

UNEQUAL ERROR PROTECTION METHODS FOR PERCEPTUAL AUDIO CODERS

Deepen Sinha and Carl-Erik W. Sundberg

Bell Laboratories, Lucent Technologies
Murray Hill, NJ 07974

ABSTRACT

In most source coded bit streams certain bits can be much more sensitive to transmission errors than others. Unequal error protection (UEP) offers a mechanism for matching error protection capability to sensitivity to transmission errors. A UEP system typically has the same average transmission rate as a corresponding equal error protection (EEP) system but offers an improved perceived signal quality at equal channel signal to noise ratio. In this work we introduce methods of UEP to the perceptual audio coder (PAC). An error sensitivity classifier divides the bits in classes of different sensitivity. Different channel codes are then applied to each class. We show how punctured convolutional codes can be used for UEP of the PAC bitstream. Experimental results for channels with uniform as well as non-uniform noise/interference level indicate that the systems with UEP exhibit graceful degradation and extended range for applications such as digital audio broadcasting (DAB).

1. INTRODUCTION

Audio coders such as the Lucent Technology's Perceptual Audio Coder (PAC) [1], attempt to minimize the bit rate requirement for the storage and/or transmission of digital audio data by the application of sophisticated hearing models and signal processing techniques. In the absence of channel errors, PAC is able to achieve transparent stereo compact disk (CD) quality at the rate of approximately 128 kbps. At the bit rate of 96 kbps the quality is still fairly close to that of CD audio. The rate of 96 kbps is particularly attractive for the FM band type of in-band DAB system (also known as hybrid IBOC and all digital IBOC systems) [1, 2, 3]. The transmission channels for these systems are severely bandlimited and noise limited at the edge of the coverage area with potential fading problem for a mobile receiver. Therefore it is important to design an error-protection scheme that is closely matched to the error sensitivity of the various bits in the compressed audio bit stream.

PAC and similar audio compression techniques are inherently packet-oriented in nature; i.e., audio information for a fixed interval (frame) of time is represented by a variable bit length packet. Each packet consists of certain control information followed by quantized spectral/subband description of the audio frame. The key idea behind an unequal error protection scheme is that different components of a packet exhibit varying sensitivity to channel errors. This may happen because of two reasons: (i) It may be more difficult to conceal errors in some bits than others. For example, corrupted control information leads to loss of synchronization and breakdown of the error concealment algorithm. (ii) Transmission errors in different audio components have varying perceptual implications.

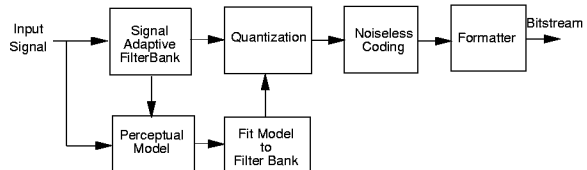


Figure 1: Overview of the Perceptual Audio Coder (PAC)

In previous work on Digital Audio Broadcasting [2], a two level error protection scheme was employed using Reed-Solomon (RS) code. There the control information was protected more robustly with a repeat code and the RS decoder was used to produce a “failure to decode” signal to activate the error mitigation algorithm.

In the present work we introduce a multi-level unequal error protection (UEP) scheme which exploits the unequal impact of transmission errors on various audio components by partitioning the PAC encoded bitstream into two or more classes. Each may then be protected at a varying level using a variety of options; e.g., block or convolutions channel codes of different rates [5]. In the latter case a class of codes called Rate Compatible Punctured Code (RCPC) is particularly suited to the UEP scheme [9]. For error mitigation purposes, an outer CRC code or RS code is typically used in conjunction with the convolutional code. The proposed UEP method can, in principle, be used with other audio codecs, as well as other transmission channels than radio. In this paper, we also present another method for UEP which bases bit placement on the varying interference levels in the channel.

