Multimedia presentation development using the Audio Visual Connection

by D. J. Moore

This paper describes the technology behind the creation of multimedia presentations using a new IBM program product, the Audio Visual Connection® (AVC™). The multimedia approach represents a major innovation in computer technology involving new concepts such as the digitization of audio and video, the involvement of the Musical Instrument Digital Interface (MIDI), and the creation of a multimedia story. Two hardware adapters support the AVC: the Audio Capture and Playback Adapter and the Video Capture Adapter. When used with these adapters, the AVC digitizes both stereo audio and color video, performs powerful edit and synchronization functions, defines an interactive user environment, and creates a multimedia presentation using standard IBM Personal System/2® hardware. Additional input is available from MIDI song files, screen capture, and non-AVC image systems. Final results range from passive presentations to interactive applications to sophisticated database front ends.

In the computer industry, *multimedia* is a term used to describe the interaction of the human auditory and visual senses with the computer. A multimedia computer presents the user with many kinds of image—text, graphics, video, animation—and several types of audio, including speech, music, and sound effects. The multimedia computer involves its user in many levels of interactive decision and control, altering the way information is conveyed.

The purpose of multimedia is to improve communication between computer and user and among users. Ideas may be expressed in many forms other than written words that are read sequentially. Image and audio may now be intertwined with the written word and presented interactively to readers. In this sense, multimedia is a new communication tool.

The interactive nature of advanced multimedia presentations gives users control of the presentation and access to desired types of information. One user may want an overview of a topic, another may be interested in its financial aspects, and a third user may be interested in the theory underlying a concept. Each user may receive what is pertinent in a self-determined sequence. The user is in charge of the information flow and is not subject to the predefined ideas of the information provider.

Although the usefulness of multimedia for graphic artists and musicians is self-evident, consider its use by a mathematician. A multimedia contribution in a field of mathematics may consist of graphics illustrations of the concept being presented, text, and the author's voice discussing the subject. How much more personal would be a mathematics text that visually animates the results of a mathematical process than one that merely describes the results. In the same way, the ability to hear the actual voice of a person whose biography you are reading and to see the creative results of that person's lifetime make the educational experience more meaningful.

The Audio Visual Connection brings the ability to create a complex multimedia presentation within an establishment.

A multimedia document need not be experienced in any predetermined sequence. Hypertext-like techniques allow a user to peruse a document in many directions. The selection of a word or an image may trigger the display of additional images and audio to reinforce the subject. Once a path through a document has been completed, the reader may select another word or image to continue on the path of discovery or return to the point of original departure. A number of important considerations must be taken into account to make sure that a reader obtains all relevant information about a topic, when so many different directions are possible. This possibility was conceived in the mind of Vannevar Bush in 1945, including an imagined implementation using components that were then available.[1]

The Audio Visual Connection system

The Audio Visual Connection® (AVC™), which is one of the IBM multimedia program products, is used in this paper to illustrate the process of creating a multimedia presentation. Storyboard Plus™, LinkWay™, and InfoWindow® are other IBM multimedia products. Like them, the AVC is a tool to develop the final application, the multimedia presentation. The AVC is not itself the end product.

The Audio Visual Connection brings the ability to create a complex multimedia presentation within an establishment, eliminating the need for outside production facilities. Changes to the presentation may similarly be incorporated immediately within the establishment. The AVC software may be used with two associated Personal System/2® (PS/2®) hardware adapters, the Audio Capture and Playback Adapter (ACPA) and the Video Capture Adapter (VCA). The ACPA is used for both audio *capture* (i.e., the conversion of an analog signal into digital format for storage in computer memory) and *regeneration* (i.e., the conversion of a computer-stored digital signal back into analog format). The VCA is normally used only for video capture as images are regenerated on standard PS/2 displays without the need for special hardware. An exception occurs when a video output of an AVC presentation is desired, such as when using the VCA to generate a video output signal for recording a presentation by a video cassette recorder.

The multimedia creation process

A multimedia presentation is created in the following three stages: capture, authoring, and presentation.

Capture. The capture process converts various electronic media forms into digital representations of the original media and stores them as computer files. Examples of these media might include a television camera's image of a flower, a scanner's image of a document, a microphone's audio signal of human speech, and a compact disc's audio output of a symphony. The ACPA and VCA adapters and the AVC software perform this conversion, and accept files created by certain external programs as well. Additional equipment, such as a video camera, video monitor, video cassette recorder, compact disc player, digital audio tape recorder, microphone, audio amplifier, and speakers complete the multimedia capture system.

