2 \quad (10)$$

$$CC'_{bwd} = 3\delta \sum_{k=1}^{\eta} {M'^{(k)}}^2 \quad (11)$$

$$CC'_{lft} = CC'_{rgt} = 3\sum_{k=1}^{\eta-1} M'^{(k+1)}M'^{(k)} \quad (12)$$

in terms of the required number of floating point operations. It is clear that any choice of $M' < M$ substantially reduces the complexity.

In order to classify the codebook in a way that the equation (6) holds, we suggest LBG quantization of the source with M' levels and defining the classes as the quantization Voronoi regions. Subsequently, we can classify the codewords of the size M codebook. This is motivated by the fact that the closely positioned codewords tend to behave similarly. We note that this does not guarantee the validity of the assumption of equation (6) and it is only an approximate technique. However, our numerical results demonstrate its fruitfulness.

IV. Packetization Strategy

Interleaving during packetization, as a Multiple Description Coding strategy, creates robustness to packet loss

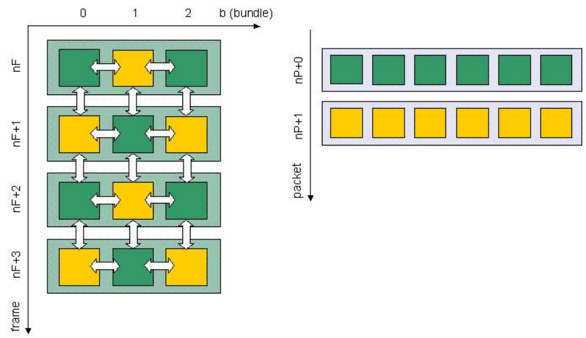

Fig. 2. Packetization for $F = 4$, $B = 3$ and $P = 2$. The arrows represent the dependencies due to residual redundancies.

through proper diversification (distribution) of the data across different packets (descriptions). In this work, we propose a packetization strategy that is matched to the presented source decoders exploiting both interframe and intraframe dependencies. Consider the scenario where the source coder outputs certain number of bits at each time instant (frame). Let us assume that these bits may be partitioned into B bundles of data where in general, there are dependencies between the (neighboring) data bundles at each time instant, i.e., within a frame, and over time, i.e., between frames. The proposed packetization rule takes into account these dependencies and is given by

$$b + f + nF \equiv p + nP \pmod P, \tag{13}$$

which distributes a set of F frames of data, each composed of B bundles, across P packets (P indicates the diversity depth). In equation (13), we have $n, b, B, f, F, p, P \in \mathbb{Z}$, where $0 \le b < B$ is the bundle number, $f + nF$, $0 \le f < F$, indicates the frame number, $p + nP$, $0 \le p < P$, indicates the packet number and $(\mathrm{mod}\, P)$ is the modulo operation in base P. Also, $n \ge 0$ is a notion of time indicating the set of frames to be interleaved. Figure 2 illustrates the case for $B = 3$, $F = 4$ and $P = 2$. As seen, if one of the packets is lost, the other packet contains all the neighboring bundles of the missing data which are utilized by the source decoder to reconstruct the missing bundles. For the case of $B = 1$, equation (13) is simplified to the popular frame interleaving scheme.

V. Performance Evaluation

Using the proposed techniques, we consider reconstruction of speech encoded with the GSM-AMR codec [8] and transmitted over a channel disturbed with noise and packet loss. We focus on the reconstruction of the Linear Predictive Coding (LPC) coefficients. These coefficients represent the short-term spectral information of speech within a frame and preserving them play a major role in the quality of the reconstructed speech. In this codec, the LPC coefficients are quantized in the Line Spectral Frequency (LSF) representation using a 3-split Split-VQ with $[8, 9, 9]$ bits for an overall rate of 26 bits per frame. A similar scheme is also used in the IS-641 standard. We use the packetization scheme of Figure 2 with the bundles now corresponding to the three splits of the Split-VQ. Taking the low complexity approach, with a class size of $M' = 32$ for each of the splits ($M^{(1)} = 256, M^{(2)} = M^{(3)} = 512$), results in a substantial reduction in memory requirement, as compared to the original algorithm; also the computational complexity is reduced by an impressive factor of 118 for the forward recursion and a factor of 192 for the backward, left and right recursive terms. Table I presents the performance gain of the proposed decoders with respect to the error concealment method of the standard. The gain is given in terms of improvement of signal to reconstruction noise ratio for each of the reconstructed LSF parameters. The results demonstrate the effectiveness of the proposed schemes. The simplified MS2 decoder provides an interesting performance with much lower complexity.

Method	LSF1	LSF2	LSF3	LSF4	LSF5	
MS1	3.58	4.52	4.95	4.63	4.30	
MS2	3.44	4.39	4.68	4.38	3.76	
Method	LSF6	LSF7	LSF8	LSF9	LSF10	Mean
MS1	4.80	4.03	4.01	3.33	3.30	4.14
MS2	3.91	3.84	3.80	3.18	3.00	3.84

TABLE I

Gain (dB) in signal to reconstruction noise ratio of the proposed error concealment schemes as compared to the scheme used in GSM-AMR codec, in presence of noise (channel SNR=2dB, BER=0.037) and packet loss (PLR=5%).

References

[1] K. Sayood, J. C. Borkenhagen, "Use of residual redundancy in the design of joint source/channel coders," *IEEE Trans. Commun.,* vol.39, No.6, pp. 838-845, 1991.

[2] N. Phamdo and N. Farvardin, "Optimal detection of discrete Markov sources over discrete memoryless channels- Applications to combined-source channel coding," *IEEE Trans. Inform. Theory,* vol. 40, pp. 186-103, 1994.

[3] D. J. Miller and M. Park, "A sequence-based approximate MMSE decoder for source coding over noisy channels using discrete hidden Markov models,"*IEEE Trans. Commun.,* vol.46, No.2, pp. 222-231, 1998.

[4] F. I. Alajaji, N. Phamdo and T. E. Fuja, "Channel codes that exploit the residual redundancy in CELP-encoded speech," *IEEE Trans. Speech and Audio Proc.,* vol. 4, No. 5, Sept. 1996.

[5] T. Fingscheidt, T. Hindelang, R. V. Cox, and N. Seshadri, "Joint source-channel (de-)coding for mobile communications" *IEEE Trans. Commun.,* vol. 50, No. 2, pp. 200-212, 2002.

[6] F. Lahouti and A. K. Khandani, "Approximating and exploiting the residual redundancies- Applications to efficient reconstruction of speech over noisy channels," *Proc. IEEE Int. Confs. Acoust., Speech and Sig. Proc.,* vol.2, pp.721-724, Salt Lake City, UT, May 2001.

[7] F. Lahouti and A. K. Khandani, "Soft reconstruction of speech in the presence of noise and packet loss," Technical Report 2003-4, Department of E&CE, University of Waterloo, May 2003, available at www.cst.uwaterloo.ca.

[8] "Adaptive Multi-Rate (AMR) speech codec; Transcoding functions," 3GPP TS 26.090 V5.0.0, June 2002.

[9] K. K. Paliwal and B.S. Atal, "Efficient vector quantization of LPC parameters at 24 bits/frame," *IEEE Trans. Speech and Audio Process.*, vol.1, no.1, pp. 3-14, 1993.