2. THE PERCEPTUAL AUDIO CODER (PAC)

PAC at its core is a perceptually driven adaptive filterbank or transform coding algorithm. It incorporates advanced signal processing and psychoacoustic modeling techniques to achieve high level of signal compression. Details of the PAC algorithm may be found in [1]. Fig. 1 is a high level view of PAC. In brief, PAC uses a signal adaptive switched filterbank which switches between a Modified Discrete Cosine Transform (MDCT) and a wavelet transform to obtain compact description of the signal. The filterbank output is quantized using non-uniform vector quantizers. For the purpose of quantization, the filterbank outputs are grouped into the so called *codebands* so that quantizer parameters, e.g., stepsizes are independently chosen for each codeband. These stepsizes are generated by a psychoacoustic model. Quantized coefficients are further compressed using an adaptive Huffman coding scheme. In PAC a total of 15 different codebooks are employed and for each coder-

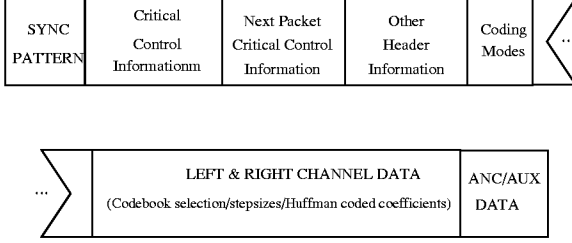


Figure 2: PAC Stereo Bitstream (Packet) Description

band the best codebook may be chosen independently. For stereo and multichannel audio material sum/difference or other form of multichannel combinations may be encoded.

PAC Bitstream Description

The compressed audio information in PAC is formatted into a packetized bitstream. PAC is a block processing algorithm and at a 44.1 kHz sampling rate, each packet corresponds to 1024 input samples from each channel, regardless of the number of channels. The Huffman encoded filter bank outputs, codebook selection, quantizers, and channel combination information for one 1024 sample block is packed into one packet. It should be clear that the size of the packet corresponding to each 1024 input audio samples is a variable; although over a long term a constant average packet length may be maintained if necessary.

Depending on application, various extra information is added to the first frame or to every frame. For unreliable transmission channels, like DAB, a header is added to each frame. This header contains critical PAC packet synchronization information for error recovery and may also contain other useful information such as sample rate, transmission bit rate, audio coding modes, etc. The critical control information is protected further by repeating it into two consecutive packets (this is an example of a very simple form of UEP in the PAC framework). An illustration of the PAC bitstream is shown in Figure 2.

3. ERROR SENSITIVITY CLASSIFICATION

Here we focus on a two class UEP scheme which is based on a partitioning of the PAC audio bits in each packet into two classes, 1* and 2*. In the simplest implementation the relative sizes of the two classes, (1* and 2*) are constant from packet to packet; i.e., a fixed $P\%$ of audio bits in each packet are classified as *critical*. The classifier for allocation of audio bits into the two groups as well as the parameter P specifying the relative rate of the code groups are chosen based on perceptual principles and subjective measurements. Two simple perception rules in the classifier are as follows: (i) Partial loss of stereo separation in the signal is typically less annoying to an average listener than spectral distortions in the center channel; and, (ii) The frequency weighting due to the middle ear implies that the low to mid frequency coder bands (corresponding to the frequency range of 100 Hz - 4kHz) are typically more critical than the rest of the spectral information.

Experiments indicate that for the case of an audio coder a relatively large value for P ($0 < P < 50\%$) is desirable. This is because a relatively large number of bits are consumed by the low to mid frequency coder bands (as high as 50 – 70% for certain packets). Therefore in our initial work a value of $P = 50\%$ was employed. The classifier then works as follows. For each packet

with a size K bits, a *bit bucket* of size $P * K/100$ is created. The bucket is then filled with bits from this packet by applying a sequence of perceptually driven selection rules. First, the control information is placed in the bucket. Then low and mid, frequency components from the center channel followed by low and mid frequency components from the side channel are sequentially placed into the bucket. If it is not possible to incorporate all of the bits in the mid frequency range, a *sub-critical band comb* is applied to the spectral bits to select part of these bits. The process terminates at the time when the bucket is full. It may be at times be necessary to include some or all of high frequency bits from center and side channels in the bucket. The classifier for monophonic spectral information may then appear as in Figure 3.

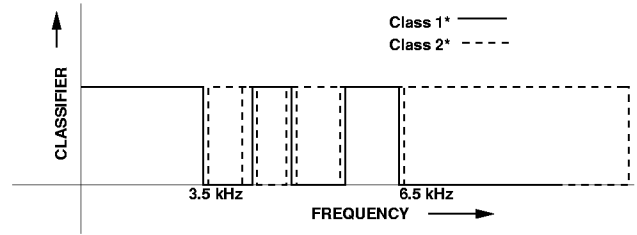


Figure 3: Classification Strategy for Monophonic Spectral Information

The bits in the bit bucket represent the more critical audio information and receive a higher level of channel error protection as Class 1* bits. The remainder of the bits form Class 2* for this packet and are given less channel protection. Some channel coding and transmission schemes for the two classes are discussed in the next section.

