Speech Coders: From Idea to Product

Richard V. Cox
Peter Kroon
Juin-Hwey Chen
Reed Thorkildsen
Kevin M. O'Dell
David S. Isenberg

Speech coders are used to transmit and store speech in digital form for various applications. This paper describes how to set requirements for given applications, the interactions between research and business units, and the role that standards bodies play. Creating a coder that meets the requirements of its customers and dealing with the issues involved in implementing the coder are an integral part of the development process. After the development is complete, the speech coder must be integrated into a product. As an example, the design and integration of the AUDIX® voice messaging coder (VMC) are described. The future holds ever-increasing challenges and new technology in which speech coders will play an integral role.

Introduction

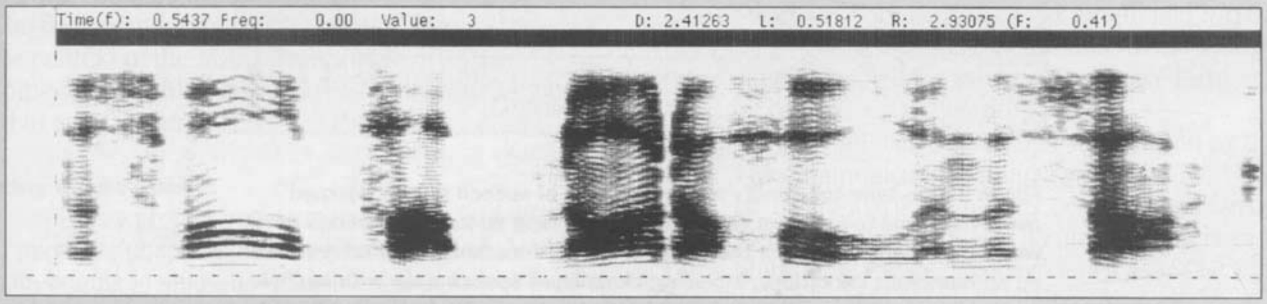
A speech coder is a device that compresses and decompresses the digital information necessary to represent a speech waveform.[1] As part of a larger system, speech coders are used in transmission or storage applications. For example, in digital cellular service, the speech is encoded into a low-rate bit stream that is transmitted over a digital radio channel and then decoded to reproduce the speech waveform. Other examples of transmission applications include network telephony, secure voice, and audio and video teleconferencing. In each of these, a live conversation is taking place between two or more people. The AUDIX® voice messaging system is an example of a storage application. The voice message is compressed and the bit stream is stored for later retrieval. Other storage applications include broadcasts, network announcements, and digital telephone answering machines. These applications involve only one-way communication.

Speech coders compress signals by exploiting the natural redundancies in speech and the properties of human hearing. (See Panel 1 for an explanation of some of the basic properties of human speech that allow coders to represent the acoustic properties of these signals efficiently.) The compression techniques used in speech coders are all known as lossy compression, because the reproduced signal is not identical to the original. Decoded speech sounds like the original because of a masking property of the human ear that renders a certain amount of noise inaudible.

Speech coders can be evaluated on the basis of four attributes—bit rate, complexity, delay, and quality—the last of which is a function of the first three. As a rule, the quality will improve if any or all of the first three attributes increase in value.

Bit rate reflects the degree of compression that the coder achieves. Telephone bandwidth speech (i.e., 2,000 to 3,400 hertz [Hz]) is sampled 8,000 times per second (8 kilohertz [kHz]), and quantized with an 8-bit logarithmic quantizer, making the normal bit rate used for transmission in telephone networks 64 kilobits per second (kbits/s). Two different standard coders are commonly used. North America and Japan use μ-law pulse-code modulation (PCM), and the rest of the world uses A-law PCM. (See Panel 2 for definitions of abbreviations, acronyms, and

(a)



(b)

Panel 1. Viewing Speech Using Speech Spectrograms

Figure a is a wideband speech spectrogram, and Figure b its narrowband counterpart. Both spectrograms are for the sentence, "The small pup gnawed a hole in the sock," spoken by a male. A spectrogram is a two-dimensional plot showing the frequency content of a signal as it evolved over time. Narrowband spectrograms have high resolution in the frequency domain and low resolution in the time domain. Conversely, wideband spectrograms have high resolution in time and low resolution in frequency. The time axis is from left to right, and the vertical axis represents the frequency, with low frequencies at the bottom. The sound energy at a given frequency is represented by the shading, the darker shading indicating more energy.

