Exhibit I

G.722, A New CCITT Coding Standard for Digital Transmission of Wideband Audio Signals

Paul Mermelstein

Introduction

The rapid increase in digital connectivity of telephone networks, brought about by the gradual removal of most analog links, suggests a new look at enhancing the quality of audio transmitted over the telephone network. Pulse-code modulation (PCM) with 64 kbit/s μ-law or A-law (G.711) arose in response to the need for multiple analog-digital-analog conversions of standard 300–3,400-Hz audio signals. Such signals are considered here to be narrowband audio. Modern speech coding techniques permit reduction of the transmitted bit rate, while preserving audio quality, as in CCITT (International Telegraph and Telephone Consultative Committee) Recommendation G.721, where the customary 300–3,400-Hz-wide telephone signal is encoded at 32 kbit/s [1]. Alternatively, one can provide improved audio quality and maintain the transmission rate at 64 kbit/s. Such improvements are most important for audio or audio-visual conferencing applications where one would like to approach the quality of face-to-face communication. CCITT Study Group XVIII recognized the need for a new international coding standard on high-quality audio to allow interconnection of diverse switching, transmission, and terminal equipment and, thus, organized an Expert Group in 1983 to recommend an appropriate coding technique. Hummel [2] provides a good introduction to the working methods of the CCITT. The coding method described in this paper constitutes the group's recommendation, which was approved by the CCITT through an accelerated procedure in 1986. The algorithm represents the results of a joint effort of contributors from around the world,* and is best described in a series of papers presented at Globecom '86 [3]. This paper is meant to be a tutorial discussion, responsibility for which lies completely with the author. Bit-level particulars of the algorithm, although important for correct implementation of the standard, are not discussed in detail. For more complete information, the reader should refer to the forthcoming CCITT document.

Requirements

The main objective for the new standard is to allow speech transmission at 64 kbit/s with quality as high as poissible and significantly better than that provided by 8-bits/sample, 8-kHz sampled PCM coding. If the signal is sampled at 16 kHz, or twice the PCM rate, the spectrum of the signal to be encoded can be extended to about 7 kHz (3-dB point), and this results in a major improve-

* A complete list of the participants in the Expert Group can be found in Appendix 1 to the Report CCITT, COM XVIII-R 17-E, April 1986. Participating organizations included BNR, Canada; CNET, France; FTZ, Federal Republic of Germany; CSELT and SIP, Italy; NTT, Fujitsu and KDD, Japan; PTT, Switzerland; BT, United Kingdom; Bellcore and Comsat, United States. Technical contributions of all participants are recognized and acknowledged without specific credit on individual items.

0163-6804/88/0001-0008$01.00

ment in audio quality. For speech signals, little additional enhancement is achieved by extending the cutoff frequency even higher. Although the need to encode music signals was recognized by the designers, a further gain in quality of transmitted music, which might have been achieved through use of an even higher cutoff frequency, would have required a transmission rate exceeding the 64-kbit/s target and would have increased the cost and complexity of the coding system. Figure 1 compares the attenuation curve chosen for wideband audio to that used in traditional (narrowband) telephony. Note that the end-to-end digital transmission conditions permit the low-frequency response of the coding system to be enhanced further by extending the cutoff frequency to 50 Hz, which improves the naturalness of the audio and results in the wideband audio signal considered here.

Today's voice networks carry a fair amount of voiceband data and analog signaling information. Such traffic can be carried more economically in digital form wherever end-to-end digital connections are available. The codec performance requirements for voiceband data are substantially different from those of voice signals. If the codec were required to encode voiceband data signals as well, its cost and complexity would increase substantially, as evidenced by the G.721 32-kbit/s codec [4]. In very large scale integration (VLSI) realizations, the cost penalty for the additional complexity is less severe than for discrete implementations and manifests itself mainly in increased chip area. If the wideband codec is not required to carry voice-band data, it can be optimized for best performance on speech signals without such penalties.

Speech quality is generally degraded through accumulation of quantization noise introduced at signal conversion points. Since the transmission systems envisioned are digital end-to-end, we require only one analog-digital conversion at the source and one digital-analog conversion at the destination. Robustness to multiple conversion sequenses of analog to digital-encoded representation is not required. However, interconnection of multiple sources of audio at conference bridges is best carried out with a uniformly quantized representation of the digital signal. To allow for multiple bridges in one connection, provision is made for a small number (up to three) of digital encoding/decoding sequences. Furthermore, both narrowband and wideband audio signals may arrive at audio bridges and bridge output should also be available in narrowband or wideband form.

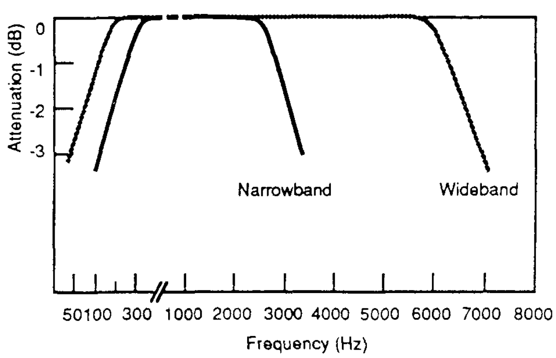

Fig. 1. Frequency Characteristics of Wideband and Narrowband Audio Channels.

