

GTE Communications
Transmission Systems

DEMODULATOR NOVEMBER/DECEMBER 1982

Lebra

PCM Channel Bit Rate Reduction

Today, pulse code modulation (PCM) common carrier telephone systems use a 64-kb/s encoding scheme for each voice channel. However, a substantial amount of research is underway to develop schemes that use an encoding rate considerably less than 64 kb/s per channel. The advantage of a lower bit rate is that more voice channels can be carried by a given transmission system, considering its digital rate or bandwidth.

The first reference in the bibliography divides speech quality into four categories: toll, broadcast or commentary, communications, and synthetic. Figure 1 shows these classifications in terms of voice channel rate in kb/s.

Toll quality is comparable to analog speech transmitted over the telephone network. Experimental laboratory model encoders can achieve toll quality at rates of 16 kb/s, but 64 kb/s is the present standard in the U.S public switched network.

Broadcast or commentary quality is higher than toll quality. It is used when the input signal bandwidth is wider than the normal telephone channel. Rates greater than 64 kb/s are required for broadcast quality transmission.

Communications quality is lower than toll quality. Although the speech is distinctly intelligible, sometimes the ability to recognize the talker is noticeably reduced. Digital encoders operating at rates below 16 kb/s can provide communications quality.

Synthetic quality is the lowest of

the four. Speech does not sound natural, talkers may not be recognized, and encoder performance is dependent upon talker pitch and volume. Encoders operating at 4.8 kb/s and below can provide only synthetic quality.

The speech quality we are interested in is toll quality. The advantage of reducing the encoding rate can be seen by considering that a 32-kb/s rate would double the capacity of existing PCM systems. This would be particularly advantageous to satellite communications where bandwidth is the limiting factor on channel capacity. In fact, the International Telecommunications Satellite Organization (Intelsat) has done much research in this area.

Published experimental data from tests of various 32-kb/s techniques show that this encoding will be capable of providing toll quality speech. Bell Telephone Laboratories has indicated they will recommend a technique sometime in 1983, and the Consultative Committee on Interna tional Telephone and Telegraph

Figure 1. Speech Coding Transmission Rates and Associated Quality

(CCITT) Study Group XVIII plans to recommend a standard in 1984.

Four of the more promising techniques are:

- (1) Adaptive Differential PCM (ADPCM)
- (2) Adaptive Delta Modulation (ADM)
- (3) Nearly Instantaneous Com panding (NIC)
- (4) Sub-Band Coding (SBC)

Each of these four techniques is discussed after the following brief review of PCM theory. The review is presented as a background for understanding the principal subject: bit rate reduction.

PCM Review

Basic operations required for PCM voice transmission systems include Sampling, Time Division Multiplexing, Quantizing, and Encoding. The following paragraphs discuss each of these operations in turn.

Sampling

An analog signal has a characteristic of continuously varying in amplitude with time. Sampling pro-

duces a representation of this signal in the form of equally spaced discrete pulses. The continuous analog signal can be recovered from the discrete representation by passing it through a low pass filter. Theoretically, both the sampling and recovery process can be accomplished without impairing the signal if the rate at which the samples are taken is at least twice the rate of the highest frequency signal component. This is known as the Nyquist rate.

The frequency band of the message is replicated around each harmonic of the sampling frequency in the form of sidebands. If sampling is at less than the Nyquist rate, the sidebands will overlap and aliasing or foldover distortion will result. This effect is illustrated by Figure 2.

Figure 2a shows the frequency spectrum for a 3-kHz message channel sampled at 8 kHz, substantially above the Nyquist rate. The message frequency sidebands are shown around successive harmonics of the sampling frequency. Note that in this theoretical case, the sidebands are of equal amplitude.

(a) Shows sidebands for a 3-kHz signal sampled at 8 kHz (F_s) , which is above the Nyquist rate, (b) Shows the irreversible distortion that results if the signal bandwidth is increased to 6 kHz and the sampling frequency remains at 8 kHz, which is now below the Nyquist rate.

