The VideoPhone 2500—Video Telephony on the Public Switched Telephone Network

Scott H. Early
Andrew Kuzma
Eric Dorsey

The AT&T VideoPhone 2500 is the first video telephone to provide color, motion video over the analog loop on the public switched telephone network. The enabling technologies include video and audio compression, as well as a high-speed modem (modulator–demodulator). This paper describes the overall architecture of the VideoPhone 2500 and operation of individual subsystems, including design considerations. It also provides details about the market research and business decision process that led to AT&T's entry into the video-telephone market.

Introduction

The AT&T VideoPhone 2500, which was announced in early 1992, provides video telephony over the local analog loop connected to the public switched telephone network. This means the VideoPhone 2500 can be connected directly to the normal telephone line. The telephone terminal features an easy to use, intuitive, user interface. Market studies that suggest this telephone provides significantly enhanced communications at an affordable price.

Technologies that were critical to the design of the VideoPhone 2500 included video and audio compression, the high-speed modem, overall system architecture including protocols, the user interface, and physical design. In this paper, we describe each of these areas and provide a summary of the market research that led to the product definition and design. (Panel 1 defines terms and acronyms used in this paper.)

Market Research

In 1988, an AT&T Consumer Products team began a review of video telephony in response to the announcement of several competitive entries into the still-frame, video-telephone market. As a result of simultaneous reviews of the technical feasibility and market opportunity in this product category, AT&T Consumer Products decided to fund forward-looking work and move the project into its New Business Development organization.

Early market research provided direction for the team. Consumers said that they were not satisfied with the capabilities of the black and white, still-frame, video telephones. Instead, they wanted and expected color, motion video, and simultaneous voice.

In the summer of 1989, close work between project management and AT&T Consumer Products Labs in Indianapolis, Indiana, led to the first quantitative market-research effort. Consumers were asked to review several different video and audio combinations that were played over a product simulator. In addition, different industrial designs were tested for consumers' preferences.

The research results summarized in late 1989 were very positive. A business case team was then assembled to determine the value of AT&T's entry into the video-telephone market with a color, motion, video telephone that would operate over a single, analog telephone line. The senior management team of AT&T Consumer Products supported the business case team's recommendation for an independently funded project to develop AT&T's video-telephone product, and the project became effective July 1, 1990.

Prototypes were developed, tested internally, and again submitted to market research with consumers. This research reinforced the positive decision for AT&T to enter the video-communications business. AT&T was considered a natural provider of such a product. In addition, video telephony would both support and enhance the AT&T brand name.

Panel 1. Abbreviations, Acronyms, and Terms

A/D — analog to digital

CAMIL2 — core-access modem-interface logic, version 2; an AT&T application-specific integrated circuit

CCITT — International Telegraph and Telephone Consultative Committee (Geneva, Switzerland)

CELP — code-excited linear prediction

CELP+ — a code-excited linear prediction algorithm that operates at 6.8 kb/s and uses fractional-pitch prediction and constrained stochastic excitation

chrominance pixel — color pixel

codec — coder–decoder

D/A — digital to analog

DCT — discrete cosine transform

DPCM — digital pulse-coded modulation

DSP — digital signal processor

EMI — electromagnetic interference

ESD — electrostatic discharge

LAPB — link-access procedure-balanced; the CCITT-specified data-link-level protocol for the B or bearer channel

LCD — liquid-crystal display

LPC — linear-prediction coding

luminance pixel — black-and-white or intensity pixel

modem — modulator–demodulator

MOV — metal-oxide varistor

NTSC — National Television Systems Committee. Also refers to the television standard for the U.S.; i.e., 525 lines of resolution transmitted at 60 half frames (interlaced) per second.

pixel — picture element; unit of resolution is pel per inch. Pixels, the smallest display elements (usually, dots or clusters of dots) on a video display screen, are used to construct the screen image.

POTS — plain old telephone service

RAM — random-access memory

RFI — radio-frequency interference

RISC — reduced instruction set computer

ROM — read-only memory

V.32bis — CCITT standard for sending data at 4,800, 9,600, or 14,400 b/s over standard voice telephone lines; *bis* defines this as an extension of the V.32 standard for 4,800 or 9,600 b/s.

X.25 — packet-switching protocol at link and packet layers; a CCITT packet-switching protocol and standard

Y, R – Y, B – Y — color-separated video component of the RS-170A standard; where Y is luminance and R – Y and B – Y are color differences

VideoPhone 2500 Architecture

A useful perspective for the architectural design of the VideoPhone 2500 is the protocol layer that takes data from all the inputs and hands data to the outputs. By using this perspective, we could better maintain a modular-design approach for both hardware and software. Furthermore, the protocol layer most directly affects the user and interacts with the hardware subsystems.

The processor of the VideoPhone 2500 acts like a traffic cop at a busy intersection. At this intersection, the data taken from the video codec (coder–decoder) and audio processor must be bundled and sent through the modem for transmission. Meanwhile, data from the far-end set is coming in through the modem and must be checked for integrity, unbundled into video and audio data, and sent to the video codec and audio processor, respectively. In addition, while the modem traffic is being handled, the user is making requests through key presses that are sensed by the set's POTS processor and forwarded to the host processor for action. (POTS stands for *plain old telephone service.*)

Figure 1 contains a photograph and a simple block diagram of the VideoPhone 2500. The rest of this section provides a functional breakdown of each block.