The delay introduced by an encoding-decoding process should be limited to 4 ms. This requirement is imposed primarily for echo control purposes. Echo control is simplified by the fact that the common sources of echo on the telephone network, namely the hybrids at two-/four-wire conversion points, are absent in end-to-end digital connections. However, interconnection of existing narrowband links with wideband conferencing systems may introduce sources of echo whose control may be difficult unless the end-to-end delay of the wideband signal is carefully limited.

For some applications, it may be desirable to provide an "in-band" data channel, i.e., to replace part of the 64-kbit/s transmission channel used for speech by data. To permit the least degradation in audio quality consistent with a time-varying demand for data transmission, three speech transmission modes were defined: 64, 56, and 48 kbit/s. Where full 64-kbit/s transmission channels are available, this makes 0, 8, or 16 kbit/s available for data transmission. On those North American channels that are limited to 56 kbit/s, only 0 or 8 kbit/s of data can be transmitted.

An important practical requirement is that the coder provide acceptable performance (maintain intelligibility) for transmission bit error rates up to 10^{3}. This requirement ensures that performance degrades gently even under the worst transmission conditions that one may encounter on the telecommunication network.

Audio Coding Considerations

A tutorial review of modern speech coding techniques can be found in Crochiere and Flanagan [5]. Highly intelligible speech may be transmitted today at rates as low as 800 bit/s. However, most low-bit-rate techniques are restricted to the transmission of speech and are inappropriate for other forms of audio. If one wishes to encode any audio signal, only waveform coding techniques enter into consideration. Forward adaptive techniques require the transmission of some adaptation parameters in addition to the signal obtained by inverse filtering with a prediction filter. Transmission of such parameters, generally considered as side information, requires the use of special framing bits to allow identification of the parameter-carrying bits in the encoded signal. Backward adaptive techniques eliminate the need to transmit such framing bits. Once synchronization between receiver and transmitter has been achieved, groups of eight consecutive bits can be identified, each group carrying similar information about the signal. Recent work on the G.721 standard indicated that 4-bit sample adaptive differential PCM (ADPCM) coding results in toll-quality narrowband speech [6]. Making the quantizer more precise by increasing the number of bits allotted to each sample generally yields no audible improvement beyond 6 bits/sample. Thus, in the absence of coding for bit-rate reduction, 6-bit quantization at a sampling rate of 16 kHz or 96 kbit/s represents an upper limit to the transmission requirements.

Subband coding techniques separate the signal into components occupying contiguous frequency bands and encode the components separately [7]. By appropriately

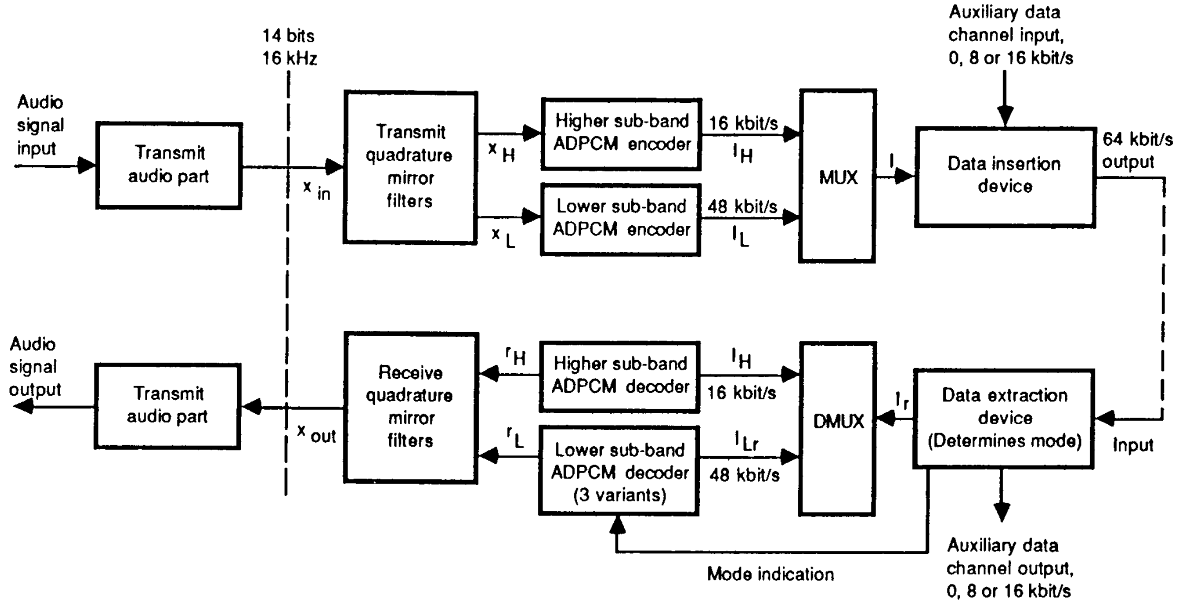

Fig. 2. Block Diagram of the 64-kbit/s (7-kHz) Audio Codec.