As we speak, we move our lips and tongue, which change the resonant frequencies, or *formants*, of our vocal tracts. Figure a is made up of vertical lines with varying amounts of darkness. Each vertical line corresponds to a single pitch period of speech, illustrating the high time resolution of wideband spectrograms. The dark areas along the lines indicate vocal tract resonances corresponding to the formant frequencies. These resonances account for much of the intelligibility in speech. They are nicely modeled by an all-pole filter, as is used in LPC. The dark horizontal bands in Figure b correspond to pitch harmonics for voiced speech. Because of its high resolution in frequency, the narrowband spectrogram can capture each individual pitch harmonic. Both the pitch and the formant information change slowly. The information describing them can be efficiently represented with a relatively slow update rate.

terms.) The degree of compression is measured by how much the bit rate is lowered from 64 kbits/s. International standards exist for coders operating at 40, 32, 24, and 16 kbits/s, and are planned for rates as low as 4 kbits/s. Regional cellular standards span the range from 13 to 3.45 kbits/s. Secure voice coders operate at 4.8, 2.4, and 0.8 kbits/s, and are planned for even lower rates.

Complexity is another important attribute. In general, the lower the rate of a coder that can still maintain speech quality, the higher its complexity. Complexity affects a speech coder's power consumption and cost. Cost is almost always a factor in selecting a speech coder for a given application, and power consumption has become more critical as the importance of wireless and portable communications has increased.

Complexity typically has three components:

- The number of instructions per second that must be executed to operate the coder in real time, which is a measure of the required processor speed. The usual unit of measurement is millions of instructions per second (MIPS). Generally, higher-speed processors cost more. In addition, power consumption is usually

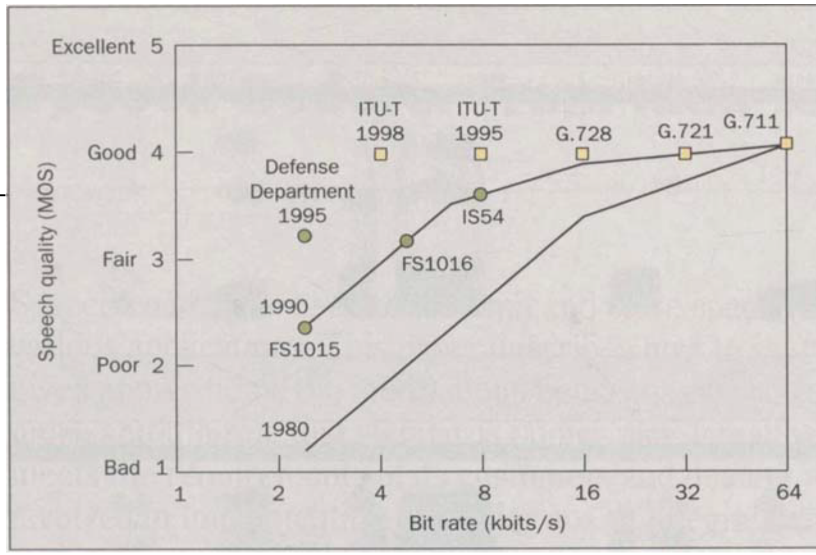


Figure 1. The term toll quality refers to a level of speech quality deemed acceptable end-to-end for a call carried over a long distance telephone network. Coders submitted for testing by the ITU-T for standardization are tested for numerous conditions, including clean input speech even with multiple encodings, input levels that are too high or too low, noisy background, and noisy channel conditions. Only when acceptable performance was achieved for all these conditions were the coders standardized and considered of toll quality. This figure shows the results of subjective testing for clean input conditions for a single encoding. The curves for 1980 and 1990 show the results of subjective tests conducted during those years by AT&T Bell Laboratories. The scores are for the best coders at the given bit rates. The boxes indicate scores for current ITU-T toll network coders (G.711, G.721, and G.728) or for planned new ITU-T standards. The circles indicate other current or planned future standards of other standards bodies. Although dramatic improvements in quality have been made in the past 15 years, improvement is still needed at the lower rates.

linearly related to the clock rate of the processor.
- The amount of random access memory (RAM) needed to store the variables used in the algorithm.
- The amount of memory required for storing the instruction sequence and constant values used in the algorithm.