Host Processor. The host processor is a Motorola MC68302 processor, which has a Motorola 68000 microprocessor as its core central-processing unit (CPU).

On-board, RISC, peripheral processors handle the serial communications protocol, reducing the core CPU load. (RISC stands for *reduced instruction set computer.*) A major reason for choosing this part was that it contained these communications processors.

Modem. The modem is capable of full-duplex, synchronous data transmission at either 19.2 or 16.8 kilobits per second (kb/s). An unusual feature of this

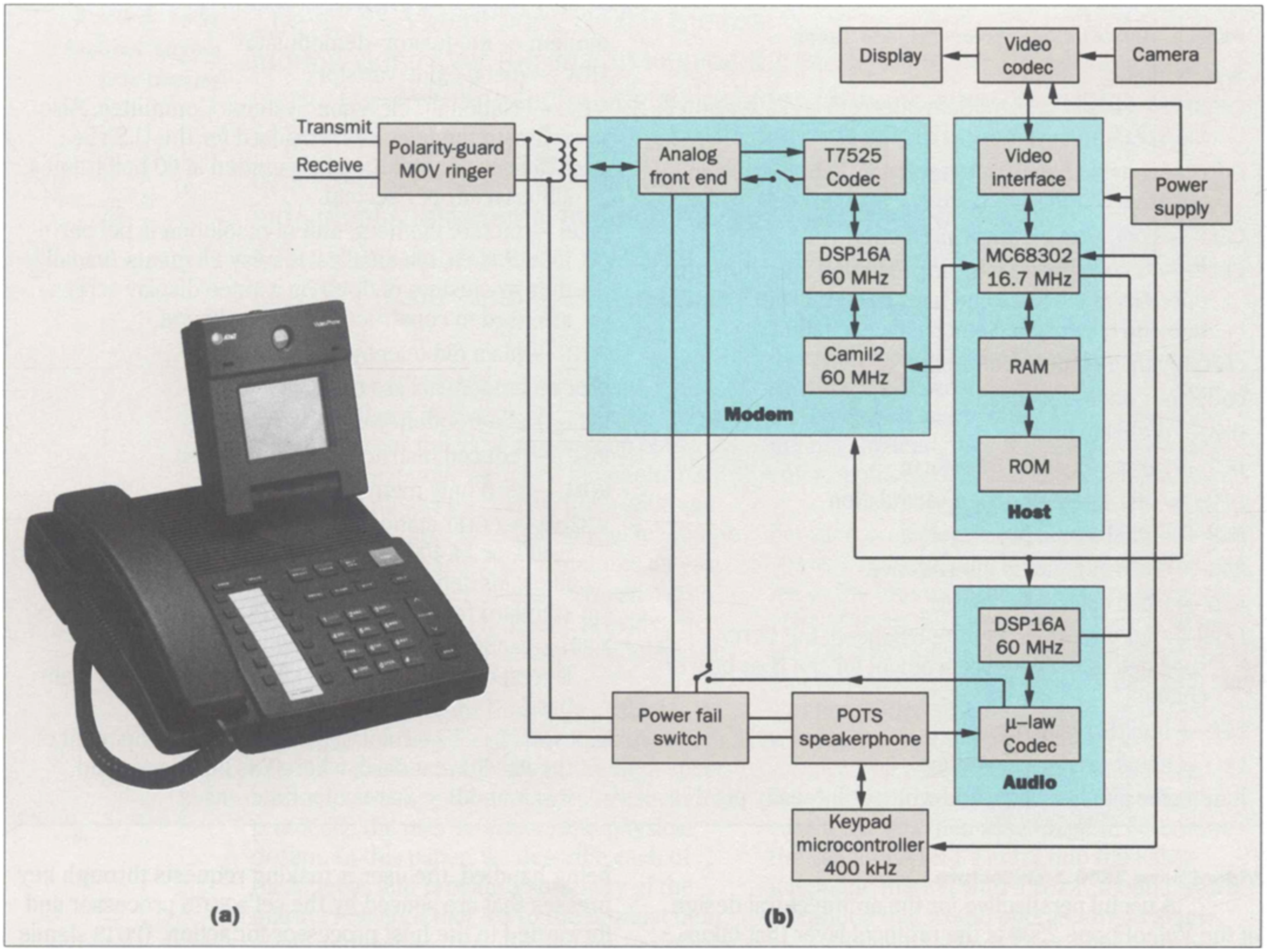


Figure 1. The AT&T VideoPhone 2500 is the world's first motion, full-color video telephone that works over existing telephone lines. (a) The unit is ready to use immediately after purchase. Users remove it from the packaging, plug it in, and start using it. (b) Block diagram of the AT&T VideoPhone 2500. Enabling technologies include video and audio compression and a high-speed modem.

modem is its fast start-up time. Fast start-up is a system requirement because of this nontraditional use of a modem in a person-to-person communications system where a short start-up time is preferred.

During a call, either the caller's modem or the called party's modem can become the answering modem at any time. Therefore, the modem also listens to the voice call to determine when to it should begin modem initialization.

In a later section, we provide more information about the high-speed modem.