allocating the bits across the different bands, the error variance in the reconstructed signal can be shaped with frequency. With the audio signal subdivided into two 4-kHz-wide bands, a high signal-to-noise ratio in the lower band becomes perceptually more important than in the higher band. An advantage of a design that uses two equally wide subbands is that each component is subsampled to 8 kHz and the total transmission rate may be reduced in 8-kbit/s steps by reducing the number of bits assigned to samples of one or the other band. While the bit rate may also be reduced by reducing the sampling rate, those processes generally are more complex to implement. These considerations led to design and evaluation of two alternative subband ADPCM systems, one using 5 and 3 bits/sample for the low- and high-band components, respectively, the other, 6 and 2 bits/sample.

The G.721 ADPCM design employs an adaptive predictor with two poles and six zeros. A fixed predictor design was also tested for the wideband codec, but it led to a generally lower speech quality. A time-varying adaptive allocation of bits to the two subbands according to the short-time signal characteristics was also tried. For voiced sounds carrying significant low-frequency energy, one can assign additional bits to the low band; for fricative sounds, the additional bit may be assigned more advantageously to the high band. However, for the two-band, 4-kHz-wide subband design, the advantage of an adaptive bit assignment is only apparent at the 4 low-2 high bits/sample assignment and is found too small to warrant the additional complexity.

Overall block diagrams for the wideband encoder and decoder are shown in Fig. 2. These blocks are discussed in greater detail in the following sections.

Subband Filtering

The nominal 3-dB band of the codec was chosen as 50–7,000 Hz. Two sets of identical quadrature mirror filters (QMF) are used to divide the wideband signal sampled at a 16-kHz rate into two 8-kHz sampled components to be transmitted, a low band and a high band, and reconstruct the wideband signal from its received low- and high-band components. QMF filters are finite-impulse response, impose a fixed delay without phase distortion, and ensure that aliasing products resulting from subsampling the input signal at the transmitter are canceled at the receiver. However, quantization noise components introduced in coding the low- and high-band signals may not be eliminated completely by the receiver QMF filter. Because the level of the high-band component of the signal may be as much as 40 dB lower than the low-band component, aliasing noise introduced into the high-band frequencies due to coding the low-band signal might be inadequately masked by the high-band signal component. To achieve

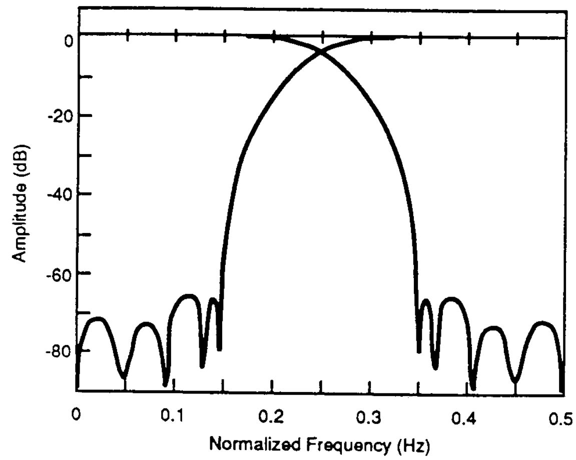

Fig. 3. Frequency Response of the QMF Filters.

a stop-band rejection of 60 dB, we employ a 24-tap filter design, introducing a total signal delay of only 3 ms (see Fig. 3). The resulting signal distortion is below 1 dB over the 100–6,400-Hz band.

The numerical precision with which the partial sums in the QMF filters are accumulated has an important bearing on the accuracy of the low- and high-band signal components generated. The overall goal for the wideband signal representation (after analog-to-digital conversion at the input and before digital-to-analog conversion at the output) is a precision of 14 bits. To this end, the internal coding computations are performed with 16 bits. For the subband signals to have 16 significant bits, the partial sum computations were found to require a precision of 24 bits. Since the wideband signal is accurate to 14 bits, the sum and difference signals to which the QMF filter coefficients are applied are only precise to 13 bits. To prevent the introduction of noise due to differently specified analysis and synthesis filters, the QMF filter coefficients are also represented with 13 bits.

ADPCM Coders

Two ADPCM coders are required, one for the low-band signal and one for the high-band signal. The coders employ identical adaptation strategies to modify the quantizers and predictors based on the previously observed characteristics of the input signal. The low- and high-band coders are very similar, except for small differences due to the need to vary the number of bits output by the low-band coder and the fact that the high-band quantizer output is always 2 bits/sample.

