SOFT DECISION UNEQUAL ERROR PROTECTION SCHEME FOR MPEG ADVANCED AUDIO CODING

T. H. Yeo (*), W. C. Wong (**) and D. Y. Huang (**),

(*) Department of Electrical Engineering, National University of Singapore, Singapore
(**) Institute for Infocomm Research, National University of Singapore, Singapore

ABSTRACT

In this paper, a system using soft decision decoding of convolutional codes is investigated for MPEG-4 Advanced Audio Coding. Under severe channel random bit error rates of $4.00 \times 10^{-2}$ and above, the proposed scheme is able to achieve above 90% improvement in residual bit error rate performance over the current scheme. The average percentage overhead incurred using the proposed scheme is about 21%, which is the same as that of the original UEP scheme. An analysis with the theoretical bounds shows that within experimental error, the proposed scheme can perform to near optimum as compared to the current scheme.

1. INTRODUCTION

In recent years, rapid advances in technology have resulted in a rapid growth in the field of wireless communications. With this explosive growth, the need for reliable transmission of high fidelity digital audio data over wireless links is becoming an increasingly important application requirement. In view of this, efficient error protection schemes are introduced into the original bit data to protect it against channel errors during transmission.

Through the years, MPEG audio codecs have been widely adopted for both storage and transmission of digital audio data. One requirement for MPEG-4 audio version 2 was to increase the robustness of the audio bitstreams transmitted over error prone channels. Hence, error protection tools, which involve the use of error control strategies and error resilience tools, which comprise of Huffman codeword reordering, virtual codebooks and variable length code, are used.

For Advanced Audio Coding (AAC), the MPEG audio codec designed to provide the best audio quality, an unequal error protection (UEP) scheme has been provided by NTT DoCoMo and tested with AAC during the standardization process. However, it fails to demonstrate high performance at high bit error rates.

In this paper, a system using soft decision decoding of convolutional codes is proposed to improve the performance of UEP scheme with the same amount of redundancy added. Section 2 details the implementation of UEP scheme using the Error Protection Tool (EPTOOL) adopted by NTT DoCoMo in the process of error protection over noisy channels. Section 3 investigates the use of soft decision convolutional codes to improve the error protection of EPTOOL. The theoretical bounds for different classes are presented in Section 4. Simulation results on the performance of the EPTOOL and the improved EPTOOL are presented and compared with theoretical bounds in Section 5. Section 6 concludes the paper.

2. IMPLEMENTATION OF UEP SCHEME

For most of the audio codecs, error robustness capabilities can be improved through a specific bitstream format. It assigns the bitstream elements to the different “Error Sensitivity Categories” based on their error sensitivity. This bitstream format enables unequal error protection available supported by MPEG-4 error protection tool.

2.1. UEP

UEP is one of the most effective and efficient methods to improve the error robustness of source coding. It is widely used for various audio/speech codecs in error-prone channels such as mobile networks. The main idea of UEP is that it classifies a bit stream to be transmitted into several groups according to its error sensitivity and encodes them with different strength of forward error correction (FEC) codes. The bits, which are extremely crucial in the correct decoding of the source coded data, are classified to be more error sensitive. Thus, these bits are given a higher level of error protection by adding more parity bits. The advantage of such a scheme is that it is able to minimize the amount of redundancy added to the bitstream. If UEP is not used, the decoded audio quality is determined by how the most error sensitive part is corrupted and thus the strongest FEC codes has to be
applied to the whole frame, thus requiring more redundancy.

2.2. EPTOOL

The EPTOOL is the standard proposed by NTT DoCoMo to provide unequal error protection capability to MPEG audio codecs. The AAC format of output bitstream will be used for evaluation purposes. In the AAC format, each audio frame consists of 3 classes, where each class consists of bits of varying error sensitivity. Class 1 consists of bits that are the most error sensitive while class 3 bits are the least error sensitive shown in Figure 1.

<table>
<thead>
<tr>
<th>Header</th>
<th>Class 1 bits</th>
<th>Class 2 bits</th>
<th>Class 3 bits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Least error sensitive</td>
<td>Most error sensitive</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 1: Structure of UEP frame

As decoding a header in error leads to the lost of the whole audio frame, a stronger code is required. The header uses a rate compatible punctured convolutional (RCPC) code of rate ¼ with a puncturing rate of 8/16 + 4-bit Cyclic Redundancy Check (CRC) code and with hard decision decoding to protect it. Table 1 summaries the recommended puncture rate and CRC length to protect the bits from class 1, class 2 and class 3 respectively by NTT DoCoMo.