Figure 2b shows the sideband overlap that would occur if a 6-kHz signal were sampled at substantially below the Nyquist rate. After sampling, frequencies are present outside the desired bandwidth. They are translated downward to fall within the original band and cause aliasing distortion.

In practical systems, aliasing distortion is controlled by using a bandwidth filter at the message input and choosing a sampling rate somewhat higher than the Nyquist rate. Therefore, 8 kHz has become the international standard rate for voice telephony where the speech bandwidth is 300 to 3400 Hz.

Other distortions are incurred during the sampling process. One of these, aperture distortion, results from holding the sample for a period of time while other operations, such as coding, are performed. This is compensated for in the output filter.

In summation, real sampling systems introduce distortion. The engineering design problem is to hold this distortion within acceptable limits.

Time Division Multiplexing

In time division multiplexing (TDM), each channel has exclusive use of the transmission path for short time intervals. Each channel transmits successive samples of its signal instead of a continuous signal. At the receiving end, a replica of the original continuous signal is con structed from these samples.

Figure 3 is a simplified presentation of the TDM process. When the clock pulses go high, they connect their assigned channel to the transmission path. Between clock pulses, the channel is isolated from the path. The clock pulses are staggered so at any instant, only one channel is connected to the transmission path.

Figure 3. Diagram of a Simple TDM System.

The time during which the transmission path is allocated to a particular channel is the time slot for that channel. Time slot width is related to the clock frequency (sampling frequency) and the number of channels.

For the 24 channel/24 time slot standard used in the USA and Japan, the time slot width is:

$$
\delta = \frac{T}{24} \text{ where: } T = \frac{1}{F} =
$$

$$
\frac{1}{9000} \text{ second} = 125 \text{ microseconds}
$$

Therefore:

$$
\delta = \frac{125}{24} = 5.2 \text{ microseconds}
$$

The time slot width for the European 30 channel/32 time slot system is

 $\delta = \frac{T}{32}$ (It is the number of time slots, not the number of channels that is important.)

$$
\delta = \frac{125}{32} = 3.9 \text{ microseconds}
$$

Both systems are CCITT standard primary systems.

Until now, what we have been discussing could be a pulse amplitude modulated (PAM) system. The pulses on the transmission path could be varying in amplitude in proportion to the amplitude of the message signals at the sampling instants. However, PAM systems are not used at all because they are semi-analog and would require a virtually distor-

tionless transmission path. The next few paragraphs discuss quantizing and encoding, which overcome the disadvantages of PAM and produce the digital signal used on PCM systems.

Quantizing and Encoding

When a signal sample is quantized, it is limited to one of a fixed set of amplitude levels. There are some disadvantages to assigning a single value to all PAM samples within a specific range. Most samples will not be exactly represented by the quantizing voltage for that particular range, which introduces an error called quantizing noise. This noise results from a coding error and is therefore signal dependent. Nevertheless, the mean square value is interpreted as noise power. Furthermore, a linear coding device with equal separation between quantum steps can be shown to have a constant value of quantizing noise power for all signal levels.

The noise level is therefore the same for both high and low levels of speech. Furthermore, quantizing noise accounts for the total noise in a well-designed PCM system. Since the dynamic range of telephone speech traffic can be greater than 40 dB, it is apparent that a coder-decoder (codec) with linear quantization could have a 40-dB difference between the signalto-noise ratios for weak and strong signals. This condition is unacceptable for telephone transmission.

Quantizing noise could be reduced by increasing the number of levels and thereby reducing the magnitude of the coding error. However, this would require an increase in the number of unique codes used to identify each level and also an increase in the number of pulses required to represent each sample.

As shown previously, the time available for sending each sample is determined by the number of channels in the system and the sampling rate per channel. If the number of bits per sample increases, the transmission bandwidth must in crease. In a practical system, tradeoffs must be made between minimizing quantizing noise and economically using bandwidth.

Non-linear quantization overcomes the quantizing noise problem and provides a more uniform signalto-noise ratio over the dynamic speech range. The quantizing levels are closer together for low level signals than for high so the quantization is much more accurate at low levels. The way a linear quantizing scale can be changed to a non-linear scale is shown in Figure 4.