The adaptive predictor design is borrowed directly from that investigated in detail in developing the G.721 standard. The two-pole, six-zero design combines good prediction gain for speech with relatively simple stability control. Robust adaptation is assured by leaky integrators allowing the effects of transmission errors to dissipate rapidly [8]. Transmission errors may introduce differences between the predictor memories at the transmitter and receiver. Adapting the predictors and quantizers using the residual signal alone and not the reconstructed signal ensures that the predictors at the transmitter and receiver recover tracking rapidly for all signals [9]. The adaptive quantizer design is also borrowed directly from the G.721 standard since the low-band signal component resembles the narrowband speech signal in most of its properties. G.721 employs a dual-mode quantizer, a locked or slowly adapting mode for voiceband data signals, and an unlocked or rapidly adapting mode for speech signals. Since the G.722 standard was not required to encode voice-band data

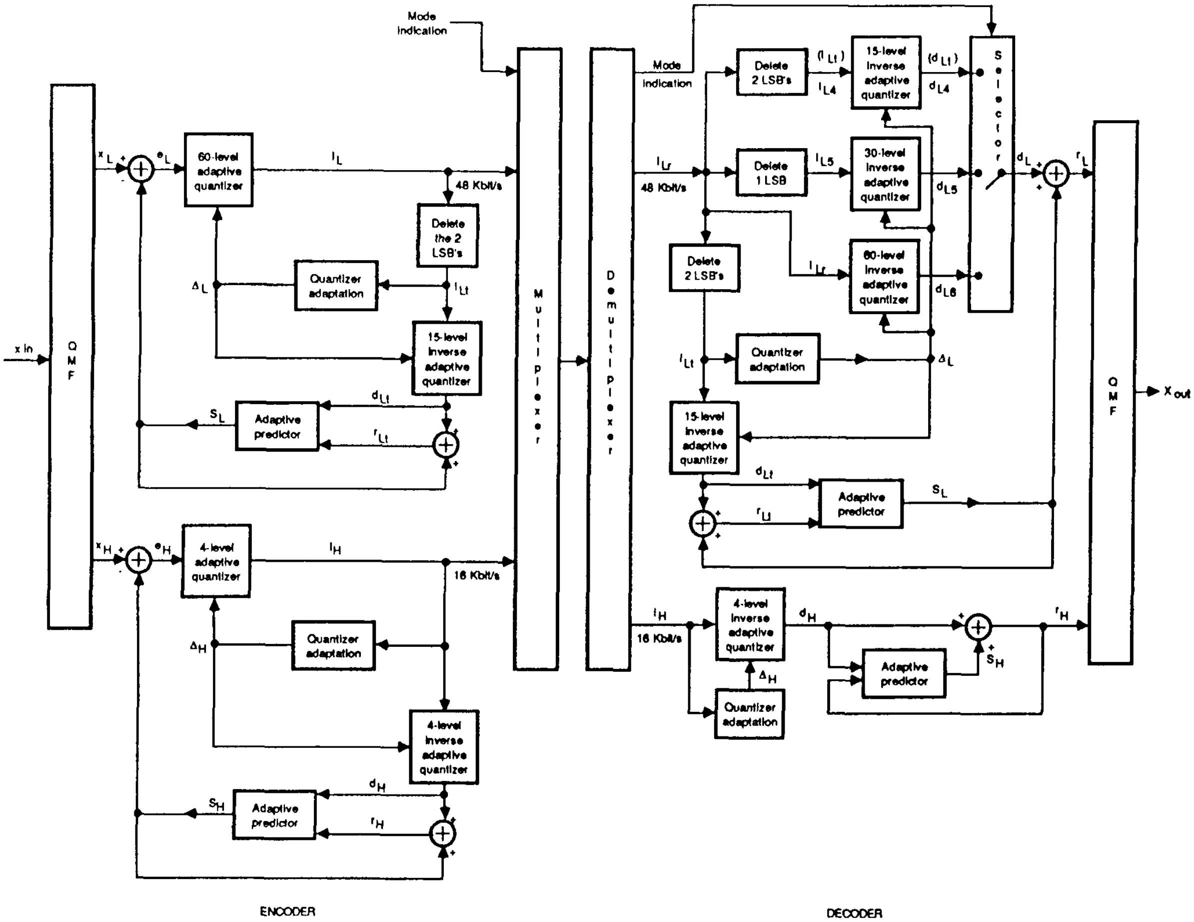

Fig. 4. Detailed Block Diagram of the Subband—ADPCM Encoder and Decoder.

signals, only a single rapidly adapting mode had to be implemented. The dynamic range of the low-band signal quantizer was set to be the same as for G.721, namely 54 dB. A higher dynamic range is allowed for the high-band quantizer, 66 dB, mostly to accommodate music signals. Robust adaptation is employed also for the quantizer scale factor to combat the effects of transmission errors.

Embedded quantizers allow for possible stripping of the less significant bits from the quantized signal during transmission by not making use of those bits in the quantizer adaptation process [10]. The low-band quantizer anticipates that the one or two least significant bits may be stripped from the transmitted code word and adapts the quantizer and predictor using only the four most significant bits. This results in 4- and 5-bit quantizers that are slightly suboptimal in quantization noise-to-signal ratio compared to the quantizers that may be designed without this constraint. The embedding property requires that the 4- and 5-bit quantization boundaries coincide with a subset of those employed for the 6-bit quantizer. Many transmission systems require a minimal number of zero-one alternatives to maintain synchronization. To prevent the all-zero code from appearing even in the 4-bit data representation, only 15 quantizer levels are used in that mode; this also restricts the higher modes to 30 and 60 levels. Experimental evaluations have shown only a fraction of a decibel is lost in quantizing speech signals with the embedded quantizer compared to an unconstrained quantizer design. In terms of subjective performance, the embedded design was not found to be significantly different from the nonembedded design, even after four transcodings.

