MULTIPLE-DESCRIPTION CODING OF LOGARITHMIC PCM

Kai Clüver, Thomas Sikora

Communication Systems Group, Technische Universität Berlin
Einsteinufer 17, 10587 Berlin, Germany
phone: +49 30 314-24588, fax: +49 30 314-22514, email: {cluever;sikora}@nue.tu-berlin.de
web: www.nue.tu-berlin.de

ABSTRACT

A practical approach for the design of multiple-description scalar quantization of speech is presented that conforms to standard G.711 PCM. The method chiefly consists of an index assignment algorithm that enables the side decoders to exhibit SNR characteristics comparable to those of the standard logarithmic quantizer. With two-channel transmission of multiple descriptions, an increase in robustness to lossy channels is obtained without violation of the standard coding method. The method found is suitable for the design of multiple descriptions of any given scalar quantizer, e. g. one within a complex speech coder.

1. MULTIPLE-DESCRIPTION CODING

Multiple-description coding [1] provides a transmission link with diversity in order to improve robustness to channel breakdown. The coded signal is split into two or more descriptions, or partial codes, which are transmitted over the same number of different channels. These channels may indeed consist of different physical links, or of different packets transmitted through networks like the internet. The priciple of a two-channel multiple-description (MD) coded transmission is shown in fig. 1. From the input signal, $x(n)$, the encoder generates two descriptions C_1 and C_2 to be sent over two lossy channels. If no loss occurs, both descriptions will be used by the central decoder to reconstruct the signal $y_0(n)$ with high quality. If one of the descriptions is lost, the received part of the code will enable its corresponding side decoder to yield a reduced-quality version of the output signal, $y_1(n)$ or $y_2(n)$. The transmission will be interrupted only when both descriptions are lost.

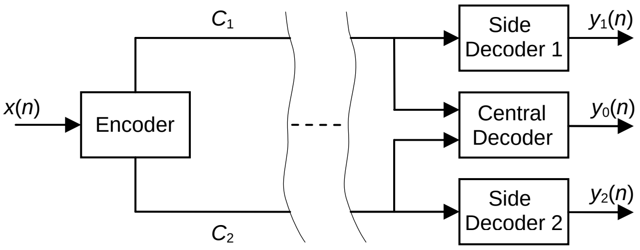

Fig. 1 Structure of a multiple-description coded transmission with two channels

The design of MD coders is subject to conflicting requirements [1]. If the side decoders were optimized for high signal quality, given the bit rates R_1 and R_2 for C_1 and C_2, little would be gained by combining both descriptions in the central decoder, which would then yield a similarly high quality, but at a considerably increased bit rate of $R = R_1 + R_2$. If, on the other hand, the central decoder were designed for minimum distortion at a bit rate of R, any splitting of the code would result in poor performance of the side decoders. Therefore, the objective is to find a compromise for central and side decoder qualities.

For multiple-description scalar quantizers, a design algorithm was introduced in [2] which iteratively optimizes both the side decoders and the central decoder for a known probability density function of the input signal. The results obtained depend on the method of assigning pairs of code indices (for two MD channels) to the quantizer steps.

While recently much attention has been given to MD image coding [1], relatively few publications have dealt with MD coding of speech. An early method consists of splitting PCM or DPCM coded speech signals into odd and even samples [3] [4]. MD scalar quantization within a transform coder was investigated [5], and diversity approaches for several speech coders were proposed [6] [7]. Another approach, which consists of transmission of the full-rate coded speech signal and low-rate additional information via two channels, results in unbalanced MD codes [8] [9] [10] [11].

This work focuses on the design of balanced multiple descriptions for a standard logarithmic PCM speech codec. In the following, the design of the MD PCM codec is described and some results are discussed.

2. DESIGN OF THE MD CODEC

2.1 Objectives

The two-channel MD codec is based on a standard G.711 PCM speech codec [12]. In the encoder, the signal sample $x(n)$ is quantized according to the logarithmic A-law compression characteristic. Then, the quantizer step index is converted into a pair of indices which form the desriptions C_1 and C_2 to be transmitted. If both indices are received, the central decoder recovers the quantizer step index and issues the corresponding reconstruction value $y_0(n)$, again according to G.711. The design aims at balanced descriptions, i. e. equal bit rates ($R_2 = R_1$) and equal distortions of the side decoders. Furthermore, it is desirable that the side decoders show a behaviour comparable to logarithmic quantization with regard to their signal-to-noise ratio (SNR) characteristics, i. e., like the central decoder, yield a virtually constant SNR for a wide range of signal levels.

