SUB-BAND CODING OF AUDIO USING RECURSIVELY INDEXED QUANTIZATION

Yang-Jeng Chen
Robert C. Maher

Department of Electrical Engineering and Center for Communication and Information Science
University of Nebraska-Lincoln
209N WSEC 0511
Lincoln, NE 68588-0511
phone: +1 402-472-2081
rmaher@unl.edu

ABSTRACT

A low-complexity audio data compression technique using sub-band coding (SBC) and a recursively indexed quantizer (RIQ) is proposed in this paper. The characteristics of the proposed system are investigated and several relevant implementation issues are described. The objective performance of the proposed system is compared to conventional coding techniques. A real time implementation of the proposed system running on a DSP microprocessor is also discussed. The results indicate that the RIQ-SBC system exhibits objective signal-to-quantization noise ratio 2 to 5□dB higher than other coders of similar computational complexity when processing a variety of wideband audio signals.

1. INTRODUCTION

Lossy compression techniques are commonly used in speech, audio, and image coding. Practical systems appropriate for real time audio coding require consideration of the tradeoff between subjective quality and computational complexity, i.e., the ability to obtain an acceptable level of quality within the available hardware/software limitations.

The study reported in this paper involves an objective performance evaluation and real time implementation of a proposed low-complexity coder for wideband audio, based on the well-known sub-band coding (SBC) framework [1]. The proposed system uses recursively indexed quantizers (RIQs) to code the sub-band channels. The RIQs are non-saturating variable rate fixed quantizers which eliminate overload distortion [2].

In this paper the objective performance of the proposed three-band, non-adaptive RIQ-SBC is compared to a three-band SBC using conventional quantizers. A real time demonstration system incorporating the RIQ-SBC structure is also presented.

2. SUB-BAND CODER

The basic block diagram of the proposed three-band sub-band coding system is shown in Figure 1. This nonuniform three-band SBC is designed by means of a two-level quadrature mirror filter (QMF) bank. In the first level a pair of 32-tap FIR filters are used for the half-band subdivision, followed by a pair of 16-tap FIR filters for the second subdivision in the lower frequency bank only.

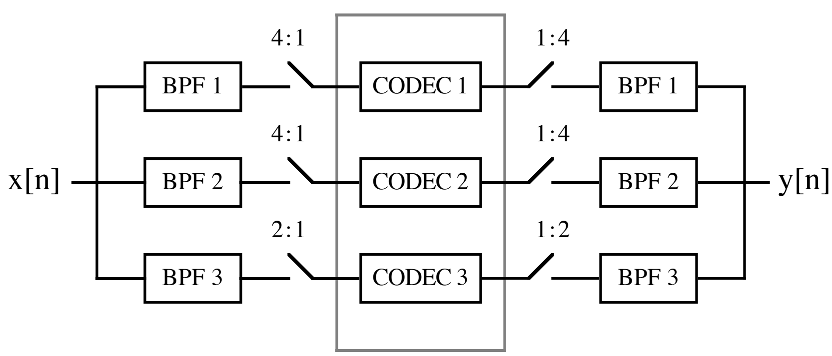

Figure 1: Block diagram of three-band sub-band coder

3. RECURSIVELY INDEXED QUANTIZER

Conventional scalar quantizers suffer from two types of quantization error: granular distortion and overload distortion. Granular distortion occurs when the input signal presented to the quantizer falls in one of the inner quantizer levels and is therefore bounded to the size of the quantizer steps. Overload distortion occurs when the input signal falls in one of the outer levels of the quantizer, resulting in a clipped quantizer output signal. Unlike the finite error size of granular distortion the error due to quantizer overload is essentially unbounded.

Sayood and Na [2] have proposed the use of a recursively indexed quantizer in order to eliminate overload distortion. The RIQ produces only granular distortion by providing a mechanism for an instantaneous increase in output bit rate whenever the base range of the quantizer is exceeded.

The RIQ concept can be summarized as follows [3, 4].

A uniform round-to-nearest-level quantizer is defined with step size Δ and minimum and maximum quantizer levels denoted by X_L and X_H, respectively.