Audio Processor. The audio processor encodes and decodes audio data, using a code-excited linear prediction (CELP) algorithm at 6800 b/s (bits per second). The enhanced algorithm, dubbed CELP+, was developed at AT&T Bell Laboratories in 1991. AT&T Consumer Products in Indianapolis, Indiana, added the interface code for start-up and exchange of data and commands. The section on speech compression in this paper provides more detail about the CELP+ coder.

Video Codec. The video codec encodes and decodes video images. It manages three interfaces:

- Host command and data interface. This interface permits robust data exchange, a start-up sequence, character generation on the display, adjustment of video coding parameters, color adjustment, diagnostics, and other features. (By *robust*, we mean it detects its own problems during data exchange.)
- Analog video input interface.
- Analog video output interface.

Compression Labs, Incorporated, in San Jose, California, designed and supplies the video codec to AT&T specifications. A later section contains details about the operation of the video codec.

POTS/Speakerphone and Keypad. The POTS and speakerphone provide the set's analog audio interface and power-fail operation. (The latter means that the telephone will work without power from a wall outlet.) The circuits are based on the AT&T 730 telephone, which has been modified to work with additional circuits. Similarly, the keypad is based on the AT&T 730 telephone's keypad. In addition, a serial link to the host processor has been added, and the software feature flows were altered considerably.

Display. The 3.3-inch liquid-crystal display (LCD) shows color video images, using field overlay rather than a full video frame, which requires frame interlace. In addition, the display input is an NTSC-variant; i.e., it uses NTSC's timing, but not the voltage or synchronization. [NTSC is the National Television Systems Committee, although the term also refers to the television standard this committee set for the United States. In that context, NTSC means 525 lines of resolution, transmitted at 60 half frames (interlaced) per second.]

The LCD is a custom modification of a mass-produced device from Epson America, Inc.

Camera and Lens. The camera provides an NTSC-variant color-video output signal and an auto-white balance function. (*Auto-white balance* means that the color is balanced automatically, using the white areas of the image as a reference.) Sony Corporation designed the camera specifically for this application and now manufactures it for AT&T.

During a call, one to three users may be using the set at each end. Therefore, the lens provides a wide angle of view and adequate depth of field for the image expected in normal use.

Communications Protocol

The architecture of the VideoPhone 2500 treats the data sources and sinks (e.g., the video codec, audio processor, modem, keypad, display, and camera controls) as peripheral devices to the host processor. Some—such as the video codec—are sophisticated devices, while others—such as the camera's On/Off control—are simple ones. From the application layer, the far-end set is seen as another peripheral, whose link is through the modem and the communications protocols.

The VideoPhone 2500 multiplexes data from three independent sources over the physical line; i.e., video data, audio data, and supervision information. The basic vehicle used for this multiplexing is the X.25 protocol, which creates permanent virtual circuits for each channel, assigning a different channel number to each one. This channel number serves as a prefix to each frame transmitted and is used on the receive side to demultiplex the data stream. The X.25 packet is encapsulated within a LAPB frame. The LAPB address is constant and is not used to demultiplex the received data stream. (LAPB stands for *link-access procedure-balanced* and is the data-link-level protocol for the B or bearer channel specified by CCITT, the International Telegraph and Telephone Consultative Committee in Geneva, Switzerland.)

The X.25 protocol automatically ensures that a reliable connection has been provided for each of the channels. After data has been transmitted, the data frame is retained in the transmitter memory until acknowledged by the remote peer. If not acknowledged, the data frame is retransmitted. Eventually, either the data will be acknowledged or the channel will be reset.

Because audio is a real-time contrained process (i.e., it must be continuous, even with errors), the audio frames are implemented using an unnumbered information, LAPB header (a type of LAPB frame).

High-Speed Modem

The VideoPhone 2500 has an integrated, high-performance, two-wire, full-duplex modem that supports data rates of 19.2 and 16.8 kb/s.

To operate reliably over the telephone network at these higher data rates, the modem had to signal at a higher rate. A higher signaling rate uses more of the available channel bandwidth, improving performance. The 19.2-kb/s mode effectively transmits 6 bits per baud at 3200 baud, and the 16.8-kb/s mode transmits 6 bits