Read-only memory (ROM) is often used for this purpose. Speech coders are most often implemented on very large-scale integration (VLSI) circuit chips, either custom chips or programmable digital signal processor (DSP) chips. Either way, MIPS, RAM, and ROM determine the physical size, speed, and power needed to implement the coder. These attributes, in turn, determine the cost.

The communication *delay* of the coder is more important for transmission than for storage applications. In two-way conversations, a large communication delay can impose an awkward protocol on talkers. Communication delays of 300 milliseconds (ms) or more are particularly objectionable to users. Multiparty conversations are carried on through a bridging site, in which the bit streams of all parties are decoded, summed, and then re-encoded for transmission. This doubles the delay because two complete encodings and decodings are taking place. By contrast, in a storage application, a delay of one second would be unnoticeable to the user.

The attribute of *quality* has many dimensions. Ultimately, quality is determined by how the speech sounds to a listener. Some factors that affect the performance of a coder are whether input speech is clean or noisy, whether the bit stream has been corrupted by errors, and whether multiple encodings have taken place. When a coder performs well under all these circumstances, it is considered robust. Figure 1 illustrates the progress that has been made toward achieving better speech quality during the past 15 years.

This paper also describes:
- How to set requirements for given applications, including the interactions between research and business units,
- The role of standards bodies in this process,

- How to create a coder that meets the requirements,
- The issues involved in implementing the coder,
- The systems integration process of creating the product that contains the speech coder,
- The design and integration process, demonstrated in the context of the AUDIX VMC, and
- Some future challenges and the new technology needed to meet them.

Setting Requirements

Every project begins with an idea. Initially, the idea may take shape in just one person's mind. As that person begins to share it with others, the project develops and takes on a more substantial form. For the purposes of this paper, this project is assumed to include a speech coder. The genesis for the project could have begun in a business unit or a research area. In the best circumstances, research and the business unit will form a partnership and make it a joint project. This partnership begins with the task of defining the requirements of a speech coder to meet the needs of a business unit.

The description of the four attributes—bit rate, quality, delay, and complexity—indicates that there are many tradeoffs in selecting a speech coder for a particular application. To illustrate, this paper will highlight two particular applications, digital cellular systems and the AUDIX voice messaging system.

Digital cellular systems transmit speech over radio channels; digital speech compression enables cellular service providers to increase system capacity and robustness simultaneously. Unfortunately, radio channels are subject to both interference and fading. Interference can cause random errors in the bit stream, and fading can either cause a burst of errors or wipe out the bit stream entirely. To avoid these problems, it is essential to transmit the bit stream with some form of error protection. The system design choice is complicated, but a system must balance capacity, speech quality, cost, and power consumption. For example, as the percentage of channel capacity (transmitted bits) used for error protection increases, the number of bits available for the speech coder decreases. Although this creates a more robust system, it also results in lower quality, higher complexity, or even higher delay. But it is not always necessary to shift the entire burden to the speech coder. A more efficient modulation scheme or a better antenna design could make the requirements for the speech coder less stringent.

Voice messaging has a different set of requirements. Lowering the bit rate does increase system capacity, but a tradeoff exists between the coder cost and the cost of memory used for voice storage. While

Panel 2. Abbreviations, Acronyms, and Terms

ADPCM—adaptive differential pulse-code modulation. A coding technique applied within the network for 32 kbits/s (ITU-T Recommendation G.726).
CELP—code-excited linear prediction. A method of speech coding that combines linear predictive coding with vector quantization of the excitation signal. CELP is used in many speech coders that operate between 4 and 16 kbits/s.
DSP—digital signal processor
FS1015—Federal Standard 1015 is a 2.4-kbits/s LPC speech coder specified for use in secure voice terminals by both the U.S. government and NATO.
FS1016—Federal Standard 1016 is a 4.8-kbits/s CELP coder specified for use in secure voice terminals by the U.S. government.
IS54-TDMA—Interim Standard 54 was created by the Telecommunications Industry Association (TIA) as the standard for the U.S. cellular Time-Division Multiple Access System. Included in IS54 is an 8-kbits/s CELP speech coder known as VSELP.
ISDN—Integrated Services Digital Network
ITU—International Telecommunications Union, the body that sets international telephony standards. The ITU Telecommunications Standardization Sector (ITU-T) is the portion of the ITU that has jurisdiction over speech coding standards. It was formerly known as the CCITT.
LD-CELP—low-delay CELP. A CELP coder using backward adaptive prediction to reduce delay, used for ITU-T Recommendation G.728.
LPC—linear predictive coding. This refers to the method used to remove correlations in the speech signal.
MIPS—millions of instructions per second
PBX—private branch exchange
PCM—pulse-code modulation, another term for the digitization of speech. For telephone bandwidth signals, it is usually combined with 8-bit logarithmic quantizers to achieve 64 kbits/s (ITU-T Recommendation G.711).
PSTN—Public Switched Telephone Network
RAM—random access memory
ROM—read-only memory
VLSI—very large-scale integration
VMC—voice messaging coder