In extensions of the above two class UEP scheme, the audio information may be classified in more than two classes. Another more elaborate extension of the UEP scheme utilizes a multi-band audio signal loudness profile over a multiple audio frame duration to choose an error protection profile for the corresponding group of packets. This profile (from a finite number of possibilities) will then be transmitted as control information to the receiver. Such a multipacket error correction profile is attractive for PAC operating at lower bit rates because they provide lower audio bandwidths. The classification strategy as described above may not work successfully for an individual packet if audio bandwidth is only 8 – 9 kHz. On the other hand a multipacket error correction profile will be able to exploit the variability in the bit demand as well as the interpolative capabilities of the mitigation algorithm to classify a suitable multipacket subset of the critical information as a separate class.

4. UEP CHANNEL CODING

In this section we discuss two approaches for achieving UEP on the classified PAC bitstream. The first approach is based on the use of different channel codes for each class. The second approach is based on a single channel code.

UEP Scheme Based on Multiple Channel Codes

A simple unequal error protection (UEP) scheme for the two class partitioned audio bitstream is illustrated in Fig. 4. The audio bits are divided into classes, 1* and 2* as above, and these are protected independently first by an “outer” coder (CRC of equal

block length and strength) and then with “inner” coder of differing strength; code 1 and 2 respectively (as seen in Fig. 4 control information is further protected with a repeat code). The relative allocation of bits into one of the code groups and the rates of the two codes are referred to as the *error protection profile* for the UEP scheme. All error protection profiles are constrained to ensure that the overall rate is equal to the rate in the case of equal error protection scheme. In the simplest embodiment, code 1 and code 2 are fixed rate error protection codes. In addition the relative data bit allocation to a code group is also a constant parameter, $P = 50$ for all the packets.

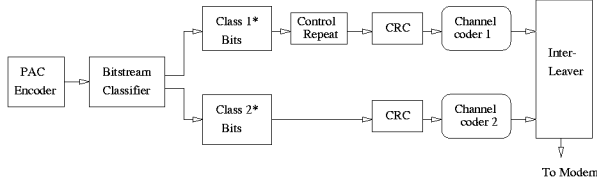


Figure 4: Conceptual transmitter of a baseline 2/3 level error protection scheme. A matching receiver is used

The convolution codes used in the introductory UEP experiment [6] are punctured codes from a rate $1/3$, memory 6 mother code with free distance $d_f = 14$. For the equal error protection case we use a rate $4/10$, $d_f = 11$ code. For the unequal error protection case we have a rate $4/11$, $d_f = 12$ code for the more sensitive Class 1* bits and a rate $4/9$, $d_f = 10$ code for the less sensitive Class 2* bits. The puncturing period is $P = 8$ for these memory 6 (64 states) codes. A code search was performed to find the best code for EEP and UEP. In general, the rates for the codes protecting the different classes of bits are chosen to meet some average rate constraint. If the fraction of Class 1* bits is f and the fraction of Class 2* bits is $1 - f$, then the rates of the codes must satisfy

$$\frac{1}{R} = \frac{f}{R_{1*}} + \frac{1-f}{R_{2*}} \quad (1)$$

where R is the average rate, and R_{1*} and R_{2*} are the rates of the codes protecting Class 1* and 2* bits, respectively. For example, with $R = 2/5$ and $f = 1/2$, the pair of rates $(R_{1*}, R_{2*}) = (4/11, 4/9)$ satisfies the above constraint.

Note that we are comparing the codes at equal E_s/N_0 rather than E_b/N_0 for QPSK. In general all simulations were run on an additive white Gaussian noise channel, characterized by the E_D/N_0 , energy per dimension over noise power spectral density. This figure is related to more conventional measures by

$$\frac{E_s}{N_0} = N_D \frac{E_b}{N_0} = R \frac{E_b}{N_0} \quad (2)$$

where R is the rate of the code in information bits per dimension (e.g., convolutional code rate for BPSK or QPSK signaling) and N_D is the number of dimensions per symbol (e.g., 1 for BPSK, 2 for QPSK) [1].