Referring to the figure, the height of each projected non-linear level is established by the rate of change in slope of the quantizing law. Figure 4 also shows that the errors are much smaller for low level signals than they are for high level signals, which is the desired result.

The quantizing law can be changed to produce even smaller errors for low level signals. So, by choosing a suitable mathematical law, the separation of the quantizing steps can be arranged to fit the requirements of particular transmission systems.

Quantizing laws are also know as compression laws because the overall effect is to compress the analog signal.

Figure 4. Changing a Linear to a Non-Linear Quantizing Scale.

Two natural logarithmic compression laws have been accepted as CCITT standards: the European A-law and the North American μ -law. The formulas for these laws are:

The A-law:

$$
Y = sgn(x) \frac{A|x|}{1 + \ln A}
$$
 for $0 \le |x| \le 1/A$

and

 $Y = \text{sgn}(x) \frac{1 + 1 \text{ mod } 1}{1 + \text{ mod } 1}$ for $1/A \le |x| \le 1$

The μ law:

Y = sgn(x)
$$
\frac{\ln (1 + \mu |x|)}{\ln (1 + \mu)}
$$
 for $-1 \le x \le 1$

where: $1n$ = natural logarithm $x = normalized input$ amplitude restricted to the range 0 to 1 $Y = normalized output$ amplitude restricted to the range 0 to 1 A and μ = constants $sgn = sign$, positive or negative

In this form, the laws express smooth compression characteristics that would be applied to an analog waveform prior to its encoding for transmission. At the receive end, after uniform decoding, the complementary expansion characteristics would be applied to the input signal amplitude distribution. This process is a form of companding that has been used for many years in analog systems. However, unlike analog compandors, which operate at a syllabic rate, PCM compandors must operate within a fraction of a microsecond because each TDM signal sample must be companded.

In early American systems of the DI type, this was accomplished by a diode-type compressor, which produced the compression law, followed by a linear coder. At the receive end, a linear decoder and an expander performed the translation back to the original signal. A value of $\mu = 100$ was used for systems of this type.

The problem with this type of operation is its reliance on properties

of diodes that are time and temperature dependent. This can make it difficult to maintain sufficiently close tracking between the compressor and expandor. The problem is more acute at higher compression ratios.

Modern American systems use non-linear codecs without the intermediate diodes. In the coder, samples of the voice signal are compared to a set of discrete steps and the step nearest the sample is represented by a digital code. At the receive end, the decoder portion of the codec produces an output dictated by the digital code.

The compression law for these systems has a μ equal to 255. Using this law produces an 8-bit code with unequal coding steps (Figure 5), but which may be readily converted to a 13-bit linear code with equal coding steps. Therefore, the companding advantage of the μ = 255 law provides a 13-bit accuracy.

This conversion from a compressed code to a linear code can be done by digital means using binary logic. To achieve a digitally linearizable compression law, the characteristic is approximated by a number of straight line segments as shown in Figure 5.

The characteristic is made up of eight segments positive (0-7) and eight segments negative (0-7). However, the two segments nearest zero are colinear. Therefore, they are considered as one, so the curve is approximated by 15 segments.

Each segment of Figure 5 is composed of a number of coding steps that meet the following requirements:

Figure 5. Segmented Compression Characteristic.

- (1) Each segment has the same number of coding steps.
- (2) Coding steps within each segment are equal.
- (3) Step sizes on adjacent segments are related by powers of two.
- (4) Steps on all segments are integral multiples of the smallest step.

If a μ -law curve is fitted to a segmented compression curve constructed with 15 segments and 16 equal steps per segment, step sizes in a ratio of 2:1 on adjacent segments, and all step sizes an integral multiple of the smallest step, then the closest fit will cause μ to equal 255.

This segmented characteristic is referred to as the μ /255/15. For the A-law, $A = 87.6$ gives a good fit to a 13-segment curve, so the characteristic is termed A/87.6/13.