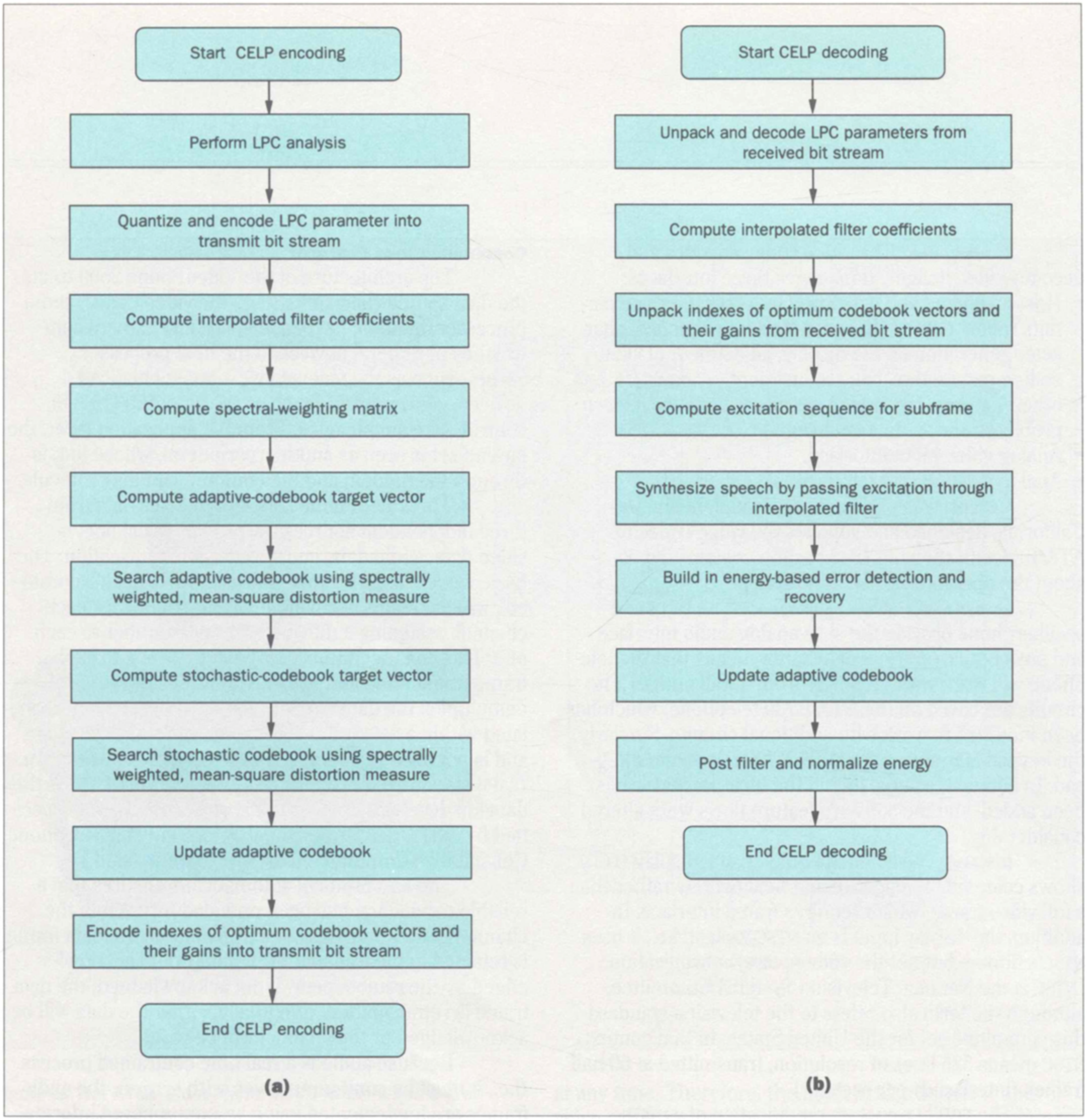


Figure 2. To compress speech, the VideoPhone 2500 uses a CELP algorithm that operates at 6.8 kb/s and incorporates fractional-pitch prediction. Here, we show the steps involved in (a) encoding and (b) decoding a frame of compressed speech using the algorithm.

per baud at 2800 baud. This contrasts with the CCITT V.32bis standard, which transmits 6 bits per baud at 2400 baud to achieve 14.4 kb/s. (A *baud* is an alphanumeric character, such as "a." In ASCII, the American Standard Code for Information Interchange, seven bits are needed to describe "a," even though it is only one character.)

Like other high-speed modems, the VideoPhone 2500 modem uses trellis-coded modulation (i.e., Viterbi encoding) to improve performance. But instead of the two-dimensional, 8-state code employed in the V.32bis protocol, the VideoPhone 2500 uses a four-dimensional, 16-state trellis code. This trellis code provides more coding gain and decreases constellation density, which is the number of constellation points needed to signal the baud.

Today's telephone network uses predominantly digital transport, which means that quantization noise is a primary noise source on the channel that the modem sees. The encoding techniques used in the network produce quantization noise that is signal dependent; i.e., larger signals are encoded with more noise and smaller signals with less noise. An efficient mapping scheme is used to translate the data into constellation points to minimize the peak-to-average ratio of the constellation. This helps to minimize the quantization noise introduced by the network encoding process.

Telephone channels can introduce amplitude distortion that, after equalization in the receiver, can color the channel noise. Other impairments can also be correlated over several symbol intervals. This correlation reduces the coding gain that the trellis code achieved. To combat this problem, interleavers are used to "shuffle" the symbols sent at the transmitter and "unshuffle" the symbols at the receiver. Impairments that the channel introduced have not been through the shuffling process at the transmitter and, are randomized at the receiver. This "whitening" effect improves the coding gain.

The hardware technology used is simple, powerful, and cost-effective:

- A single, ROM-coded, AT&T DSP16A digital signal processor is used for all signal processing.
- An AT&T application-specific integrated circuit, the CAMIL2, acts as an interface device to the VideoPhone 2500's host processor and supplies other ancillary functions, such as clocking circuitry and a data interface. (CAMIL2 stands for *core-access modem-interface logic, version 2*.)
- An AT&T T7525 linear codec provides the analog-to-digital (A/D) and digital-to-analog (D/A) conversion to interface to the telephone line.

All three parts are manufactured by AT&T Microelectronics in Allentown, Pennsylvania.

Speech Compression

To compress speech, the VideoPhone 2500 uses a CELP algorithm or coder.[1] The coder operates at 6.8 kb/s and incorporates fractional-pitch prediction[2] and constrained stochastic excitation.[3] Because of these enhancements, the algorithm is labeled a CELP+ coder. This section contains an overview of its operation.