Unequal Error Protection Using One Channel Code

The unequal error protection approaches discussed as far are such that the improvements are obtained for any type of channel. The channel noise is assumed to be averaged over time and frequency by interleaving in both time and frequency for each channel code class. Thus a UEP scheme with a more powerful channel code

properly matched to the most sensitive source bits always out performs the EEP scheme.

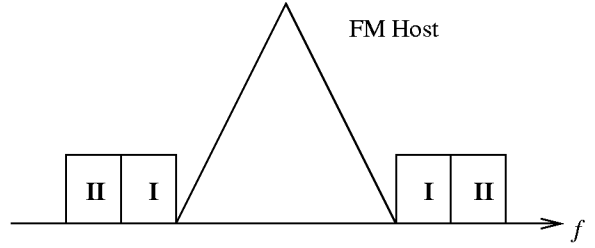


Figure 5: Hybrid in band on channel FM spectrum with two sidebands divided in two classes. Class I is assumed to be less susceptible to interference. The most sensitive source bits are transmitted here.

An alternative UEP approach, which basically only gives improvements for certain nonuniform interference channels can be described as *frequency division UEP* (FD-UEP), in contrast to time division (TD) UEP, which was considered earlier. A simple FD UEP scheme with two classes of bit sensitivity is illustrated in Figure 5, which shows a simplified power spectrum of a hybrid IBOC FM system.

It has been suggested that the part of the power spectrum at the highest and lowest frequencies are more susceptible to interference from other broadcasting radio stations than the part closest to the analog host FM spectrum. For such a scenario, a two class UEP scheme like the one in Figure 5 will give improvements over a conventional EEP scheme. In the simplest scenario, the same channel code is used for classes I and II, however with separate interleaving in time and frequency. In the simulations we used rate $2/5$, memory 6 code with $d_f = 11$.

It is clear that the FD UEP scheme using the same channel codes gives no improvement for channels with uniform interference level in frequency (using different channel codes will). However, it does provide an increased robustness against increased interference levels in region II.

5. EXPERIMENTAL RESULTS

A classifier for the 2 class UEP scheme discussed in section 3 and 4 was implemented for a proof of the concept; i.e., to determine the potential advantages of the UEP scheme in terms of perceived audio quality in the presence of simulated channel errors. In the first set of bounding experiments, based on 96 kbps stereo PAC, only one of the two classes (1* and 2*) of compressed audio bits were exposed to simulated random channel errors with certain bit error probabilities. Informal listening tests clearly indicate that perceived audio quality is significantly more sensitive to the bit errors in Class 1* than to the bit errors in Class 2*. Such bounding experiments are useful in determining the best pair of UEP codes matched to the sensitivity of the two classes since it is clear from section 4 that several different code pairs may satisfy the UEP rate constraint. More details on this code selection procedure based on a combination of subjective listening and code performance data may be found in [6],[7].

In the next set of experiments, the two class partitioned UEP PAC bitstream was channel coded using the codes discussed in the preceding section (with an average channel coding rate of $2/5$ for

both UEP and EEP). The PAC bitstreams were transmitted over simulated Gaussian channels and compared at equal E_s/N_0 . For simplicity we used ideal binary phase shift keying (BPSK) in these introductory simulations. Brief subjective tests were conducted to assess the subjective quality of received audio material under different channel conditions. Results of these tests using 3 audio source clips and 4 listeners are summarized in Table 1. The scores

Table 1: PAC Subjective Test Results for 96 kbps UEP vs. EEP for two different channel conditions.

Channel Condition (E_s/N_0 dB)	UEP Preference Score
-0.5	0.88
-1.0	1.33

measure the relative preference for the UEP scheme over the EEP scheme (under identical channel condition) on a 5 point scale. On this scale, which ranges from -2 to +2, a score of +2 reflects strong preference for UEP while -2 reflects a strong preference for EEP. As an indicator of the “quality of the channels”, E_s/N_0 of -1.0 dB corresponds to a decoded bit error rate of about 2.10^{-4} for the EEP channel code. (The EEP system here operates close to the point of failure). The somewhat cleaner channel at -0.5 dB E_s/N_0 , corresponds to a decoded bit error rate of about 6.10^{-5} for the EEP channel code.