To summarize our discussion of quantization, the process has an in herent error. A-law and μ -law compression distributes this error in a way that improves system performance without increasing bandwith. This fact is demonstrated by the 30-dB maximum companding advan tage of the μ -law, which gives 13-bit accuracy with 8-bit coding.

The required bit rate for a 24-channel system with an 8-kHz sampling rate and 8-bit encoding is 1.544 Mb/s. If the same system used linear 13-bit encoding, the required bit rate would be 2.504 Mb/s.

The A-law has a maximum advantage of 24 dB, so 8 bits provide a 12-bit resolution. Note that each 6 dB of advantage is equivalent to a 1-bit improvement.

It has been experimentally determined that 2048 linear quantum

steps would be required to cover the voice range of a telephone channel. Without compression, this would require an 11-bit code.

Coding

The final step in the PCM process is to code the quantum steps into digital form. If each quantum step is numbered in decimal form, then some type of digital code can be developed to represent each of the numbers. Ordinarily, a binary code is used that consists of a combination or code group of binary 1 's and 0 's, each group representing a decimal number. Once the code is established, a series of pulses representing the code groups can be used for transmission.

Typically, the coded line signal is a pulse train in which binary 1's are represented by positive or negative pulses and binary O's are represented by spaces (or no-pulses). A binary 1 in any bit position means that the value of the position is to be summed. A binary 0 means that the value is not to be summed.

Intelsat, which needs to provide toll-quality, international telephone channels, is one of the organizations investigating coding schemes that use a rate considerably less than 64 kb/s. From Figure 1, it is apparent that this limits their consideration to rates above about 16 kb/s. Early tests indicate that 32 kb/s is the most likely rate.

In February 1982, Intelsat issued a "progress report" that included the results of the R & D programs up to that time. A portion of the report describes the results of tests of the four previously mentioned promising techniques. These are:

- (1) Adaptive Differential PCM (ADPCM)
- (2) Adaptive Delta Modulation (ADM)
- (3) Nearly Instantaneous Com panding (NIC)
- (4) Sub-Band Coding (SBC)

The balance of this article discusses these four techniques and the test results.

ADPCM

This technique is based on the fact that speech samples are highly correlated if sampled at the Nyquist rate or higher. Variances in the differences between speech samples are much smaller than that of the original set of speech samples. So, a difference or error signal can be generated, quantized, coded, and transmitted using fewer bits than would be required for full samples. The result is a reduction in the overall bit rate. This process is known as Differential PCM (DPCM). It is implemented by a coder with a predictor connected in a feedback configuration.

Because the sample correlations of speech waveforms are non-stationary, the coder would be much more efficient if the coder and/or quantizer circuit characteristics were also variable and could track the correlations. This coding technique is referred to as Adaptive Differential PCM (ADPCM).

ADPCM is also more efficient for mixed traffic; a combination of speech and voiceband data, for example. The reason is that there is a marked difference in the correlations between these types of signals.

An adaptive transversal filter could be used to provide the predictor in an ADPCM coder. This type of filter has the same operating principles as the echo canceller described in the July/August 1981 issue of the Demodulator, "Satellite RTD Parts 1 and 2." However, a filter for ADP-CM would be far simpler than the echo canceller.

ADM

Adaptive Delta Modulation (ADM) is a special case of ADPCM where one bit is used to quantize the prediction error signal. The quantizer output assumes only two values: a positive one or a negative one. Oversampling of the input signal is used to compensate for the coarse quantization. Also, in ADM the quantizing step size can be made to adapt to input signal variations so the process is also known as Continuously Variable Slope Delta Modulation (CVSD).

NIC

Nearly Instantaneous Companding (NIC) is based on the fact that the dynamic range of small groups of speech samples is much smaller than the dynamic range of the whole speech signal. An NIC speech coder accepts blocks of companded PCM speech samples and requantizes each block. The requantizing process is adaptive.

Since the adaptive parameter must be transmitted, the bit rate of an NIC coder is slightly higher than a multiple of 8 kb/s. For example, 34.66 instead of 32 kb/s or 26.66 instead of 24 kb/s.