High-pass filtering removes low-frequency noise from the input speech. Then, a Hamming window (a common, digital-filtering technique) is applied to the speech buffer before linear-prediction-coding (LPC) analysis is done. An autocorrelation method of LPC analysis generates the filter coefficients, which are bandwidth-broadened (so the speech sounds better when decoded) and converted to log-area ratios before being quantized. The log-area ratios are interpolated and converted to filter coefficients that determine the target vector for both the adaptive and the stochastic codebook searches. (Both filter coefficients, i.e., the original value and the value from the second conversion, are compared so that the error can be calculated. The codebooks provide the values used to encode speech.)

Excitation parameters, pitch delay, and spectral weighting are synthesized. The adaptive codebook index, stochastic codebook index, and associated codebook-gain quantization indexes are determined for best fit using a spectrally weighted, mean-square distortion measure. Figure 2a shows the various steps involved in encoding a frame of compressed speech.

To decode a frame (Figure 2b), the ten log-area ratios are recovered from the compressed-speech bit stream, and are interpolated and converted to filter coefficients. Next, the excitation sequence is reconstructed, using both codebook indexes and the codebook-gain quantization indexes. Speech is synthesized by filtering the scaled excitation sequence with the interpolated filter. An optional, energy-based error detection scheme can aid recovery under adverse channel conditions. Figure 2b shows the steps involved in decoding a frame of compressed speech.

Again, the hardware is straightforward and cost-effective. A single, ROM-based, AT&T DSP16A digital signal processor does all the speech processing. In addition, an AT&T T7513 μ-law codec provides the A/D and D/A conversion for the speech input and output.

Video Compression

The video subsystem accepts analog component signals in Y, R − Y, B − Y format from the camera, along with horizontal and vertical synchronization information. (The timing format is part of the Electronic Industries Association RS-170A standard for NTSC video signals. This format defines color separation. The Y represents luminance or intensity, R − Y is one color difference, and B − Y is the other color difference. Y, R − Y, and B − Y are all that are needed to describe an image.) The analog

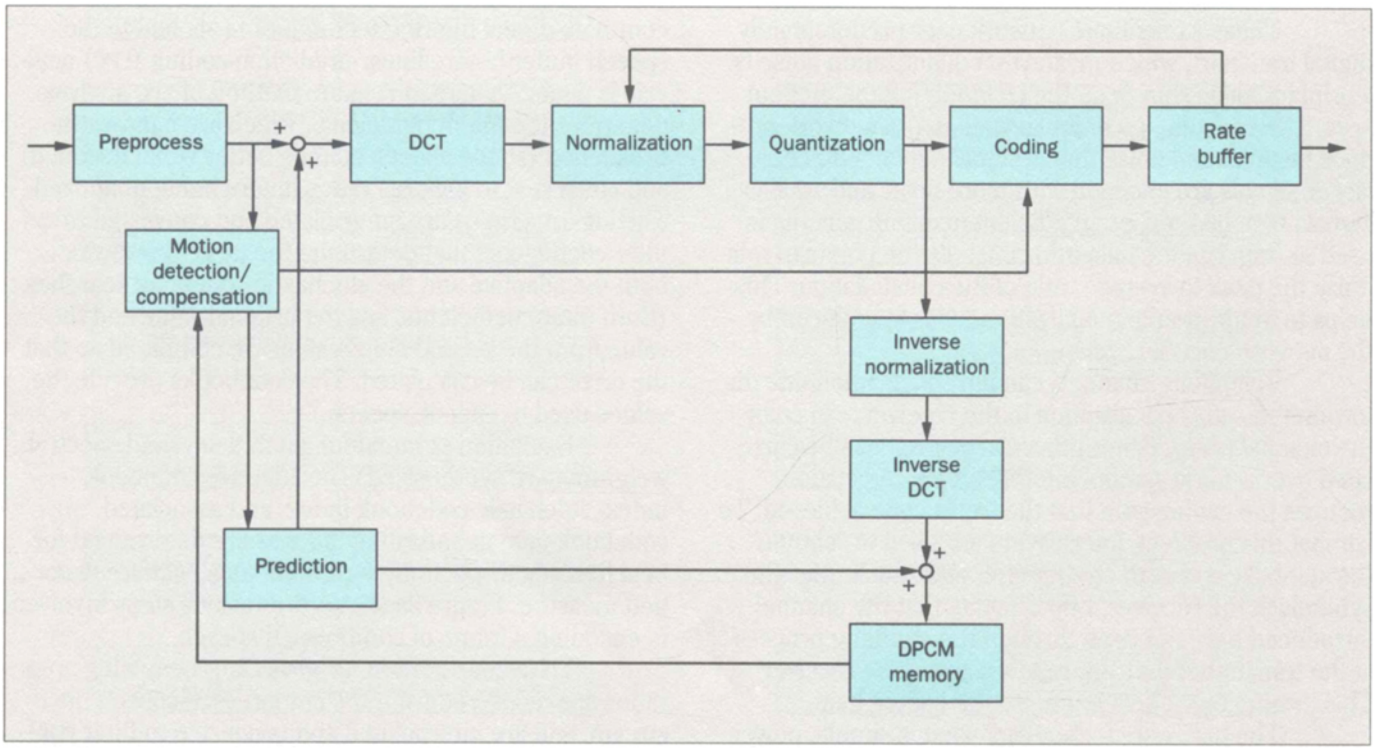


Figure 3. Motion-compensated interframe and intraframe coding system. The luminance and chrominance frames are segmented into 16-by-16 blocks before being injected into the DPCM loop.

signals are digitized and preprocessed with spatial and temporal filters to produce three separate video frames. One frame contains luminance pixels, and the other two contain chrominance (color) pixels. The luminance resolution is 128 pixels by 112 lines, but the chrominance frames are decimated to 32 pixels by 28 lines.