In order to preserve conformity to the standard PCM codec, neither the quantizer part of the encoder nor the central decoder of the MD codec are subject to the optimization process. Thus the design problem reduces to finding an appropriate index assignment strategy and then calculating the side decoders which result from the central decoder and the actual index assignment.

2.2 Index Assignment

An example of an index assignment for the side decoder rate of $R_1 = 3$ bit/sample is shown in fig. 2. The quantizer step codewords are renumbered according to their reconstruction values and then progressively assigned to cells of the matrix from upper left to lower right. The row index i and the column index j form the two descriptions of the corresponding quantizer index. For minimum side decoder distortion, the occupied cells are as close to the main diagonal as possible [2]. In order to obtain balanced distortions, care was taken to find a progression pattern that results in increments and decrements evenly distributed among both side decoder indices.

i \ j	0	1	2	3	4	5	6	7
0	0	2	3					
1	1	4	6	10				
2	5	7	9	12	13			
3		8	11	14	16	20		
4			15	17	19	22	23	
5				18	21	24	26	30
6					25	27	29	32
7						28	31	33

Fig. 2 Example of an index assignment matrix for balanced multiple descriptions

The standard G.711 PCM codec has 256 quantizer levels, which corresponds to a bit rate of $R_0 = 8$ bit/sample. Side decoder rates of 4 to 7 bit/sample were considered. With $R_1 = 4$ bit/sample, all cells of the index assignment matrix are occupied. As there is no redundancy in the descriptions ($R = 2R_1 = R_0 = 8$ bit/sample), the side decoder performance will be poor. Increasing R_1 results in an incompletely populated matrix ($R > R_0$), which means introducing redundancy and therefore correlation between the descriptions, and thus in decreasing side decoder distortion.

Usually, the selection of the main diagonal and additional pairs of diagonals of the matrix results in a higher number of index pairs than necessary. Consequently, a number of index pairs on the outer diagonals have to be discarded. In order to preserve the logarithmic behaviour, the assigned cells should be spread evenly along the diagonals, while retaining the symmetry of the assignment pattern. For this purpose, a two-pass discarding algorithm was found. In the first pass, every second index pair on the outer diagonals is discarded, starting from the centre of the matrix. If necessary, remaining index pairs are discarded in a second pass, again beginning at the centre. The algorithm is stopped as soon as the number of assigned index pairs equals the desired number of quantizer levels (i. e. 256). Fig. 3 shows a part of the resulting index assignment matrix for $R_1 = 6$ bit/sample (or $R = 12$ bit/sample).

i \ j	0	1	2	3	4	5	6	7	8	9	10	11	12
0	0	2	3										
1	1	4	6	10									
2	5	7	9	12	X								
3		8	11	13	15	18							
4			14	16	17	20	X						
5				X	19	21	23	26					
6					22	24	25	28	X				
7						X	27	29	31	34			
8							30	32	33	36	X		
9								X	35	37	39	42	
10									38	40	41	44	X
11										X	43	45	47
12											46	48	49

Fig. 3 Upper left corner of the index assignment matrix for $R_0 = 8$, $R_1 = 6$ bit/sample; "x" denotes a discarded index pair.

2.3 Calculation of the Side Decoders

For squared error measures, the optimum reconstruction level of a side decoder for index i (or j) is equal to the expectation value of the combined input signal distributions within the quantizer intervals in row i (or j, respectively) [2]. Since the logarithmic quantizer is designed for a wide range of input levels, no distinct known distribution can be used for the optimization of the side decoders. Instead, equal probabilities of the quantizer intervals and uniform signal distribution within each interval were assumed for simplification. Then, the optimum reconstruction level for any side decoder index is equal to the mean value of the central decoder levels involved.

3. RESULTS

3.1 SNR Measurement

The side decoders were evaluated by measuring the SNR as a function of the variance of a zero-mean Gaussian white noise input. The results for side decoder rates of $R_1 = 7$, 6, and 5 bit/sample are shown in fig. 4. MD coding without redundancy, i. e. with $R_1 = 4$ bit/sample, turned out to be useless, as it resulted in extreme distortions of the side decoder output. With some redundancy, however, the signals reconstructed by the side decoders show the desired behaviour. But for very low input levels, the SNRs are perfectly balanced.