the priority of delay ranks low, that of speech quality is high, because customers want the voice quality of wireline telephone service.

Most cellular systems are part of a national or global communications system, making it essential to define standards to enable the equipment of various manufacturers to operate together and to interface with wireline systems. These standards are usually drafted by regional standards bodies, which are formed by members of different interested parties, such as manufacturers, service providers, and customers (see Reference 2 for an overview of this process). As a result, the requirements for a speech coder used in cellular applications are defined in a far more formal fashion than proprietary products such as the AUDIX voice messaging system.

Creating a Coder

This paper has discussed the attributes of speech coders and how these attributes determine requirements for new coders. But designing a speech coding algorithm requires more than just combining blocks that typically constitute a speech coder. Once the industry acknowledges that a certain technique can meet the requirements, it is up to the coder designer to combine the proper techniques. A given set of coder requirements for delay and complexity will limit the number of possible choices for certain parameters. Within these constraints, the insights and experience of the algorithm designer are combined to pinpoint the optimal choices. Often, many versions with varying parameter choices need to be evaluated to find the one that performs best. The ever-increasing availability of standard coders for different bit rates sometimes makes it possible to use or modify a standard coder. However, unique requirements or patent and royalty issues may justify the design of a completely new coder. This happens more often when the requirements are pushing the boundaries of the current state of the art. In this case, basic research results of many people are combined into innovative algorithms, which perform better than any of the known coding techniques. Obviously, such an approach is not always possible and requires more time than traditional, obvious ways.

Building a Prototype

Once a speech coding algorithm has been selected, the coder must be implemented in real time using either a floating-point or fixed-point VLSI device. The implementation step may either follow the previous step, coder creation, or it might be done in parallel. The latter approach reduces the time needed to deliver a final product. It also allows the hardware implementor to talk to the algorithm designer, usually about complexity issues, such as procedures to reduce the number of operations to perform a certain algorithmic step, or to find ways to reduce the amount of memory needed. For a speech coder that uses well-established techniques, such an approach works, because many of its components are well understood and can be implemented independently. For a speech coder that uses a completely new approach, however, a parallel approach might not always be suitable.

Floating-point DSPs have roughly the same numerical precision and dynamic range as single-precision floating-point high-level simulation code being run on computers. A floating-point high-level simulation code can generally be converted to a floating-point DSP implementation with little trouble. More difficulties arise when a coder must be implemented using 16-bit fixed-point arithmetic. The limited dynamic range and accuracy of fixed-point devices often cause overflows, underflows, and other numerical problems that degrade the quality of the coder's output speech. Without a well-defined procedure to transform a floating-point simulation to a 16-bit fixed-point DSP implementation, the outcome often depends on the skills and insights of the DSP programmer. In most cases, however, the fixed-point implementation of a speech coder, if properly done, can match the speech quality produced by its floating-point counterpart. It is not always clear if the resulting implementation is the most efficient in terms of memory usage, development time, and the number of operations. To improve this whole process, research is under way to make the conversion from floating-point algorithms to fixed-point implementations more streamlined and efficient.

As soon as a real-time implementation is available and its correct operation has been verified, more exhaustive subjective testing can be performed. Preferably, these tests should be done in an environment that mimics the intended application of the speech coder. By processing more speech material through the coder, potential problems can be discovered and fixed. It also allows certain coder parameters to be "tuned" to the types of input speech. For example, if only certain types of microphones are used with the coder, one could adjust the coder parameters to obtain the best possible speech quality for that type of microphone. Usually, it takes a cycle of tests to pinpoint a coder that performs optimally for the intended application.

