DIGITAL SPEECH PROCESSING METHODS

S. J. CAMPANELLA
COMSAT Laboratories
Clarksburg, Maryland

The digital transmission of speech has,,in recent years, advanced at a rapid pace. This is due in part to the fact that speech in a digital form enjoys the advantages of binary data transmission techniques and also due to the fact that common digital carriers can be used for both voice and data communications.

This paper is devoted to a survey of methods for converting the analog speech signal into a digital data stream. The most commonly known technique is that of pulse code modulation (PCM), wherein speech sampled at a Nyquist rate of 8000 samples/second is coded in binary form. PCM is currently used in commercial and military communications systems for speech transmission at bit rates from 32 to 64 kbits/second. Such processing is influenced by a number of variables, such as the type of code employed, the number of bits per sample, and the companding law used, and the influence of these will be considered.

Another form of processing closely related to PCM is differential pulse code modulation (DPCM). By this method, it is the objective to take advantage of the correlation that exists from sample to sample as a consequence of the properties of speech. In DPCM such correlation is used to reduce the number of bits/sample required to achieve a given level of performance as gauged in terms of signal-to-quantizing noise ratio.

Going further in the properties of speech, predictors have been used to achieve significant reductions in the bit rate required for transmitting speech. Rather sophisticated predictors have been demonstrated using computer simulation which promise to provide very good quality speech communications at bit rates as low as 2400 bits/second. A simpler application of the predictor principle is currently being developed for use in multitrunk telecommunications which cuts in half the bit rate required for such communications.

Delta modulation provides yet another means for achieving digital transmission of speech over a wide range of transmission bit rates. Simple forms of delta modulation have been used for over two decades. In recent years, relatively sophisticated formats of delta modulation incorporating feedback-controlled variable slope have made it possible to achieve good quality speech at data rates as low as 20 kbits/sec. with acceptable performance extending to as low as 10 kbits/sec.

Another area which has received attention in recent years has been the use of orthogonal transformation techniques, wherein the speech is first transformed from the time domain into the transform coefficient domain, the coefficients coded for transmission and at the receiver, converted once again to the time domain. This form of processing has provided a reduction in the bit rate required for communications of speech of a given quality. Transformation techniques which have been investigated included Hadamard, Fourier, and Karhunen-Loeve. Reductions in bit rate on the order of 10 kilobits per second have been achieved using these techniques.