Proc. Biennial Symp. Commun. (Kingston, ON), pp. C-2.6–C2.7, May 1984

Variable Rate Coding of Speech Signals

C. Emond
INRS-Telecommunications
University of Quebec

now with
Bell Canada
1060 University
Montreal, Quebec H3B 3A2

P. Kabal
Electrical Engineering
McGill University
Montreal, Quebec H3A 2A7
and
INRS-Telecommunications
University of Quebec
Verdun, Quebec H3E 1H6

This paper examines variable rate digital coding of speech signals. This type of scheme tries to accommodate the non-stationary nature of the input signal by coding the incoming speech samples in blocks. The total number of bits for a block is fixed, but within a block, the number of bits allocated for the quantization of a given sample can vary. One approach approach to variable rate coding is to try to maximize the signal-to-noise ratio (SNR) of the coded signal. The bit allocation is based on an estimate of the local energy of the speech samples. It is shown that this technique can theoretically offer significant improvement in SNR over fixed rate coding for speech signals. To be effective, the variable bit allocation requires a good estimate of the short term energy variation of the speech signal. The same estimate must be available to both the coder and decoder. This implies an overhead for the transmission of side information to the decoder.

The first approach taken in this work was to use a good signal variance tracker based on past input values. This system realizes a significant SNR gain over a fixed rate coder for moderate block lengths. However this increase in SNR does not manifest itself as a comparable increase in perceived quality. A more perceptually

meaningful measure, segmental SNR, shows only a marginal increase. The accuracy of the energy estimate also has a significant effect on the SNR gain. For time constants of the energy estimator corresponding to practical amounts of side information, much of the SNR gain is lost.

A system proposed by Dubnowski and Crochiere [1] was also simulated. Results again show an improved SNR without an attendant increase in perceptual quality. It is felt that a time domain variable rate coder can achieve at most modest perceptual gains over a fixed rate coder. This is due to the fact that variable rate coding tends to make the distortion constant throughout a block of speech. For low level segments, the quantization distortion can overwhelm the speech signal. The success of variable rate coding used to minimize the mean square error seems not to be in time domain coders but in frequency domain techniques such as Adaptive Transform Coding.

1. J.J. Dubnowski and R.E. Crochiere, "Variable Rate Coding of Speech", BSTJ, vol. 58, March 1979, pp. 577-600.