Results in Table 1 clearly indicate that the UEP technique yields improved quality for the same transmission channels. Furthermore the advantage of UEP increases with a more severe channel. This implies extended range for digital audio transmission. More structured experiments with specific RCPC codes and RS codes are underway. See [6],[7] for further details.

A third set of experiments were conducted to determine the viability of FD-UEP scheme over a channel with unequal interference level. In these simulations, the channel conceptually consists of two disjoint segments or partition, I & II. With suitable interleaver depth, the channel quality may be assumed to be constant over a particular segment. The two segments then can be parameterized by the corresponding noise level measured in terms of E_s/N_0 . Gaussian channel conditions are assumed in these simulations. and a single convolutional code with a rate of 2/5 was used throughout. Further detail about the experiments may be found in [1]. Summary of the subjective audio quality evaluation for the EEP and FD-UEP systems under different channel conditions is summarized in Table 2. Expectedly, if the channels conditions on the two segments are roughly equivalent, both EEP and FD-UEP systems perform similarly (simulation 1 in Table 2). On the other hand from simulations 2-3 in Table 2 it is clear that when the conditions in the two channel segments are substantially different the FD-UEP system exhibits much more graceful degradation. More specifically, if moderated channel conditions exist in segment I and segment II is approximately 2.0 dB worse the EEP system is unacceptable with muting nearly half the time. The FD-UEP system on the other hand survives with reduced audio bandwidth and some increase in distortions. When the channel condition in segment II is about 2.5 dB worse than segment I, EEP system mutes more than 75% time while FD-UEP survives albeit with lower audio bandwidth. In other words as the interference in segment II increases, the audio quality in FD-UEP “bottoms out” with a lower yet often

acceptable level of audio quality; the EEP system in comparison mutes almost completely under these similar conditions.

Table 2: Simulated FD-UEP and EEP systems for a channel with unequal interference condition, in E_s/N_0

Exp. No.	E_s/N_0 dB Segments I/II	EEP Quality	FD-UEP Quality
1.	-0.5/-0.5	Good	Good
2.	-0.5/-2.5	breakdown 50% muting	some distortion reduced BW some noise bursts
3.	-0.5/-3.0	Breakdown > 75% muting	reduced BW some distortions

6. DISCUSSION AND CONCLUSIONS

The present work establishes that for DAB, UEP results in improved range and graceful degradation compared to EEP. The UEP schemes for PAC is in principle also applicable to a wide range of noisy channels, e.g., Internet, cellular multimedia, satellite channels, etc. Investigations of UEP schemes based on multipacket error protection profiles are being considered for lower rate PAC. Multichannel and multiprogram audio coding scheme could offer other avenues for more powerful UEP schemes.

7. REFERENCES

- [1] Sinha, D., Johnston, J. D., Dorward, S. M., and Quackenbusch, S. R. “The Perceptual Audio Coder (PAC),” *The Digital Signal Processing Handbook*, pp. 42-1:42-17, CRC/IEEE Press, Editors: V. K. Madisetti and D. B. Williams, 1998.
- [2] Jayant, N. S., and Chen, E. Y. “Audio Compression: Technology and Applications,” *AT&T Technical Journal*, pp. 23-34, Vol. 74, No. 2, March-April, 1995.
- [3] Sundberg, C-E. W., Sinha, D., Kroon, P., and Juang, B-H. “Technology advances enabling ‘In-Band on Channel’ DSB systems,” *1998 International Conference on Broadcast Asia*, Conf. Rec. pp 289–296, Singapore, June 1998.
- [4] Lin, S., and Costello Jr., D. J. *Error Control Coding: Fundamentals and Applications*. Prentice Hall, Englewood Cliffs, NJ, 1983.
- [5] Cox, R. V., Hagenauer, J., Seshadri, N., and Sundberg, C-E. W. “Sub-band speech coding and matched convolutional channel coding for mobile radio channels,” *IEEE Transactions on Acoustics, Speech & Signal Processing*, pp. 1717-1731, Vol. 39, No. 8, August 1991.
- [6] Sinha, D. and Sundberg, C-E. W. “Enhanced Perceptual Audio Coders with Unequal Error Protection for Digital Audio Broadcasting,” *Preprints of the AES 104th Convention*, Amsterdam, May 1998.
- [7] Sinha, D. and Sundberg, C-E. W. “Unequal Error Protection for Compressed Audio Bit Streams Using Single or Multiple Channel Codes,” *In Preparation*.