- If the current quantizer input sample falls within the inner steps of the quantizer, the RIQ output is simply the quantizer symbol corresponding to the input value. This is known as the *non-recursive mode*.
- If the current quantizer input sample is extreme in magnitude so that the uniform quantizer output would be X_L or X_H, the RIQ enters the *recursive mode*. In the recursive mode the difference signal ,X_1, is calculated between the current quantizer input and the nearest quantizer level, i.e., $X_1 = X - X_H$ or $X_1 = X_L - X$. If X_1 falls in the inner levels of the uniform quantizer the recursive process terminates. On the other hand, if X_1 exceeds the uniform quantizer range a second difference signal, X_2, is calculated and compared to the base range of the quantizer. This recursive process continues until the difference value, X_m, falls in the inner levels of the quantizer.

The quantizer output for X exceeding the level of X_H in the recursive mode consists of the sequence:

$$Q(X) = \underbrace{X_H, X_H, \ldots X_H}_{m}, X_m$$

where m denotes the number of recursions needed to quantize input value X. The output sequence for X below X_L is similar, except the repeated output value is X_L rather than X_H:

$$Q(X) = \underbrace{X_L, X_L, \ldots X_L}_{m}, X_m$$

Thus, the X_H and X_L symbols appear in the RIQ output stream *only* when the quantizer operates in the recursive mode. This property allows the receiver of the RIQ data stream to reconstruct the input sequence by monitoring the stream for X_H or X_L symbols: the reconstructed sample is the "sum" of the *m* repeated symbols and the following difference symbol.

4. RIQ-DPCM

Differential pulse-code modulation (DPCM) is used to remove redundancy from an input data stream by quantizing the difference between the current input sample and a linear prediction of that sample [5]. The basic structures of a DPCM coder and decoder are shown in Figure 2.

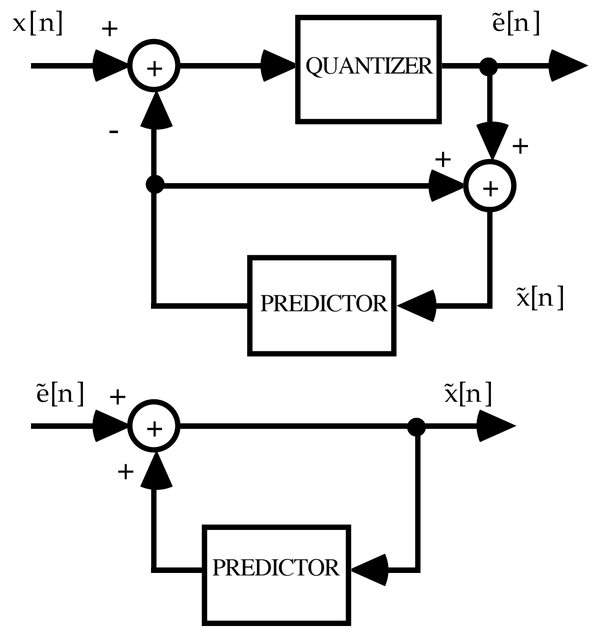

Figure 2: Basic DPCM structure (encoder and decoder)

RIQ-DPCM is essentially a conventional DPCM structure in which the difference signal is quantized with a recursively indexed quantizer. Several aspects of the objective and subjective performance of RIQ-DPCM systems have been reported previously [3, 4, 6]. In particular, RIQ-DPCM coding of wideband audio has been shown to provide subjective and objective quality better than several conventional coders of similar computational complexity, due to avoidance of overload distortion [7].

5. RIQ-SBC

In a three-band SBC system as shown in Figure 1 the lowest band is often coded with DPCM while the upper bands are coded with PCM. In the proposed three-band system the lowest band is coded with RIQ-DPCM and the upper bands are coded with RIQ-PCM. In this way it is possible to eliminate overload distortion in each band so that the reconstructed signal contains only granular error. It is important to evaluate the bit rate properties of the RIQ-SBC system when comparing the rate-distortion-computation tradeoffs among various coding schemes.

The optimum fixed bit-allocation strategy of SBC depends upon the energy distribution of the input signal and the desired quantization noise properties of each band [8]. Typical audio signals exhibit more energy in the lower bands than in the upper bands, so the lower bands are generally allocated more bits in order to maintain fidelity. A better strategy is to provide for dynamic bit allocation, but dynamic schemes are outside the computational limits associated with this investigation. The optimum quantizer step size is also dependent upon the statistical characteristics of the sub-band signals. As with the bit allocation, an adaptive strategy for the quantizer is preferred if sufficient computational resources are available.

6. OBJECTIVE PERFORMANCE OF RIQ-SBC

Non-real time implementations of three-band SBC and three-band RIQ-SBC were developed to evaluate the rate-distortion performance of these coders and to perform objective quality comparisons. Two other low complexity fixed coding schemes, DPCM and RIQ-DPCM, were also implemented to allow additional objective performance comparisons.