Creating the Product

Typically, the prototype speech coder is implemented using a development system for the chosen DSP. In the final product, the speech coder is included on a specific piece of hardware, commonly referred to as a platform. The first task is to port the speech coder so it can operate on its target platform. Because the speech coder is only a portion of the system, the next task is to integrate the coder into the rest of the system.

Once the whole system has been integrated and tested, its final design can be frozen, and the process of making one or more products based on this platform can begin. For example, AUDIX software has now been ported to the Intuity™ voice messaging system, which uses industry-standard Intel-based hardware with add-on voice port cards. (Intel is a registered trademark of Intel Corporation.) These cards contain the DSPs and telephone line interfaces. For digital cellular applications, speech coders are used in both the terminal and cell site equipment. The basic speech coder implementation on a fixed-point DSP chip can be used in both. In fact, AT&T Microelectronics sells chip sets that combine the speech and channel coders for use in the cellular terminals of other manufacturers' equipment.

Case Study of the AUDIX VMC

In 1989 the AUDIX voice messaging system was a successful product line that used a 16-kbits/s sub-band speech coder implemented on a WE® DSP20.[3] The speech coder had originally been created and implemented on a WE DSP1 by AT&T researchers in 1980. The availability of newer, faster DSPs, such as the AT&T DSP32C, presented an opportunity to create a significantly cost-reduced product (the DEFINITY® AUDIX voice mail system) as an embedded circuit pack within the private branch exchange (PBX) of the DEFINITY telecommunications system. AUDIX managers were aware that increased processing power, combined with more advanced speech coding techniques, could achieve toll-quality voice compression at 16 kbits/s. When this work was undertaken, AT&T had just submitted low-delay code-excited linear prediction (LD-CELP)[4] for ITU-T standardization as a toll-quality 16-kbits/s coder. Although this coder was an obvious candidate, it was too complex to be the new AUDIX VMC, which had to be low enough in complexity to fit three complete encoder and decoder pairs onto a single DSP32C. This represented about one-fifth the complexity of LD-CELP. Therefore, the goal for the new AUDIX VMC was to match the quality of the LD-CELP, but reduce the complexity. In addition, the AUDIX VMC required a decoder that would support speed control (i.e., either speeding up or slowing down a message without changing the pitch of the speaker). Thus, the decoder had to be less than half as complex as the encoder, so it could support real-time playback when running at double speed.

Speech quality, which has many dimensions, needs further specification. The bit streams of messaging systems like the AUDIX VMC are not generally corrupted by errors. Therefore, robustness to channel errors is not a factor in creating or selecting the coder. Neither was the VMC required to handle multiple talkers. Typically, multiple encodings will not be used with the same coder. However, for networking applications (e.g., forwarded voice mail), it is important to consider the performance of the VMC when its input is speech encoded with different speech coders in other vendors' voice messaging systems.

The AUDIX VMC has both rewind and fast forward functions. The decoder can restart the decoding process after each operation from any frame in the compressed bit stream without adding obvious distortion. For this reason, when the decoder is started with internal states that differ from those in the encoder, the decoder states must be able to converge quickly with the encoder states.

Because voice messaging is a storage application, the final attribute, delay, did not have a requirement. It became a free parameter that the algorithm creator could use either to enhance quality or reduce complexity.

The VMC was created using 16-kbits/s LD-CELP as a starting point. The LD-CELP coder needed a high level of complexity to achieve high speech quality with a very low delay (a buffering delay of 0.625 ms). Once the delay constraint was relaxed, its complexity could be greatly reduced. The 50th-order backward-adaptive linear predictive coding (LPC) was replaced by a forward-adaptive 10th-order LPC. Its coefficients were updated only once every 24 ms, greatly reducing the complexity created by

LPC updates. A computationally efficient LPC quantization scheme was also developed to keep the complexity low.

The speech quality was degraded by reducing the LPC predictor order from 50 to 10, making it necessary to include pitch prediction to improve the quality. Because the closed-loop pitch search commonly used in CELP coders has a high level of complexity, a more efficient, open-loop pitch-search method based on 4:1 decimation was used. The pitch predictor has three coefficients that were jointly quantized using an efficient codebook search method.

The excitation quantization used a slightly larger vector dimension and half the codebook size of LD-CELP, which cut the search complexity significantly. To maintain good speech quality and reduce complexity, a simplified form of backward-adaptive excitation gain was retained. An analysis of complexity showed that the resulting VMC algorithm had the potential to fit three encoders and three decoders onto a single DSP32C.

