Missing Packet Recovery Techniques for DM Coded Speech

Wen-Whei Chang\textsuperscript{1}, Hwai-Tsu Chang\textsuperscript{2}, Wan-Yu Meng\textsuperscript{2}
\textsuperscript{1}Department of Communication Engineering
National Chiao Tung University, Hsinchu, Taiwan, ROC
\textsuperscript{2}Advance Technology Center, Computer and Communication Research Laboratories,
Industrial Technology Research Institute, Hsinchu, Taiwan, ROC

Abstract

The packet loss effects of DM coded speech can be mitigated by either using an embedded DM system (EDM) or using a tree search interpolator. This paper provides theoretical and experimental results for EDM coding of autoregressive sources under random error conditions. For the tree interpolation, we explore the benefits of delayed decoding by using an interpolative DM code generator to form a tree of sample possibilities given the remaining adjacent samples.

I. Introduction

Packet communication increases the utilization efficiency of channel capacity by exploiting the bursty nature of voice traffic. Depending upon the contents of speech packets, numerous recovery techniques have been proposed to combat the packet loss due to channel impairments or excessive delay [1]. Most previous research has emphasized the missing packet recovery techniques for PCM and DPCM coded speech [2,3]. When using a DM encoder, however, the missing samples in the discarded packets cannot be reconstructed by zero-amplitude stuffing or waveform substitution techniques. This is because mistracking operations between the encoder and the decoder result in large distortion. An odd-even sample interpolation [2] of the weakly correlated DM residual signal also leads to unsatisfactory results. This paper addresses this issue by comparing embedded DM [4] to two interpolative recovery techniques. We first analyze the effect of transmission errors for EDM coding of autoregressive sources, and we substantiate the theoretical values with simulation results. We have also studied the benefits of a new speech packet interpolation method that has some similarity to the work in [3], but rather than being based on the assumption that the autoregressive model is known, the new structure utilizes the code tree generated by an interpolative DM system.

II. Embedded DM

The embedded DM scheme, which is designed for variable rate coding [4], can be applied to mitigate the degradation caused by packet loss. The reason seems to be that EDM with simple bit dropping method resolves the mistracking problem between the encoder and the decoder when adaptive parameter control information is discarded. As shown in Fig. 1, we note that EDM encoder contains a minimal DM encoder and a supplemental PCM encoder that creates the $R_{wp}$-bit digital representation of DM error signal. The decoder adds the DM and PCM coded signals and filters the sum to reject its out-of-band distortion components. Further enhancement can be achieved if DM error signal $e(n)$ is filtered and decimated to obtain $e_{id}(n)$ prior to processing by the supplemental PCM encoder. Corresponding filtering and decimation of the decoded DM signal in the receiver are then required. To operate the system at 32kb/s, each EDM codeword is equally split into two segments, one with $R_{min}$=2 MSBs (most significant bits) and the other with $R_{wp}$=2 LSBs (least significant bits) per Nyquist interval. In our analysis, we refer to three signals: $\{u\}$, the quantizer input, $\{u_1\}$, the $D$-bit representation of $\{u\}$, and $\{u_0\}$, the received version of $\{u\}$ affected by the binary error pattern $l$. Note that the cascade of 1-bit and $R_{wp}$-bit quantizers leads to the same quantization noise as that of $D=3$. The signals in the decoder are written with primes to distinguish them from those in the encoder.

When subject to packet loss, errors in the PCM bit streams are amplified by the decoder integrator, which has no effect on errors in the DM bit streams. To gain further insight, we provide here an approximate analysis for the combined effects of quantization and transmission errors on the mean-squared-distortion performance of
EDM. Suppose that errors in all bits are statistically independent and occur with probability $P$. Thus for the $D$-bit error pattern $l$, its Hamming weight $w$ and the probability $P(l) = P(l-l^D)^w$. With $p_l$ the probability that $u$ is quantized to $u_l$, we can derive the corresponding average quantization error $q_l$ and the error variance $\sigma^2_q$. In our analysis, we begin by viewing the EDM as a simplified form of embedded DPCM, and then take into account the complication due to the lowpass filtering and decimation. Proceeding in this way, the variance of the filtered, decimated error signal between $u$ and $u_l$ can be written in the form of $\left(\sigma^2_u\right)_d = \left(\sigma^2_q + \sigma^2_i\right)_d$. The effects of transmission errors on the EDM performance are summarized in the second term $\sigma^2_i$, which is given by

$$\sigma^2_i = \frac{a_1^2}{1-a_1^2} P \sum_{i=1}^{2^{p-1}} P(i) \left( \sum_{i=0}^{2^{p-1}} p_i (u_i - u_l)^2 + 2 \sum_{i=0}^{2^{p-1}} q_i (u_i - u_l) \right)$$