In this experiment a variety of different 16-bit, 44.1kHz sample rate music sources were transferred digitally from compact disc (CD) sources. The examples were monophonic excerpts (705.6 kbits/s) 10-15 seconds in duration of four musical styles: rock and roll, female vocal solo, symphony orchestra, and jazz combo. A fixed allocation of bits and a fixed quantizer step size were used to code each example (non-adaptive).

The average and segmental signal-to-quantization noise ratios were computed for each of the four musical examples using each of the four coders. The resulting overall SNR values were then averaged for each of the four coders in order to give a single figure of merit over the range of musical signal characteristics. The overall average SNR and segmental SNR of the tested coders are summarized in Figures 3 and 4, respectively. An improvement in objective quality could be expected in situations where the additional computational overhead of adaptive processing is allowable.

The RIQ-SBC system performed 3 to 5 dB better than SBC without RIQ for the average SNR comparison, and about 1 to 3 dB better in the segmental SNR test. In addition, RIQ-SBC performed 6 to 7 dB better than conventional DPCM in both the average and segmental SNR comparisons.

7. REAL TIME IMPLEMENTATION

A real time implementation was investigated to evaluate the feasibility of RIQ-SBC coding using a standard DSP microprocessor. A two-band RIQ-SBC structure with 16-tap FIR filters (QMF) was used for this test. The lower band was coded with RIQ-DPCM and the upper band was coded with RIQ-PCM. A Motorola DSP56001 processor (27 MHz) installed on an Ariel Corp. DSP-56 board was used as the target system.

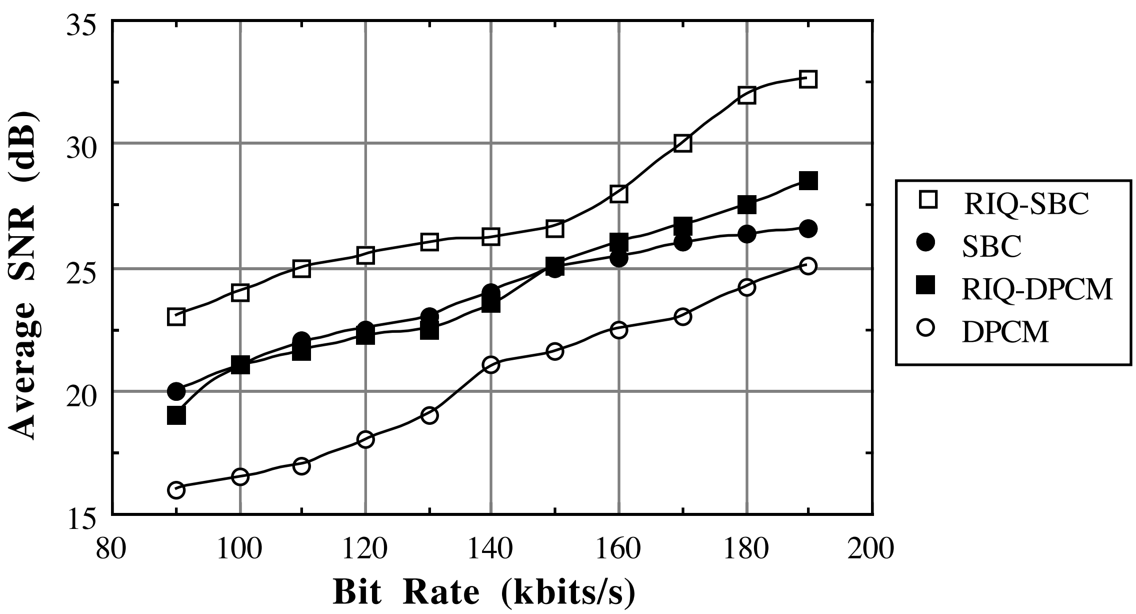

Figure 3: The overall average SNR of RIQ-SBC, SBC, RIQ-DPCM, and DPCM for the 4 musical selections.