The first version of the VMC algorithm was now ready. In subjective tests, its quality was equivalent to or slightly better than the 16-kbits/s LD-CELP and the 32-kbits/s adaptive differential pulse-code modulation (ADPCM) standard. Encouraged by this, AUDIX managers decided to proceed with the DSP32C real-time implementation. While implementing the algorithm on a DSP32C, however, engineers found that it took slightly more than one DSP32C to implement three pairs of VMC encoders and decoders. When this problem was reported, to lower the complexity even further the AT&T researcher who created the VMC algorithm reduced the codebook size and vector dimension for the excitation. This change made it possible to implement an encoder and decoder in only 3.9 MIPS on the DSP32C. Thus, three complete encoders and decoders could comfortably fit onto a single 12.5 MIPS DSP32C, with some room left for implementing system functions.

When the real-time VMC was integrated into an AUDIX system and the fast forward and rewind functions were tested, system engineers found that the VMC's convergence time was longer than desired. AT&T researchers and AUDIX engineers quickly fixed the problem by reducing the coefficient of the gain predictor to allow more leakage. This completed the changes to the VMC. Additional subjective tests showed that the final version of the VMC maintained the speech quality of the first version. Thus, the close collaboration between AUDIX engineers and AT&T research led to a VMC that not only met the complexity requirement, but also slightly exceeded the expected speech quality.

In the case of the DEFINITY AUDIX system, the speech coder needed to be integrated with additional DSP software (for call-progress tone detection/generation, automatic gain control, and silence detection). Because multiple channels had to run in a single DSP, a small real-time kernel was developed to allow switching between voice channels. Most of the application software runs on a host processor of the UNIX operating system, so communications software between the DSP and host processors needed to be developed. (UNIX is a registered trademark of Novell in the United States and other countries, licensed exclusively through X/Open Company Limited.) The host processor is responsible for storing compressed voice files, for the overall "intelligence" of the voice mail system, and for the call processing software.

Another consideration in creating the product was the need to define a storage format. In digital cellular systems, the decoder operates sequentially on the bit stream. In a voice mail operation, the user may skip forward or backward, or concatenate messages while forwarding. To reduce the necessary storage and incorporate time/date information, announcements are generally stored as separate fragments that are concatenated on playback. Rapid synchronization was also necessary. The basic unit of the sub-band encoding used in the original AUDIX system was a two-byte frame, making synchronization easy to maintain. In contrast, the VMC used a 48-byte frame. If the decoder is not correctly frame-aligned, the speech will be garbled. Developing the synchronization protocol was a major effort in the process of moving voice message coding into the DEFINITY AUDIX product.

The result of this strong collaboration between AT&T research and AUDIX engineers led to a significantly better product that successfully competes with other voice messaging products.

Challenges in Speech Coding

Speech coding is an area in which discoveries continue to be made and new applications continue to emerge. This section reviews several new applications and examines the challenges that their requirements impose.

One new application is video telephony. The high bandwidth needed for the video portion of this technology limits the bit rate available for speech, whether digital connections (Integrated Services Digital Network [ISDN]) or analog connections on the Public Switched Telephone Network (PSTN) are being used. The ITU-T is trying to standardize a telephone bandwidth speech coder that will be able to carry PSTN video telephony at 6.3 and 5.3 kbits/s. There is also a growing interest in using wideband speech for videoconferencing. Customers find that 7-kHz bandwidth speech is far more pleasant and less fatiguing to listen to than telephone bandwidth speech. Currently, the only standardized 7-kHz bandwidth speech coder operates at 48, 56, or 64 kbits/s. The ITU-T will soon attempt to standardize a 7-kHz bandwidth speech coder with a rate as low as 16 kbits/s. Because conference calls are involved, the challenge for the speech coder's creators is to maintain speech quality even with multiple encodings.

Another interesting application is a high-quality, low-bit-rate (2.4 kbits/s) speech coder for secure phones. Traditionally, 2.4 to 4.8 kbits/s have been used for secure phones, which consist of a speech coder, a modem, and digital encryption. Many of the government's secure calls are routed over satellite channels, whose capacity can only be increased by reducing the speech coder's bit rate. The government's radio channels can have a high bit-error rate, but the total bit rate of 2.4 kbits/s makes it almost impossible to include additional error protection in the form of channel coding. Instead, the coder itself must be made robust enough to withstand occasional random errors. Another difficulty is noisy background environments, such as military vehicles, in which the speech coder must attempt to preserve all the intelligibility of the input signal.