The next step of the present investigation is concerned with the out-of-band noise suppression due to the lowpass decimating filter. To a first approximation, we assume that the minimal DM coder operates in its granular mode and has unity quantizer overload load, which implies that DM error signal $e(n)$ has a uniform amplitude distribution over $[-1/2, 1/2]$. With $R_{\text{intra}}$-bit uniform quantization, the PCM step size is $\Delta = r_0 / 2^{R_{\text{intra}}-1}$ where the overload point $r_0$ can be expressed in terms of a loading factor $\lambda = \sqrt{2R_{\text{min}}}$ such that $r_0 = \lambda R_{\text{intra}}$. Due to the out-of-band noise suppression, the variance of $e_p(n)$ is lower than that of $e(n)$ in the form of $\sigma^2_{e_p} = \sigma^2_{e_p} / bR_{\text{intra}}$ with $b$ equal to 2 in granular-noise mode. Finally, the variance of the filtered, decimated quantization noise is given by

$$\sigma^2_{q,d} = \frac{\Delta^2}{12} = \frac{2^{2\beta+3}R_{\text{intra}}}{36bR_{\text{intra}}}$$

Similarly, we can derive the filtered, decimated version of $\sigma^2_q$ as follows

$$\sigma^2_{q,d} = \frac{a_1^2}{1-a_1^2} P + R_{\text{min}}^{-1} \sum_{i=1}^{2^{p-1}} P(i) \left( \sum_{i=0}^{2^{p-1}} p_i (u_i - u_l)^2 + 2 \sum_{i=0}^{2^{p-1}} q_i (u_i - u_l) \right)$$

Fig. 2 presents the theoretical and experimental results for EDM encoding of first-order Gaussian autoregressive $(AR(1))$ sources at the rate of 32 kb/s. The encoding performance is evaluated over a wide range of random error conditions in terms of the signal-to-noise ratio (SNR). In this experiment, an $AR(1)$ source $x(n)$ is generated by passing a zero mean, variance $\lambda^2$, white Gaussian excitation sequence $d(n)$ through a coloring filter with transfer function $1/(1-\alpha z^{-1})$. We use 5000 samples per experiment, and each experiment is repeated 20 times, each time using an independent realization of $d(n)$. From the observation of Fig. 2, it is evident that the theoretical and experimental values of SNR performance of EDM agree very closely.

III. IDM-Based Tree Interpolator

Classical DM systems operate without delay in the sense that for an input sample at time instant $n$, only data at times $\leq n$ are used in the coding process. Since slight delays are not critical to the communication systems, we propose the incorporation of delay at the decoder, leaving the encoder unchanged, and investigate the performance improvement available using tree search interpolation. In this experiment, the ADM codewords are loaded alternatively with odd numbered samples in one packet and even numbered samples in another packet. This allows for the missing samples to be reconstructed from adjacent samples by means of linear interpolation. The block diagram of a tree search interpolator is illustrated in Fig. 3. In this experiment, we implement an IDM code generator which calculates the output $z(n)$ as a weighted linear interpolation of local DM outputs $y(n)$ and available previously interpolated values. The basic equations describing the IDM code generator operation are

$$y(n+1) = a_1 y(n) + v(n+1)$$

$$z(n) = b_1 y(n+1) + b_2 z(n-1)$$

where the interpolator coefficients $[6]$ are calculated from the single-lag autocorrelation value of $y(n)$ as follows

$$b_{-1} = b_1 = R_y(1) / (1 + R_y(1))$$

The procedure for finding the reference signal is based on IDM decoding of the signal $v(n)$, which are calculated from an odd-even sample interpolation [2] of
the quantized DM residual signal \( u(n) \). When one of two consecutive packets is discarded, a code tree of sample possibilities is formed by feeding the path map candidates to the code generator input for reconstruction. Next, the squared error distortion between the reference sequence and each possible reconstructed sequence to some depth \( L \) in the tree is calculated. For given initial conditions, the \((M,L)\) search algorithm finds the \( M \) best paths to depth \( L \) and discards all of the other paths to this depth. Only the first sample in the best path is released, and the \( M \) or fewer paths with this root node are again extended to depth \( L \), and the process is repeated.

IV. Performance Comparisons

Extensive simulations were conducted to establish the performance improvement attainable with the IDM-based tree search interpolator. Though better performance can always be obtained by increasing the code tree, the \((M,L)\) search algorithm with depth \( L=4 \) and \( M=4 \) is empirically determined as the best compromise between coding gain and implementational complexity. Fig. 4 presents the SNR performance results for 32kb/s ADM coding of speech in conjunction with a tree search interpolator, an odd-even sample interpolator and an embedded DM system with possibility of high order packet loss. The packet length is 256 samples and the missing packet ratios varied from 0 to 8 percent. The speech database for these studies consisted of male and female utterances each of 6 second duration and sampled at 16kHz. The application of tree search interpolator to ADM coded speech shows significant improvement, especially at higher packet loss probabilities. Informal listening testes also indicate that packet loss ratios of up to 20 percent are tolerable with this approach.

V. Conclusion

This paper has discussed various compensation techniques for DM coded speech. We first emphasize the effects of random transmission errors on the performance of EDM, and we substantiate the theoretical values with simulation results for EDM coding of autoregressive sources. We also explore the benefits of tree search interpolation based on an IDM code generator. Experimental results indicate that at high packet loss ratios the tree interpolation yields the best performance, as compared with using EDM or odd-even sample interpolation.

References


Fig. 1 Reduced Rate Embedded Delta Modulation

Fig. 3 Functional Diagram of Tree Interpolation.

Fig. 2 Theoretical and experimental results at EDM 32kb/s.

Fig. 4 SNR Using EDM, Sample Interpolation and Tree Interpolation.