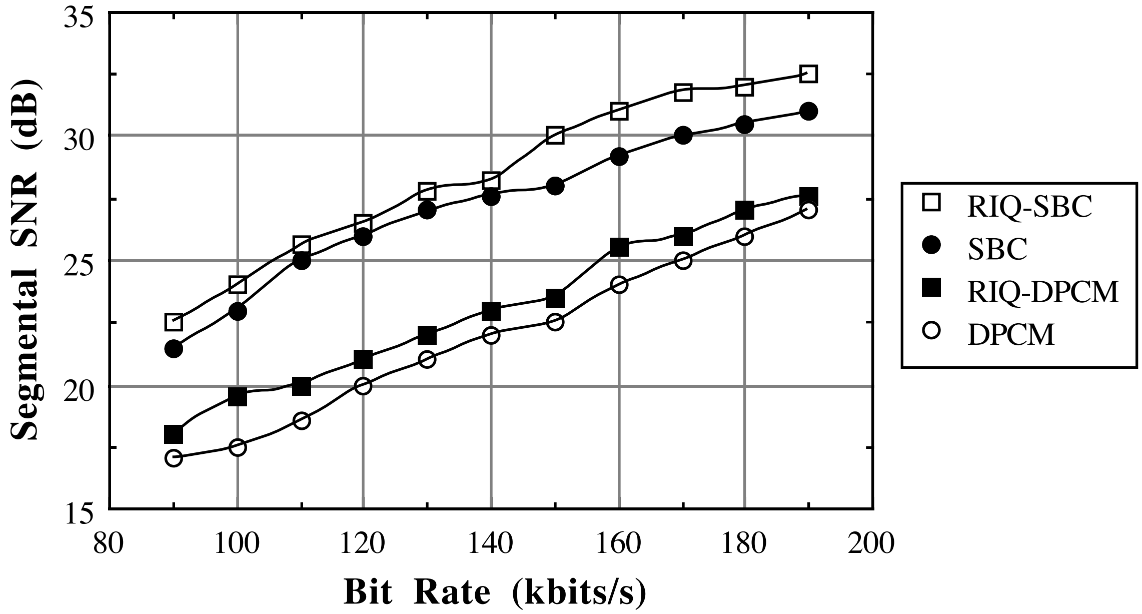

Figure 4: The overall segmental SNR of RIQ-SBC, SBC, RIQ-DPCM, and DPCM for the four musical selections.

Programs written in non-optimized assembly code were sufficiently fast to provide a sustained single channel of encoding or decoding up to a 50 kHz sample rate. Other real time programs providing a 25 kHz sample rate per channel for either two channels of encoding, two channels of decoding, or one channel of simultaneous encode/decode were also implemented. Extending these results to a larger number of sub-bands or to an adaptive algorithm would require either a reduction in sample rate, a highly optimized implementation, or a faster DSP chip.

8. CONCLUSION

In this paper the use of a recursively indexed quantizer within the sub-band coding framework has been presented. The proposed system is able to eliminate quantizer overload distortion in all sub-bands without a significant increase in bit rate. The procedure is also of low computational complexity suitable for real time processing.

The work reported here is based in part upon a thesis submitted by Yang-Jeng Chen in partial fulfillment of the requirements for the degree of Master of Science in Electrical Engineering, University of Nebraska-Lincoln.

9. REFERENCES

1. N. S. Jayant and P. Noll, Digital Coding of Waveforms, Englewood Cliffs, NJ: Prentice-Hall, 1984.
2. K. Sayood and S. Na, "Recursively indexed quantization of memoryless sources," *IEEE Trans. Inf. Theory*, vol. 38, no. 5, pp. 1602-1609, 1992.
3. K. Sayood and S. Na, "Recursively indexed differential pulse code modulation," *Joint DIMACS/IEEE Workshop on Coding and Quantization*, October 1992, Rutgers Univ., Piscataway, NJ.
4. R. C. Maher, "Computationally efficient compression of audio material by means of RIQ-DPCM," *Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*, New Paltz, NY, October, 1993.
5. N. S. Jayant, "Digital Coding of Speech Waveforms: PCM, DPCM, and DM Quantizers," *Proc. IEEE,* vol. 62, pp. 611-632, 1974.
6. R. C. Maher, "An efficient scheme for lossy realtime audio data compression," *Proc. 1994 Audio Engineering Society Convention*, Preprint #3922 October, 1994.
7. S. M. Joseph and R. C. Maher, "Subjective evaluation of four low-complexity audio coding schemes," to appear, *J. Acoust. Soc. Am.*, 1995.
8. J. D. Johnston, "A filter family designed for use in quadrature mirror filter banks", *IEEE Proc. 1980 Int. Conf. Acoust. Speech and Signal Processing*, pp.□291-294, 1980.