A significant systems advantage resulting from the embedded coder design is that the encoder may operate without regard to the momentary data transmission requirements. The speech coding and data multiplexing operations are separated logically and possibly even physically. Thus, data may be introduced at a point downstream in the transmission path removed from the encoding terminal. The receiver or decoding terminal

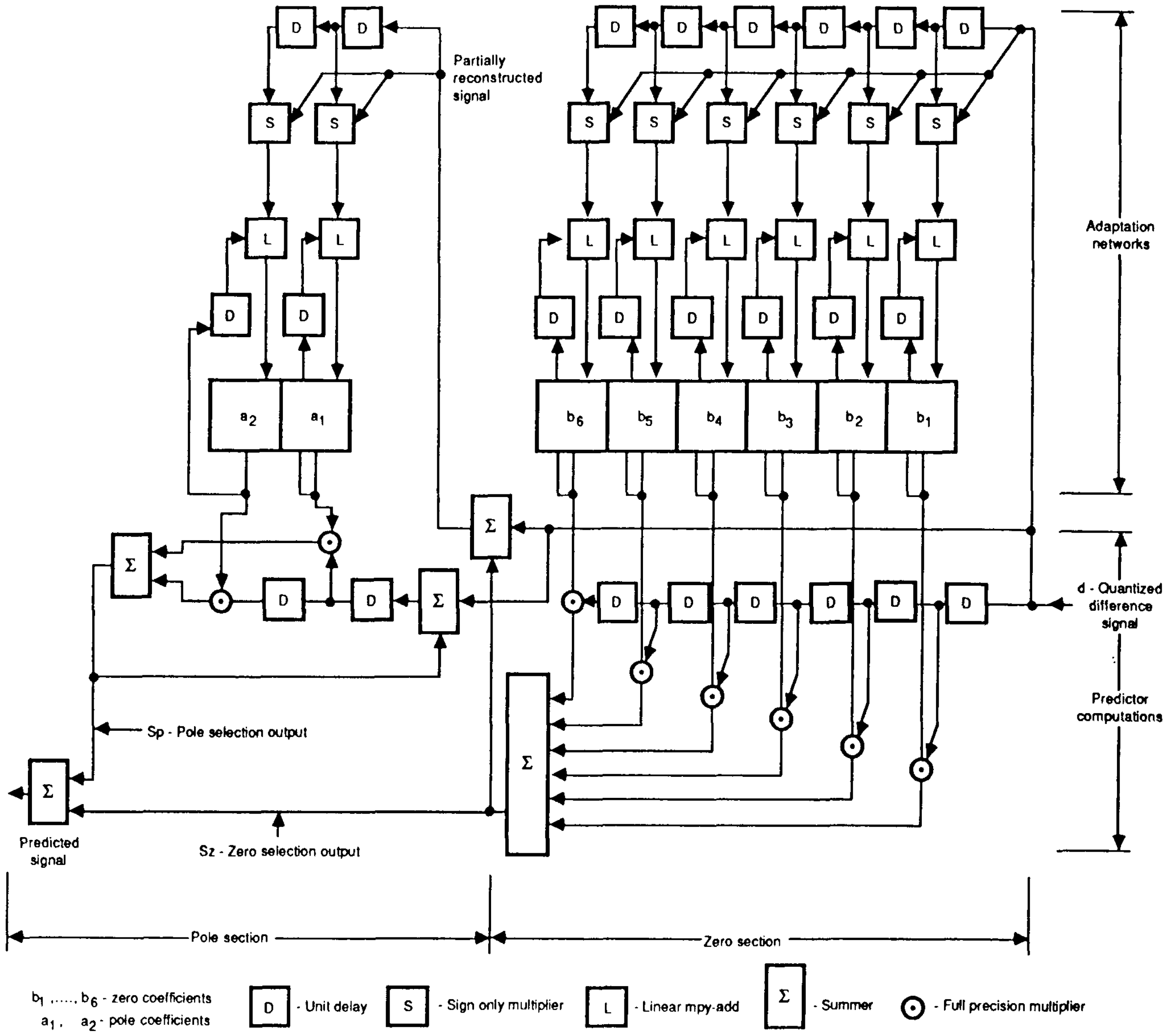

Fig. 5. Signal Flow Diagram for the Adaptive Predictors of each Subband.

must, of course, be aware of the amount of data introduced in place of speech information so that it may interpret the respective bits appropriately.

A more detailed diagram of the subband-ADPCM encoder and decoder is given in Fig. 4. As illustrated, identical encoding by the transmitter and receiver in each of the three modes of low-band quantization is ensured by stripping the two least significant bits in the feedback paths of the predictor and quantizer. The decoder interprets the received data in accordance with the current mode indication. The 6-, 5-, or 4-bit words are converted using 60-, 30-, or 15-level inverse quantizer tables to arrive at the appropriate signal estimate.