The U.S. government and its contractors are trying to create a 2.4-kbits/s standard for secure voice whose speech quality will match or exceed that of its current 4.8-kbits/s CELP standard. AT&T is actively working to meet this challenge. Assuming that this new 2.4-kbits/s standard will meet all of the requirements, it could also be usable in satellite-based services for the civilian sector. With such a phone, an individual can send or receive calls from anywhere in the world. Another application is low-bit-rate voice messaging services over pagers.

These new challenges, and the applications described earlier, demonstrate the continued interest in advancing the science of speech coding. Each new application rekindles business interest in the technology and spurs new research. During the past few years, the number of speech coding applications has grown enormously, rapidly increasing the number of different speech coding standards. Researchers continue to enjoy new challenges. When fiber-optic communications were introduced, more than a decade ago, soothsayers predicted the imminent death of speech coding. As this paper has demonstrated, speech coding is more alive than ever, and interest in it continues to grow as the number of challenges increases.

References

1. A. Gersho, "Advances in speech and audio compression," *Proceedings of the IEEE*, June 1994, pp. 900-918.
2. S. Dimolitsas, "Standardizing speech coding technology for network applications," *IEEE Communications Magazine*, November 1993, pp. 26-33.
3. R. E. Crochiere, "Sub-band coding," *Bell System Technical Journal*, Vol. 60, No. 7, September 1981, pp.1633-1653.
4. J-H. Chen, R. V. Cox, Y-C. Lin, N. Jayant, and M. J. Melchner, "A low-delay CELP coder for the CCITT 16 kb/s speech coding standard," *IEEE Journal on Selected Areas in Communications*, June 1992, pp. 830-849.

(Manuscript approved January 1995)

Richard V. Cox *is head of the Speech Coding Research Department at AT&T Bell Laboratories in Murray Hill, New Jersey. He is responsible for speech coding, combined speech and channel coding for noisy channels, real-time implementation of signal processing algorithms, and speech coding standards. Mr. Cox joined AT&T in 1979, after receiving a B.S. from Rutgers University, New Brunswick, New Jersey, and an M.S. and Ph.D. from Princeton University, New Jersey, all in electrical engineering.*

Peter Kroon is a member of technical staff in the Speech Coding Research Department at AT&T Bell Laboratories in Murray Hill, New Jersey. He is working on speech coding algorithms and standards for cellular wireless communications. Mr. Kroon joined AT&T in 1985, after earning an M.S. and Ph.D. in electrical engineering from Delft University of Technology, The Netherlands.

Juin-Hwey Chen is a distinguished member of technical staff in the Speech Coding Research Department of AT&T Bell Laboratories in Murray Hill, New Jersey. He created the 16-kbits/s voice messaging coder and the ITU-T G.728 LD-CELP coding standard, and is currently working on wideband (7-kHz) speech coding. Mr. Chen joined AT&T in 1988, after receiving a B.S. in electrical engineering from National Taiwan University, Taipei, Taiwan, R.O.C., and an M.S. and Ph.D. in electrical engineering from the University of California, Santa Barbara.

Reed Thorkildsen is a technical manager in the Wireless Systems Core Technology Department at AT&T Network Wireless Systems in Whippany, New Jersey. He is responsible for signal compression technology in voice, audio, image, and video media in wireless applications. Mr. Thorkildsen received a Ph.D. in physics from the University of Virginia in Charlottesville. He joined AT&T in 1981.

Kevin M. O'Dell is a member of technical staff in the Voice Systems Department of AT&T Global Business Communications Systems in Denver, Colorado. He develops digital signal processing software (voice compression and call-progress tone detection) for the DEFINITY AUDIX voice mail system. Mr. O'Dell joined AT&T in 1985, after receiving a B.S. in mathematics and computer science and an M.S. in physics, all from the University of Utah in Salt Lake City.

David S. Isenberg is a member of technical staff in the Technology Management Division of AT&T Communications Services Group in Murray Hill, New Jersey, where he is working on strategic planning and future service concepts. Mr. Isenberg received a Ph.D. in biology from the California Institute of Technology, Pasadena. He joined AT&T in 1985.