SBC

Sub-Band Coding (SBC) divides the speech signal into four to eight sub-bands. Each sub-band is translated to a low-pass spectrum, which is sampled at its Nyquist rate. These sampling rates are lower than the 8 kHz used for a PCM voice channel. The sub-bands are encoded according to criteria specific to each band.

Relative Performance

Over the last few years, Intelsat has sponsored work on ADPCM, NIC, and CVSD. The effort has primarily been at 32 kb/s transmission since, as discussed later, this is the rate most likely to be recommended by the CCITT. The results of the Intelsat tests are summarized in the following paragraphs.

Performance of Speech Coding Techniques

Figure 6 illustrates the relative performance of 64-kb/s PCM, 32-kb/s ADPCM, 32-kb/s CVSD, and NIC at 34.66 kb/s and 26.66 kb/s. A fourtone test signal with two tones centering around 850 Hz and the other two tones centering around 1380 Hz simulated a speech signal. Figure 6 shows that in an objective test ADP-CM consistently outperforms CVSD and NIC, in that order, over a wide dynamic range of the input test signal. The CVSD exhibits a performance least dependent on input signal level, while the performance variations of ADPCM and NIC over the range of levels can be clearly seen.

When the same test is repeated over a channel with bit errors introduced, the order of performance no longer holds true over the entire bit error rate (BER) range of 0 to 10⁻³. As can be be seen in Figure 7, at 10^{-4} BER ADPCM still outperforms CVSD and NIC, but the performance difference between ADP-CM and CVSD (or NIC) has been narrowed. At the channel error rate of 10⁻³, ADPCM and CVSD exhibit approximately the same performance, while outperforming both the NIC and even the 64-kb/s PCM.

Figure 6. Objective Signal-to-Noise Ratio.

Figure 7. Objective Signal-to-Noise Ratio as a Function of Bit Error Rate.

Figure 7 also shows the robustness of CVSD and NIC over the wide range of channel error rate. It should be noted however, that bit error rates on the order of 10^{-4} or 10^{-3} rarely occur on the telephone network.

Figure 8 shows the subjective test results of the performance of the various speech coding techniques mentioned in Figure 7. It can be noted from this figure that the subjec-

tive SNRs are considerably higher than the objectively measured SNRs for all coding techniques. At 10^{-6} BER, PCM outperforms ADPCM, 34.66-kb/s NIC, and CVSD in that order. However, at the channel error rate of 10^{-4} , PCM degrades so sharply that ADPCM performs considerably better than PCM, CVSD, or NIC. At the high channel error rate of 10^{-3} , the performance of 34.66

Figure 8. Subjective Evaluation Result.

12

kb/s NIC is the best, followed by approximately the same performance of PCM, ADPCM, and the CVSD. Again, it should be noted that these high BERs are very unlikely on the telephone network.

Performance of Voice Band Data

An important consideration in the choice of a particular speech coding technique is its performance with voice band data. Tables 1 and 2 show the relative performance of the four encoding techniques with voice band data at 2400 b/s and 4800 b/s, respectively. Analog voice band data is encoded, transmitted over a digital channel, and decoded. The resulting data error rates for various digital channel BERs are tabulated. Table 1 shows that at zero channel error rate, while PCM, NIC, and CVSD all provide essentially error-free data transmission, ADPCM generates a significant data error rate. Over the range of channel BERs up to 10^{-2} . CVSD provides the best transmission performance for voice band data at 2400 b/s.

Table 2 shows that when the data rate is increased to 4800 b/s, only PCM and NIC provide essentially error-free transmission while the channel BER is zero. PCM and NIC exhibit similar performance over the wide channel BER range tested. CVSD performance is consistently

Table 1. 2400 b/s Modem Error Rate.

Table 2. 4800 b/s Modem Error Rate.

13

worse than either PCM or NIC, with ADPCM performance being the worst.