Table 1: Recommended puncture rate and CRC length for AAC parameters

<table>
<thead>
<tr>
<th>Class</th>
<th>Puncture Rate</th>
<th>CRC Length</th>
</tr>
</thead>
<tbody>
<tr>
<td>Header</td>
<td>8/16</td>
<td>4</td>
</tr>
<tr>
<td>1</td>
<td>8/16</td>
<td>12</td>
</tr>
<tr>
<td>2</td>
<td>8/12</td>
<td>9</td>
</tr>
<tr>
<td>3</td>
<td>8/8</td>
<td>6</td>
</tr>
</tbody>
</table>

The encoding process takes on various stages. First of all, encoding by CRC is carried out. CRC, which has error detection capability, appends parity bits to the audio data before encoding by convolutional code is carried out. In order to reduce the amount of redundancy transmitted over the channel, the output of the convolutional encoder is punctured to allow for different code rates. The process of puncturing involves periodically deleting selected bits at the output of the convolutional encoder.

In the decoder, the header is de-punctured and decoded to determine the lengths of the classes 1, 2 and 3. Next, hard decision convolutional decoding using Viterbi algorithm is carried out, followed by the deletion of parity bits using syndrome decoding, to give the AAC stream, arranged in accordance to the classes.

3. IMPROVEMENT TO EPTOOL

At high bit error rates (BER), the original scheme fails to demonstrate high performance. This is due to the fact that the header is often decoded in error and this leads to a lost of the entire frame. This section discusses how soft decision decoding of convolutional codes are employed to improve error protection of the frame, including the header, thus pushing the limits of noisy channel conditions.

The improvement made is to implement a soft decision decoding system to the UEP structure. The output from the channel is no longer quantized to 2 levels. Unquantized decision decoding is performed to achieve optimum performance. For soft decision decoding of convolutional codes, the Viterbi algorithm operates the same as before. However, the metric is now squared Euclidean distance received codeword with each branch codeword,

\[ d(y_i - c_i) = \sum_{j=1}^{n} (r_{i,j} - c_{i,j})^2 \]  (1)

as compared to the metric based on Hamming distance in the original UEP scheme.

4. THEORETICAL BOUNDS

This section explores the theoretical performance of hard decision and soft decision decoding of convolutional codes. Theoretical bounds of the different punctured code rates of the sub-frames are also highlighted so as to determine the optimum performance of the header and the individual classes.

For coherent binary phase shift keying (BPSK) signals with additive white gaussian noise (AWGN) channels and two level quantization of received signals, the bit error probability per decoded information bit for hard-decision decoding is bounded by

\[ P_b < \frac{dT(D,N)}{dN} \]  (2)

where \( \rho \) is the channel transition probability, \( N \) is the path length, \( D \) is the Hamming weight of the encoder output for that branch and \( T(D,N) \) is the transfer function. The transfer function is derived from the state diagram of a convolutional code. The bit error probability per decoded information bit for soft-decision decoding is

\[ P_b < \frac{1}{2} \frac{dT(D,N)}{dN} \]  (3)

where \( E_b/N_0 \) is the average ratio of bit energy to noise power spectral density. Table 2 shows the bit error probability of the various classes using hard and soft
decision decoding of convolutional codes - a rate ¼ systematic recursive convolutional punctured code, as proposed by NTT DoCoMo.

Table 2: Bit Error Probability of the sub-frames

<table>
<thead>
<tr>
<th>Hard Decision</th>
<th>Soft Decision</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \frac{D^7}{(1 - ND)^2} )</td>
<td>( \frac{D^7}{(I - ND)^2} )</td>
</tr>
<tr>
<td>( \frac{D^4}{(1 - ND)^2} )</td>
<td>( \frac{D^4}{(2 - I - ND)^2} )</td>
</tr>
<tr>
<td>( \frac{1}{2} \text{erfc}(\frac{E_b}{\sqrt{N_o}}) )</td>
<td>( \frac{1}{2} \text{erfc}(\frac{E_b}{\sqrt{N_o}}) )</td>
</tr>
</tbody>
</table>

For header/class 1, the coding gain of soft decision decoding over the uncoded BPSK is
\[
G = 10 \log_{10} \left( \frac{rd_{\text{min}}}{(1 - ND)^2} \right) = 10 \log_{10} \left( \frac{8 \times 7}{16} \right) \approx 5.44 \text{ dB}
\]
and is approximately 2 dB over hard decision decoding.

For class 2, the coding gain of soft decision decoding over the uncoded BPSK is
\[
G = 10 \log_{10} \left( \frac{rd_{\text{min}}}{(2 - I - ND)^2} \right) = 10 \log_{10} \left( \frac{8 \times 4}{12} \right) \approx 4.12 \text{ dB}
\]
and is approximately 2 dB over hard decision decoding.