The two-pole, six-zero adaptive predictor data-flow structure is illustrated in greater detail in Fig. 5. Alternative implementations are, of course, available; the figure shows a flowchart for perhaps the most simple configuration.

Subjective Performance

Before selecting the final design, the Expert Group carried out a series of subjective experiments with different speech signals (diverse languages), music, differing transmission conditions, and various hardware coding devices. Starting with four different algorithms, as implemented in hardware, the list was pruned to two, and, finally, one compromise design. The discussion on performance that follows pertains only to the final design as embodied in the recommendation.

The subjective measure of audio quality adopted is the mean opinion score (MOS) on a five-point scale: excellent, good, fair, poor, bad. Since in the anticipated applications of audio conferencing and hands-free telephony listening over speakers and not handsets is most likely, audio was provided over loudspeakers at a level of 70-dB SPL (sound pressure level) at the listener. Seven different language tapes were processed by the same codec hardware. Listening experiments were conducted in seven laboratories; the results quoted are the average results obtained.

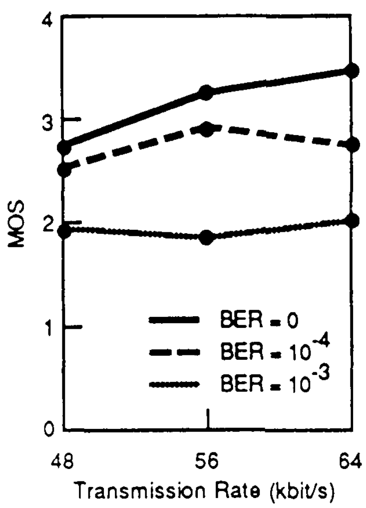

Fig. 6. Average Subjective Quality Ratings of the Final Algorithm.

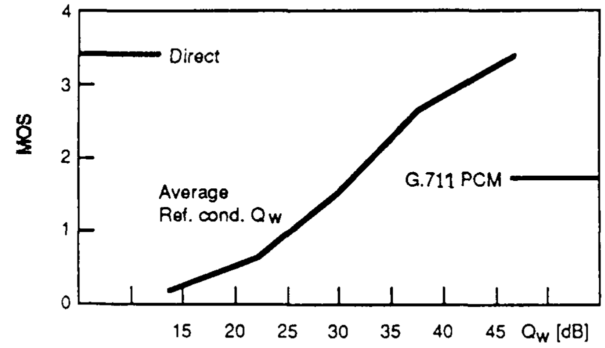

Fig. 7. Average Subjective Quality Ratings for Multiplicative Noise Reference Conditions.

Audio quality is best at 64 kbit/s, drops slightly when the transmission rate is reduced to 56 kbit/s, and more significantly with a further reduction to 48 kbit/s (Fig. 6). However, even at 48 kbit/s, the audio quality is significantly better than narrowband PCM. Received audio quality is only slightly affected by transmission bit error rates up to 10^{-4}, but a significant quality drop is noted when the error rate reaches 10^{-3}.

The subjective MOS may show differences in speech quality due to speaker and listener effects as well as the language used. To allow different experimental conditions to be compared more precisely, a set of reference conditions of speech mixed with multiplicative white noise was evaluated by each group of listeners evaluating the coded speech. In each case, Qw indicates the signal-to-noise ratio in decibels. Figure 7 gives MOS scores as a function of Qw. The best mean MOS score of 3.3, obtained under error-free 64-kbit/s transmission, corresponds to a reference condition of $Qw = 45$ dB. Note that the direct uncoded speech is assigned the same MOS rating, implying no measurable quality degradation due to one stage of wideband coding. In contrast, narrowband PCM (G.712) results in a Qw of 32 dB. Thus, the overall subjective gain for G.722 coded speech relative to G.712 coded speech is equivalent to some 13 dB of noise reduction. Wideband signals coded at 128 kbit/s, 8 bits/sample at a sampling rate of 16 kHz, are found to correspond in quality to a Qw of 38 dB. This finding suggests that roughly 6 dB of quality improvement results from expanding the signal bandwidth to the range of 50-7,000 Hz, and 7 dB of additional improvement is obtained by more precise encoding of that signal.

As the available transmission rate is reduced, the quality becomes slightly degraded. At 56 kbit/s, we observe a Qw of 43 dB, a very minimal degradation. At 48 kbit/s, Qw is 38 dB, which is still significantly better than the quality of G.712. Whenever data transmission is intermittent, conference participants may not even be aware of the audio-quality variations due to dynamic mode switching.

Mode Initialization and Mode Switching

Although the procedures for mode initialization and mode switching will be incorporated into a separate recommendation, they are discussed here to indicate how the G.722 coding procedures may be applied in practical communication systems.