Next, the luminance and chrominance frames are segmented into 16-by-16 blocks and injected into the digital pulse-coded modulation (DPCM) loop, as shown in Figure 3. The motion estimator searches for the best subpixel offset between the current block and the previous frame that results in minimum error. Then, the current block is subtracted from the best-matched block in the previous frame to produce the DPCM-error block, which is then converted to the frequency domain by the discrete cosine transform (DCT).[4] The output of the DCT is normalized and quantized, and the remaining nonzero coefficients are encoded for run length and placed in the rate buffer, ready to be transmitted.

The quantized coefficients are next quantized inversely and converted inversely with the DCT. The result is added to the DPCM memory to create a new, reconstructed block that is stored in the DPCM-reference memory. The receiving VideoPhone 2500 duplicates this reconstruction path, so that both DPCM memories are identical.

After a frame of video is received, the blocks are interpolated back to their original resolution. Each Y, R − Y, B − Y pixel is transformed into RGB (i.e., red, green, blue) and converted to analog to be displayed on the LCD screen.

Because the channel rate is fixed and the amount of data that flows into the rate buffer varies, a feedback mechanism is needed to prevent overflow of this buffer. The normalization process monitors the level of the rate buffer; hence, the severity of the normalization will vary. When a lot of motion and spatial high frequencies are present, the normalization increases to keep the number of bits per frame as constant as possible.

The VideoPhone 2500's allocated video

bandwidth is about 10 kb/s, and its nominal frame rate is 10 frames per second. This yields 1 kilobit per frame. Even at the highest normalization, the number of bits per frame often exceeds this average. When this happens, the frame rate is reduced, so that the rate buffer will not overflow. By using the FOCUS button, the user can change the maximum allowable normalization, thus trading frame rate for normalization noise.

Periodically, each block is sent as an INTRA block, which prevents the interframe DPCM process from occurring, and the DCT coefficients are sent to represent the real video data. This action will clean up the reconstruction DPCM memory, if any line errors have occurred that would have corrupted the reference memory. Because of this periodic refreshing of the reference memory, a video artifact that results from a line error will persist on the screen for no more than 15 seconds.

User Interface

A product's user interface is what a customer controls and perceives. This interface includes how the set looks, sounds, and feels and how easy the set is to use. Our goal was to give the VideoPhone 2500 an interface that was as much like a plain, ordinary, voice telephone's interface as possible.

When the Human Factors Lab at AT&T Consumer Products in Indianapolis, Indiana, conducted its first customer interviews, privacy was mentioned constantly as a concern for this mainly residential product. People were uneasy about having a camera attached to their telephone lines, if they did not have full control of whether their images were sent out. Thus, besides making the VideoPhone 2500 set as easy to use as an AT&T voice-only telephone set, we had to build in assurances of privacy in its physical attributes and its operation. We deemed it mandatory to have a mechanical, visible shutter in front of the camera's lens. When the lens is covered, even the most cautious user knows that the image is not being transmitted.

Feature operation was designed to be as familiar, comfortable, and simple as possible. Telephony functions were closely modeled after the AT&T 730 telephone, a corded set that has a two-line display. Figure 1a is a photograph of the VideoPhone 2500.

Video communications from one's home initially could be unnerving, so the actions required were made as simple as possible. The set has a large, blue button labeled VIDEO in the top right area of the button field. After an analog call has been established, users at each set must push the VIDEO button to initiate the video communication.

A privacy safeguard was the assurance that a user's image would not be sent out until he or she had pressed the local set's VIDEO button. Users would be able to see the person at the far end after the far-end set's VIDEO button had been pressed. However, two-way video communication could not occur until the users at both ends had pressed their VIDEO buttons. Thus, turning on video transmission is the only feature that requires action by both users.