Figure 2 shows the theoretical bounds of the header and the individual classes of a frame, using the equations obtained above.

5. SIMULATION RESULTS

There are two methods to test the performance of EPTOOL. The first method used is to pass sample bitstreams into the encoder, simulate noisy conditions on the channel, decode the output bitstream that is corrupted by noise and then compare the original bitstream with the decoded bitstream to determine the number of residual bit errors (BE). The residual BE is a measure of the number of residual errors that occur after decoding by FEC codes applied to the bitstream. The second method involves computing the theoretical bounds to obtain the optimum performance for the original and improved EPTOOL. Next, a comparison is made with the simulated results to find the divergence.

Both schemes are tested with several standard audio clippings provided by MPEG. These audio clippings are sampled at a rate of 48kHz and 16 bits per sample in mono. The bitstream of audio sequences are subjected to random white noise.

In the first method, the encoded bitstreams are subjected to channel bit error rates of \(1.00 \times 10^{-03}\) to \(5.00 \times 10^{-02}\). The improvement in average residual BER for the original and improved EPTOOL is summarized in Table 2. The improvement percentage is defined as
\[
\text{Improvement} = \frac{\text{Original Average BER} - \text{Improved Average BER}}{\text{Original Average BER}} \times 100\%
\]

<table>
<thead>
<tr>
<th>BER</th>
<th>R 1 BER (Original)</th>
<th>Average R 1 BER (Improved)</th>
<th>Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.00E-03</td>
<td>3.35E-04</td>
<td>3.04E-04</td>
<td>9.24%</td>
</tr>
<tr>
<td>5.00E-03</td>
<td>1.72E-03</td>
<td>1.35E-03</td>
<td>21.50%</td>
</tr>
<tr>
<td>1.00E-02</td>
<td>4.60E-03</td>
<td>2.76E-03</td>
<td>40.00%</td>
</tr>
<tr>
<td>2.00E-02</td>
<td>1.18E-02</td>
<td>5.10E-03</td>
<td>56.78%</td>
</tr>
<tr>
<td>3.00E-02</td>
<td>1.93E-02</td>
<td>8.69E-03</td>
<td>54.97%</td>
</tr>
<tr>
<td>4.00E-02</td>
<td>1.32E-01</td>
<td>1.31E-02</td>
<td>90.08%</td>
</tr>
<tr>
<td>5.00E-02</td>
<td>2.08E-01</td>
<td>2.02E-02</td>
<td>90.29%</td>
</tr>
</tbody>
</table>

From Table 3, it can be seen that there is a significant improvement of 50% to 90% in average residual BER at channel BER of \(2.00 \times 10^{-02}\) to \(5.00 \times 10^{-02}\). As decoding of a header leads to the lost of the entire frame, at high BER, the average residual BER of the original scheme is even greater than the channel BER.

The cost incurred in the proposed scheme is 21%, increased in bandwidth compared to original signal, which is the same as that of the original scheme. Hence, there is no additional redundancy. The results for both schemes are shown in Figure 3.

From Figure 3, it can be seen that at low signal-to-noise ratio (SNR), the proposed scheme is able to achieve a coding gain of about 2 dB. However, at high SNR, due to the uncoded class 3 bits for both schemes, the proposed scheme does not show an increase in performance over the original scheme.

Next, the performance of EPTOOL is evaluated by computing the theoretical bounds for the BER of the
original and proposed scheme. Figure 4 shows the comparison between the simulated and theoretical curves of the original and proposed scheme.

Figure 3: Graph of simulated BER Against SNR per bit for original EPTOOL and improved EPTOOL

The original scheme shows a divergence at low SNR values between the simulated and theoretical curves due to the fact that decoding of a header in error could lead to the loss of the entire frame. However, in the proposed scheme, the simulated and theoretical curves are very close to each other. Hence, within experimental error, the proposed scheme can perform near optimum as compared to the current scheme.

Figure 4: Graph of theoretical and simulated BER Against SNR per bit for the original and improved EPTOOL

6. CONCLUSION

An enhanced UEP scheme based on EPTOOL, which involves soft decision decoding of convolutional codes for wireless transmission of MPEG audio has been proposed and tested successfully. Comparison with theoretical bounds demonstrates a near optimum performance for the proposed scheme. It is also found that the error protection tool does help in effectively protecting error sensitive bits and at the same time maintain low overheads. The scheme will be tested in the real world wireless channel.

Another avenue of possible future research can be in the field of turbo product codes. The importance of turbo product codes is that they enable reliable communications with power efficiencies close to the theoretical limit predicted by Claude Shannon. They also significantly outperform many other coding schemes available today.

7. REFERENCES