To avoid the need for multiple audio terminals on one's desk—a high-quality wideband terminal to communicate with other wideband terminals and a normal telephone to reach parties equipped only with narrowband telephones—most wideband terminals will incorporate a narrowband PCM communication mode. Calls can then be set up with the terminal in the narrowband mode (Mode 0). As soon as the called party answers, an exchange of flags takes place. The calling terminal transmits one of two flags denoting its capabilities. Terminals may be of Type 1, which has only a 64-kbit/s transmission capability, narrowband or wideband, or Type 2, which implements at least 64- and 56-kbit/s wideband modes (Modes 1 and 2) and possibly even 48 kbit/s (Mode 3). The flag sent identifies the terminal type; the calling terminal awaits a similar response flag. Dumb terminals having only a narrowband capability cannot respond to the received flag, which leads the calling terminal to conclude, after a suitable time-out, that it must remain in the narrowband mode. Intelligent terminals acknowledge the received flag by transmitting the appropriate response flag and adopt a wideband mode (Mode 1). On receipt of that flag, the calling terminal assumes the same wideband mode. If the exchange of flags takes place on the least significant bit or bit 8 of every word, it introduces only a slight amount of noise into the audio path. Thus, narrowband communication is available even during the flag exchange. This plan allows wideband terminals to be introduced gradually into the telephone network, retaining connectivity and gradually enhancing quality with increased penetration of wideband terminals.

Many North American network connections optionally modify the least significant bit of PCM words to transmit signaling information. To allow the use of wideband terminals in such situations, the initial wideband mode could be Mode 2. It has been suggested that both bits 7 and 8 be employed here for reliable flag transmission. Corruption of the flag on bit 8 would then indicate to the terminals that their transmission channel is limited to 56 kbit/s.

Data-Speech Multiplexing

In many conference situations, it is desirable to transmit conference-related data, such as speaker-identification information or document facsimile, on the established connection without interrupting the speech communication path. The fact that high-quality audio transmission can be maintained down to 48 kbit/s with but minor degradation allows up to 16 kbit/s of data transmission.

Type 2 terminals may switch from Mode 0 (64-kbit/s speech, no data) to Mode 1 (56-kbit/s speech, 8-kbit/s service channel) by exchanging flags. It is envisaged that this service channel will provide terminal-to-terminal signaling and incorporate a frame alignment signal (FAS) of 8 bits/80 octet frame, a bit-rate allocation signal (BAS) of 8 bits/80 octet frame, and up to 6.4 kbit/s of data. Under the control of the BAS, the service channel may be expanded in increments of 8 kbit/s by stealing additional bits from the speech channel.

If more than 16 kbit/s are to be devoted to data transmission, the benefits of wideband audio become marginal. Additional modes of audio transmission, e.g., narrowband audio within 32 kbit/s (possibly G.721) or even 16 kbit/s, can be readily defined and would allow the capacity of the service channel to increase to 32 and 48 kbit/s, respectively. Speech transmission, although still quite intelligible, would be of lesser quality.

Communication between Narrowband and Wideband Terminals

To allow audio conferences between participants, some having wideband terminals, others narrowband terminals, conversion of narrowband signals to wideband representation and wideband signals to narrowband representation is required. Narrowband terminals cannot reproduce the high-band components of the wideband signal; this permits their derivation from only the low-band component by filtering and PCM coding. However, generating a wideband signal using a QMF synthesis filter with only low-band and no high-band input results in audible high-frequency distortion due to the uncanceled aliasing product. It is preferable, there-

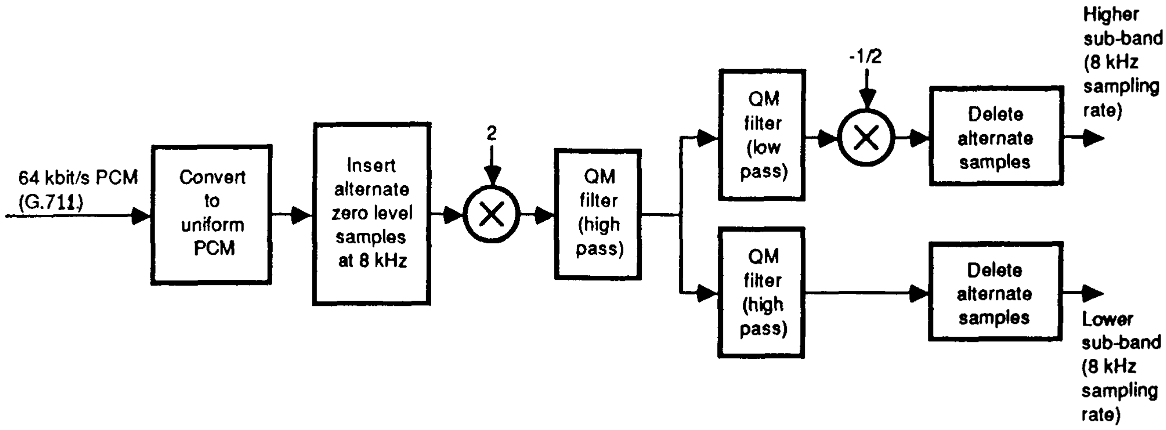

Fig. 8. Generation of Low- and High-Band Subband Signals from a Narrowband PCM Signal.

fore, to generate a pseudo-high-band signal from the narrowband input and use it later to cancel the aliasing products of the low-band signal.