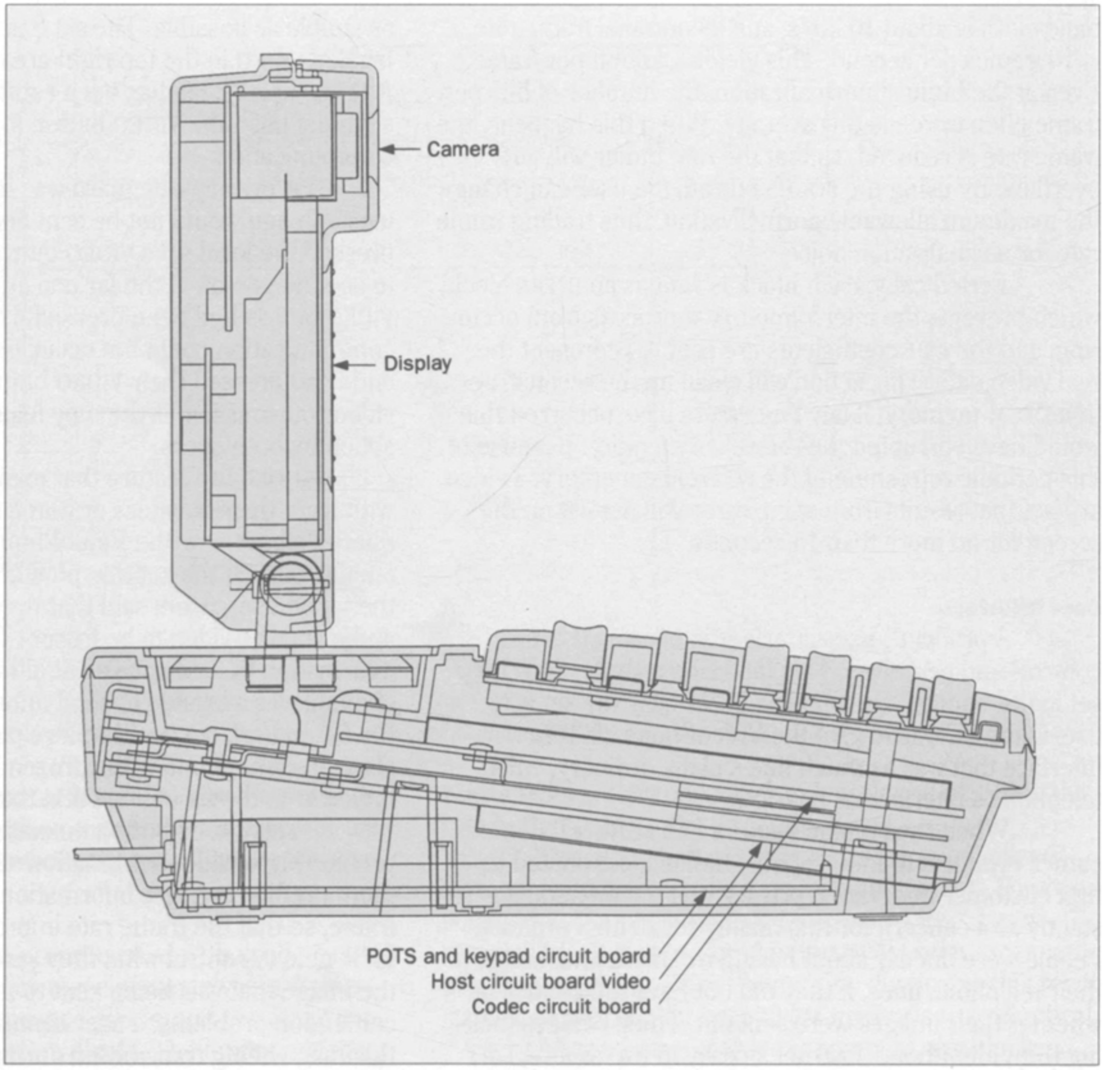
HOLD is a feature that many users are familiar with from their business or home telephone sets. When queried about how the VideoPhone 2500's HOLD feature should work at the set that puts the call on hold and at the far-end set, users said that they expected both the audio and the video to be frozen. They did not want new information to be displayed at either set. Because a data channel was available to send information to and control the far-end set, we could relieve people's possible anxiety about a silent phone with a frozen image by adding an indicator at the far-end set that the call was on hold.

The FOCUS button on each set controls the prefiltering of video information at the far-end set. With more prefiltering, less information changes in each frame, so that the frame rate improves. People expected to be able to control what they saw, rather than control the image that was being sent to the other end. To avoid contention problems, a user cannot change the focus of the image being transmitted during a video call, but can always refocus the image he or she receives.

In addition, we needed to provide a way for people to compose their pictures and check their lighting and setting. A toggle button on the set, labeled SELF VIEW, provides this capability. A user can access this feature at any time during a voice call, a video call, or even when not on a call. Because the set requires some time to react to this button press, an indicator on the display gives immediate visual feedback that this button has been pressed. In addition, all button presses have auditory and tactile feedback.

During the interviews, we learned that some people (such as children) might prefer to look at

Figure 4. Cross section of the AT&T VideoPhone 2500. The physical designers, an industrial designer, EMI and heat consultants, and electrical hardware designers worked together to develop an attractive, telephone-size, cost-effective package. Durability, customer safety, and repair considerations also influenced the design.



themselves instead of the far-end party during part of the call. That user might activate SELF VIEW but still wants to know if his or her image and voice are being transmitted. Because video transmission continues even when SELF VIEW is on, the display informs the user who is in the SELF VIEW mode that the set is still sending out picture and voice.

Physical Design

Because of its high-speed electronics, the VideoPhone 2500 would have had insurmountable problems with electromagnetic interference (EMI), radio-frequency interference (RFI), electrostatic discharge (ESD), and heat if we had not accounted for them in the initial design. The physical designers worked jointly with the industrial designer, EMI and heat consultants, and electrical hardware designers to develop a physical design architecture that addresses all the concerns in a good-looking, telephone-size, cost-effective package. Figure 4 shows a cross section of the videophone.

To mitigate electromagnetic and radio-frequency interference, a ground plane spine connects to the high-speed host and video-codec circuit boards. The ground plane provides for ground points, as needed, to short-circuit the eddy currents that induce electromagnetic interference. The high-speed circuits are grouped as closely together as possible, under metal EMI covers. The circuit boards contain EMI moats, bridges, and ground layers. This design minimizes interconnects and, where needed, treats them with flexible ground planes—as in the flexible cable that goes to the upper module camera and display.