Status of CCITT Standardi zation Efforts

International standarization efforts on speech coding techniques have been underway for the last decade. Within CCITT, the group with primary responsibility for standardizing the speech encoding algorithm and recommending uniform performance characteristics is Study Group XVIII. Study Group XVIII receives relevant inputs from Study Group XII on the assessment of speech quality and from Study Group XVI on the network performance objectives.

During the last study period of S.G. XVIII (1977-1980), Question 10 investigated methods of encoding other than PCM. ADPCM at 32 kb/s for voice and voice band data signal transmission was examined in detail. The earlier study period of 1973-1976 dealt mainly with ADM. At the end of the 1977-1980 study period, the majority of Administrations agreed that when using differential methods of encoding, the encoding bit rate

should be a submultiple of 64 kb/s and a multiple of 8 kb/s. The preferable bit rate was 32 kb/s. However, they did not have an opportunity to prepare a draft recommendation on the subject.

In conclusion, Intelsat is interested in bandwidth reduction because Intelsat V and VI satellites will use Time Division Multiple Access/ Digital Speech Interpolation (TDMA/DSI) transmission to in crease channel capacity in com parison to earlier satellite systems. If the bit rate per channel is decreased, even greater gains could be realized.

It is apparent that 32-kb/s encoding could be equally beneficial to domestic satellites and terrestrial systems. Its greatest application might well be common carrier telephone systems. However, some problems could be encountered in reconciling 32-kb/s encoding with the developing Integrated Services Digital Networks (ISDN).

These networks are being developed by GTE and others. Their intent is to provide a single digital network to carry all types of traffic including voice, data, and video. ISDN developments will be the subject of a future issue of the Demodulator.

BIBLIOGRAPHY

- 1. PCM Terminals, Theory and Practice, Hewlett-Packard, South Queensferry, Scotland.
- 2. Readings in Pulse Code Modulation, Second Edition, 1978, GTE Lenkurt, San Carlos, California.
- 3. Satellite Communications Update, Part 2, GTE Lenkurt Demodulator, May/June 1979, GTE Lenkurt, San Carlos, California.
- 4. Contribution of the Director General, BG/T-40-34E and BG/PC-20-10E, Intelsat, 18 February 1982.
- 5. AT&T Contribution to CCITT Study Group XVlll, April 1982. American Telephone and Telegraph Company.
- 6. James Flanagan, et al., Speech Coding, IEEE Transactions on Communications, Vol. COM-27, No. 4, pp. 710-737, April 1979.
- 7. Thomas R. Rowbotham and Kunihiko Niwa, The Application of Low-Bit-Rate Encoding Techniques to Digital Satellite Systems, Fifth International Conference on Digital Satellite Communications Records, pp. 265-272, Genoa, March 1981.
- 8. Contribution of the Director General, Project Completion Report, Project 52. Voice Channel Coding, BG/T-36-13, Intelsat, 2 February 1981.
- 9. CCITT Document COM XVIII No. 1-E, Questions Allocated to Study Group XVIII for the Period 1981-1984, Annex 1, page 60, February 1981.
- 10. CCITT Document COM XVIII No. R4-E, Report of the Meeting of Working Party XVIII/2 (Speech Processing), July 1981.

GTE Lenkurt Incorporated 1105 County Road San Carlos, California 94070

Address Correction Requested

Bulk Rate U. S. Postage

PAID San Carlos, CA Permit No. 37

ES 94804BADG0150SO MR. BROCK W LEE BADGER METER INC ELECTRONICS O 150 E STANDARD AVE 4 94804

9123A Digital Multiplex

The GTE Lenkurt 9123A is a new compact MX3 type digital multiplex for high-capacity digital radio and optical fiber transmission links. It combines lower speed DS1, DS1C or DS2 signals in any combination to achieve the 44.736-Mb/s DS3 rate and is end-to-end compatible with WECo type MX3 digital multiplex.

The 9123A incorporates one-forone protection against high-speed unit failure. Low-speed unit protection on a one-for-seven basis is optional.

The GTE Lenkurt Demodulator is circulated bimonthly to selected technicians, engineers and managers employed by companies or government agencies who use and operate communications systems, and to educational institutions. Permission to reprint granted on request.