Two alternative procedures suggest themselves for the narrowband/wideband conversion. The first and more straightforward is to upsample the uniform PCM representation to 16 kHz, lowpass the result to eliminate the 4-8-kHz aliasing component, and then wideband encode the result as if it were a normal wideband signal. A second procedure is simpler and avoids the need for a new low-pass filter design. As shown in Fig. 8, it generates a lower subband signal by a series of two QMF high-pass operations on the aliased narrowband signal. It also generates an artificial high-band signal by a series of high- and low-pass operations on the same aliased signal. When the two subband components are passed through QMF synthesis in a wideband terminal, a narrowband signal is heard with no additional noise.

Conference bridges combine several input signals and may transmit different output signals, depending on whether the particular port is considered active (currently speaking) or silent. For wideband audio bridges, it appears preferable to combine the low- and high-band components of the input signals from the several ports separately, as this avoids the accumulation of delays due to QMF analysis and synthesis at conference bridges. To achieve best quality when mixing narrowband and wideband inputs, narrowband inputs should first be converted to wideband form. The all-wideband bridge may then employ signal combination logic analogous to that found in narrowband bridges, but implement it separately for the low- and high-band signal components.

Concluding Remarks

The new wideband coding standard represents an important advance in two aspects: first, it improves the quality of audio on the telephone network; second, it provides for an audio-associated data channel to carry conversation- or conference-related data. Its deployment is currently limited to locations accessible by 56- or 64-kbit/s digital loops. However, since there is much current international interest in digital networks such as ISDN, the penetration of end-to-end digital connectivity will probably increase rapidly, and the cost of digital transmission will decrease simultaneously. The adoption of the new standard is timely and should not only prevent the proliferation of incompatible coding techniques, but also help in making cost-effective premium-quality audio terminals rapidly available.

References

[1] W. R. Daumer, P. Mermelstein, X. Maitre, and I. Tokizawa, "Overview of the ADPCM coding algorithm," *Conf. Record Globecom '84*, Atlanta, GA, pp. 23.1.1-23.1.4, 1984.

[2] E. Hummel, "The CCITT," *IEEE Commun. Mag.*, vol. 23, no. 1, pp. 8-11, 1985.

[3] IEEE Global Telecommunications Conference—Globecom '86, Houston, TX, Papers 17.1 through 17.5, 1986.

[4] T. Nishitani, I. Kuroda, M. Satoh, T. Katoh, and Y. Aoki, "A CCITT standard 32 kb/s ADPCM LSI codec," *IEEE Trans. Acoust. Speech Sig. Proc.*, vol. ASSP-35, pp. 219-225, 1987.

[5] R. E. Crochiere and J. Flanagan, "Current perspectives in digital speech," *IEEE Commun. Mag.*, vol. 21, no. 1, pp. 32-40, 1983.

[6] G. Williams and H. Suyderhoud, "Subjective performance evaluation of the 32 kb/s ADPCM algorithm," in *Proc. IEEE Global Telecommun. Conf.—Globecom '84*, Atlanta, GA, pp. 778-785, 1984.

[7] R. E. Crochiere, S. A. Webber, and J. L. Flanagan, "Digital coding of speech in sub-bands," *Bell Syst. Tech. J.*, pp. 1069-1085, 1976.

[8] N. S. Jayant and P. Noll, *Digital Coding of Waveforms*. Englewood Cliffs, NJ: Prentice-Hall, p. 306, 1984.

[9] D. Millar and P. Mermelstein, "Prevention of predictor mistracking in ADPCM coders," *Proc. IEEE Int. Conf. Commun.—ICC'84*, Amsterdam, pp. 1508-1512, 1984.

[10] D. J. Goodman, "Embedded DPCM for Variable Bit Rate Transmission," *IEEE Trans. Commun.*, vol. COM-28, pp. 1040-1046, 1980.

Paul Mermelstein was born in Czechoslovakia in 1939. He received the B. Eng. degree in engineering physics from McGill University, Montreal, Canada in 1959, S.M., E.E. and D. Sc. degrees from the Massachusetts Institute of Technology, Cambridge, in 1960, 1963 and 1964, respectively. From 1964 to 1973 he was a member of the technical staff in the Speech and Communications Research Department of Bell Laboratories, Murray Hill, New Jersey. From 1973 to 1977 he was a member of the research staff of Haskins Laboratories conducting research in speech analysis, perception and recognition. Over the years 1977-1986 he has been manager of Speech Communication Systems at Bell-Northern Research in Montreal. He currently serves as Manager of Man-Machine Systems at BNR. Paul is editor for Speech Communication of the *IEEE Transactions on Communications*. He served as associate editor for Speech Processing for the Journal of the Acoustical Society of America in 1983-87. He has also served on numerous expert groups on speech coding of CCITT SGXVIII recommending standards on speech coding for telecommunication applications. Since 1977 he has held appointments as Visiting Professor at INRS-Telecommunications (Université de Québec) and Auxiliary Professor of Electrical Engineering at McGill University.