The circuitry inside the set generates about 13 watts of heat, while a normal telephone's circuitry generates about one-tenth this amount. Thus, heat could be a problem. Because this set also has a speakerphone, noisy fan cooling was not an option. The design could only depend on convection and conduction cooling. An early heat analysis showed that this cooling should be possible if we:

- Used adequate venting.
- Put the hottest circuit boards and components toward the top of the air paths.
- Slanted the circuit boards to maximize air flow.
- Used a heat sink on the biggest voltage regulators.
- Provided vent holes in the EMI covers.
- Spread out the high-heat-producing components.
- Maximized all the air paths.

As a result of these design guidelines, the highest internal temperature of the package is below 70°C in a 38°C ambient environment.

The large number of vents and holes and the proximity of circuit-board assemblies to them highlighted ESD as a problem. To prevent ESD damage to components, any discharges must be directed to ground before they get too far into the circuit boards. Therefore, the design provides ESD ground paths around the top and bottom edges of the circuit boards. These paths tie into a green-wire ground or the tip side of the telephone line.

Durability, customer safety, and repair considerations led to several important design decisions. The VideoPhone 2500 set had to withstand unpackaged drop tests from heights of 6 and 3 inches, but the goal was for the set to withstand drop tests from 30 inches, if possible. At odds with this goal was the possibility of damage to the display, lens, or camera in a "freak" fall. This conflict led to a compromise. To reduce the amount of energy absorbed by the internal components, the design allows the snaps that hold the video module (i.e., the display and camera) to the telephony module to fail by shearing, which permits the modules to separate. This reduces repair costs because only the plastic housing needs to be replaced; the expensive video components can be reused in the new housing.

This approach solved two design considerations (i.e., durability and repair) but aggravated the third, customer safety. The interconnect cable that runs between the two modules carries high voltage, and physical contact with it could cause a minor injury.

To assure customer safety, we used zero-insertion-force flexible connectors in the video module and standard insertion connectors in the telephony module. Zero-insertion-force connectors require a much higher pull-out force than the other connectors. When the modules separate during a drop or because of customer abuse, the cable stays with the video module, and is disconnected from the high voltage that is developed in the telephony module. Therefore, a customer's fingers cannot make contact with the high voltage.

Summary

The VideoPhone 2500 provides video telephony across the analog loop through the public switched telephone network. This capability has been made possible through advancements in video and audio compression and in modem technology. The color, motion video significantly enhances the perceived quality of a telephone conversation. This AT&T product is the first in a series that provides the advantage of video telephony.

Acknowledgments

We would like to thank the entire AT&T VideoPhone 2500 development team. Their combined efforts made this product a reality. In addition, the following people made specific contributions to this paper: John Slevin (for market research); Jonathon Masel of Chimney Software Products in Kfar Saba, Israel (for protocol); Rick Flanagan (for the high-speed modem); Reed Thorkildsen (for speech compression); Ellen Keen (for the user interface); and Bob Beecher and Mike Groesch (for physical design).

References

1. M. R. Schroeder and B. S. Atal, "Code-Excited Linear Prediction (CELP): high quality speech at very low bit rates," *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing*, Tampa, Florida, March 26–29, 1985, IEEE, New York, 1985, pp. 937–940.
2. P. Kroon and B. S. Atal, "On improving the performance of pitch

predictors in speech coding systems," *Advances in Speech Coding*, B. S. Atal, V. Cuperman, and A. Gersho (eds.), Kluwer Academic Publishers, Boston, Massachusetts, 1991, pp. 321–328.
3. Y. Shoham, "Constrained stochastic excitation coding of speech at 4.8 Kbps," *Advances in Speech Coding*, B. S. Atal, V. Cuperman, and A. Gersho (eds.), Kluwer Academic Publishers, Boston, Massachusetts, 1991, pp. 339–348.
4. W. H. Chen and W. K. Pratt, "Scene Adaptive Coder," *IEEE Transactions on Communications*, Vol. COM-32, No. 3, 1984, pp. 224–232.

(Manuscript received June 25, 1992)

Eric Dorsey *is director of engineering at Compression Labs, Incorporated, in San Jose, California, where he is responsible for development of desktop video and videophone products. Mr. Dorsey has both B.S. and M.S. degrees in electrical engineering from Stanford University in Stanford, California.*

Scott H. Early *is a technical manager in the New Business/Forward Looking Work Department with AT&T Consumer Products at the Heritage Park facility in Indianapolis, Indiana. He is responsible for technology platform development and evaluation of new business opportunities for AT&T Consumer Products. Mr. Early joined the company in 1977. He has a B.S.E.E. from Purdue University in West Lafayette, Indiana, and an M.S.E.E. from the University of California at Berkeley.*

Andrew Kuzma *is a member of technical staff in the Video Systems business unit of AT&T Consumer Products at the Heritage Park facility in Indianapolis, Indiana. He is the technical architect for the business unit. Mr. Kuzma joined the company in 1983. He has a B.S. in physics and mathematics from Purdue University in West Lafayette, Indiana, and an M.S. in physics from the University of Illinois in Urbana.*