For comparison, the SNR measurement was also carried out for single-description logarithmic PCM with bit rates of $R_0 = 3$ to 8 bit/sample (cf. fig. 5). In all cases considered here, the central decoder is identical to 8-bit PCM, independently of the side decoder rates. The SNR of the side decoder output with $R_1 = 7$ bit/sample is only about 2 dB below the SNR of 7-bit PCM, $R_1 = 6$ bit/sample yields a characteristic almost identical to 5-bit PCM, and lowest redundancy ($R_1 = 5$ bit/sample) roughly corresponds to 3-bit logarithmic PCM.

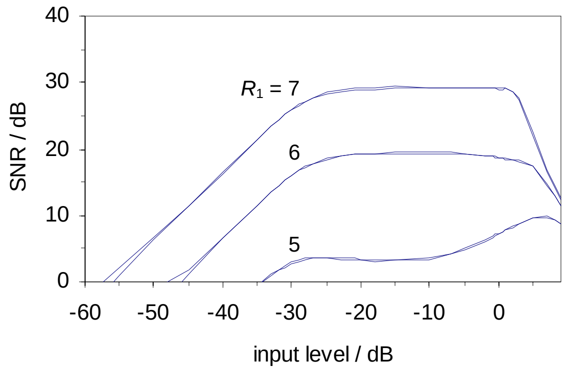

Fig. 4 SNR of both side decoders as a function of input level for three bit rates; a level of 0 dB corresponds to four-sigma loading of the quantizer.

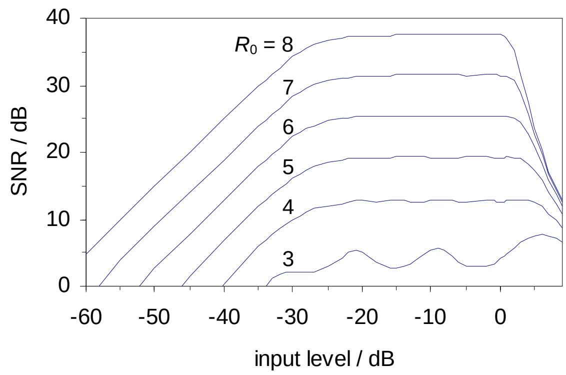

Fig. 5 SNR characteristics of logarithmic PCM for several bit rates; $R_0 = 8$ bit/sample corresponds to standard G.711 A-law PCM.

3.2 Packetized MD Coded Speech

In order to judge the behaviour of multiple-description coded transmission, an experiment on packetized PCM speech was carried out. For consecutive coding intervals of 20 ms duration, two packets were generated of which each contains one of the coded descriptions (C_1 or C_2). Statistically independent packet losses were simulated, and decoding was performed using central or side decoders according to the number of packets received for each interval. If both packets and, consequently, both descriptions were lost, the missing speech samples were replaced by zero values. The simulations were done using a speech data file of 6.5 s duration.

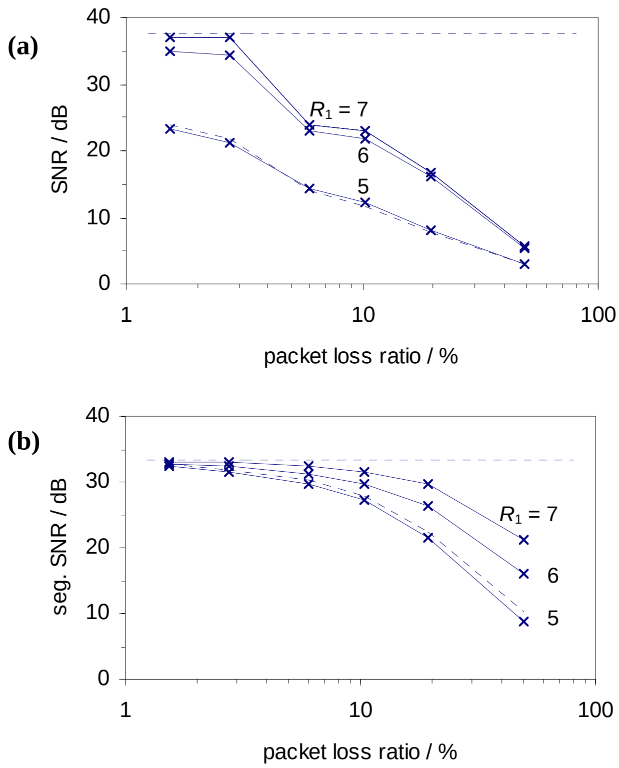

Fig. 6 (a) SNR and (b) segmental SNR of MD coded speech signals as a function of the packet loss ratio

Fig. 6 shows the SNR and segmental SNR results obtained for three side decoder bit rates. For comparison, the horizontal dashed lines indicate the SNRs of lossless transmission (i. e. of the central decoder) and the dashed curves show the results of conventional ("single-description") packetized transmission with zero substitution for packet losses.

A side decoder bit rate of $R_1 = 5$ bit/sample does not yield any improvement compared to single-description coding, neither in terms of SNR nor of segmental SNR. This corresponds to the findings of informal listening tests: the decoded speech signals are easily distinguishable from each other, but no preference can be stated. With higher rates, however, significant gains can be achieved for packet loss ratios of up to 20 %. The resulting SNRs for both rates ($R_1 = 6$ or 7 bit/sample) are significantly higher than those of single-description transmission. Again, the results were confirmed in informal listening tests: while only few zero-signal substitutions remain in the decoded speech, most packet losses are transformed into soft rustling noises which are still highly acceptable at R_1 = 6 bit/sample and almost inaudible at a side decoder rate of 7 bit/sample.

4. CONCLUSION

The practical approach for the design of multiple-description coded logarithmic PCM chiefly consists in an index assignment method that achieves "logarithmic" behaviour of the side decoder output signals. With two-channel transmission of multiple descriptions, an increase in robustness to lossy channels is obtained without violation of the standard coding method. Especially with 50 % redundancy added (i. e. a total rate of 12 bit/sample), a suitable compromise was found of moderately increased bit rate and - with one channel broken down - noisy but still highly acceptable speech quality.

Apart from link failure, multiple-description coding of PCM speech signals can also enhance robustness to losses in packetized speech transmission, e. g. within multi-media applications. With the bit rates chosen appropriately, single packet losses will only cause a slight decrease of the signal quality but not a complete drop-out of the signal.

The index assignment strategy outlined above is suitable for the design of multiple descriptions of any given scalar quantizer, e. g. one within a more complex speech coder, if the side decoders are required to preserve the characteristics of the original quantizer.

REFERENCES

[1] V. K. Goyal, "Multiple Description Coding: Compression Meets the Network", *IEEE Signal Processing Magazine*, vol. 8, no. 5, pp. 74-93, Sept. 2001.

[2] V. A. Vaishampayan, "Design of Multiple Description Scalar Quantizers", *IEEE Trans. on Information Theory*, vol. 39, no. 3, pp. 821-834, May 1993.

[3] N. S. Jayant, S. W. Christensen, "Effects of Packet Losses in Waveform Coded Speech and Improvements due to an Odd-Even Sample-Interpolation Procedure", *IEEE Trans. on Communications*, vol. COM-29, no. 2, pp. 101-109, Feb. 1981.

[4] A. Ingle, V. A. Vaishampayan, "DPCM System Design for Diversity Systems with Applications to Packetized Speech", *IEEE Trans. on Speech and Audio Processing*, vol. 3, no. 1, pp. 48-58, Jan. 1995.

[5] J.-C. Batllo, V. A. Vaishampayan, "Multiple Description Transform Codes with an Application to Packetized Speech", in *Proc. Int. Symp. on Information Theory*, Trondheim, Norway, June/July 1994, p. 458.

[6] C.-C. Lee, "Diversity Control among Multiple Coders: A Simple Approach to Multiple Descriptions", in *Proc. Workshop on Speech Coding*, Delavan WI, Sept. 2000, pp. 69-71.

[7] X. Zhong, B.-H. Juang, "Multiple Description Speech Coding with Diversities", in *Proc. ICASSP 2002*, pp. I-177 - I-180.

[8] R. Singh, A. Ortega, "Erasure Recovery in Predictive Coding Environments using Multiple Description Coding", in *Proc. Workshop on Multimedia Signal Processing*, Copenhagen, Sept. 1999, pp. 333-338.

[9] A. K. Anandakumar, A. V. McCree, V. Viswanathan, "Efficient, CELP-Based Diversity Schemes for VoIP", in *Proc. ICASSP 2002*, pp. 3682-3685.

[10] W. Jiang, A. Ortega, "Multiple Description Speech Coding for Robust Communication over Lossy Packet Networks", in *Proc. ICME 2000*, pp. 444-447.

[11] Y. J. Liang, E. G. Steinbach, B. Girod, "Real-Time Voice Communication over the Internet using Packet Path Diversity", in *Proc. ACM Multimedia*, Ottawa, Sept./Oct. 2001, pp. 431-440.

[12] N. S. Jayant, P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs NJ: Prentice-Hall, 1984.