Technical specialists are required to set up the audio and video capture equipment initially. Knowledge of the audio and video equipment, lighting techniques, and methods of recording a noise-free audio signal are essential for the successful capture of a multimedia production. Once the equipment has been set up, the capture process may be performed by less experienced personnel.

Authoring. The authoring process establishes a set of procedures called a *story* that controls the multimedia presentation. The story defines the sequence and methods in which the various media formats are displayed or heard, the synchronization between the types of media, and the human interaction. With the AVC, the human interaction capabilities are extensive. For example, the AVC story might allow control of the presentation by the movement of a cursor on the computer screen. Another possibility for presentation control might be the AVC story's evaluation of data entered by the user in response to a displayed or spoken question to start a complex database search on a remote mainframe. For a simple presentation, the interaction may be as straightforward as allowing the user to step through images by the touch of a mouse button.

Text may be overlaid on an image to provide information or pose questions.

The authoring process provides an editing capability for the computer files that have been created in the capture process. Audio files may have click and pop sounds removed or a new phrase may be inserted into a previously recorded speech. Image files may have background objects removed. Also, parts of one image may be mixed with another image, or an image may be displayed as a mirror of the original. Text may be overlaid on an image to provide information or pose questions. Drawing tools may produce images created on the computer, from simple boxes, ovals, and graphs to elaborate drawings.

The authoring environment requires an ACPA and audio playback equipment (assuming that audio is involved in the presentation), because standard PS/2 hardware is used for display. Successful multimedia presentations require preparation and a well thought-out plan. As with any presentation, proper planning is very important. Once complete, making minor changes to a multimedia presentation, such as altering price information or replacing an image, requires minimal multimedia skill.

Presentation. The presentation process is the execution of the story, which coordinates the multimedia presentation and the interaction with the user. The presentation environment requires an ACPA and audio playback equipment, again assuming that audio is used in the presentation. If the presentation is made to a large group, a video projection system may be used. A well-designed story provides the user with the knowledge and options to execute the presentation; thus it imposes no special multimedia experience requirements on the user.

Multimedia example

Consider a department store chain that captures images of its product lines and combines those images with spoken and musical overlays to create stories for each department. The multimedia presentation is sent to each store of the chain and the stores distribute specific stories to the various departments. Customers look at the products displayed on the screen. They select specific items on the screen, then type their names and addresses on the computer keyboard. Each customer's action initiates the following actions by the store: (1) the order is sent to the warehouse; (2) the product is moved to the front of the store for customer pickup; (3) the charge card is billed; (4) the store's inventory is updated; and (5) the customer's name and address are added to the store's mailing list. Note that the multimedia presentation need not perform every task, but it must be able to link into external programs, such as inventory control and charge card billing. As a variation of the basic procedure, consider the automobile department manager of one store who decides to have a sale on tires. The tire prices are changed in the database at that store and a "Sale" screen is created for that department. The multimedia presentation for the automobile department is updated and customers are then made aware of the sale. The update takes but minutes to accomplish by local personnel. A given establishment might have several capture systems, a larger number of authoring systems, and many presentation systems.

The AVC system configuration

Multimedia involves a large amount of data manipulation, which poses demands upon the processing speed and memory capacity of the computer. The AVC executes on an Intel 80286™ machine or larger under DOS 4.0, Operating System/2™ (OS/2®) Standard Edition, or OS/2 Extended Edition. Improved performance is achieved under OS/2, due to more efficient memory management. A Micro Channel® PS/2 with 4 megabytes of real memory (RAM) is required for capture, and presentation may be performed with either of the bus architectures and about 3 megabytes of RAM. Performance may improve with increased RAM, because less memory swapping is required. Standard PS/2 displays, such as the IBM 8503 monochrome or 8512, 8513, or 8514 color displays, may be used, and the IBM 8514/A display adapter may be added to provide greater color choices and resolution.

A typical AVC multimedia presentation requires about one megabyte of storage per presentation minute. This estimate may vary by a factor of two up and down, depending upon the number of images and their resolution and upon the amount and fidelity of the audio.

A capture system normally contains both an ACPA audio component and a VCA video component, and an authoring or playback system normally contains only an audio component. The Audio Capture and Playback Adapter (ACPA) is available in two models: one for Micro Channel machines and the other for non-Micro Channel machines. The Video Capture Adapter (VCA) is available only for the Micro Channel machines. Models are available for two standards, one compatible with North American broadcast television and the other compatible with most European broadcast television.

Cooperative processing extends the capability of AVC. An IBM Token-Ring Network is used to connect multiple computers for file sharing. One computer stores the audio and image files, and the other computers download the files as required. A C-language programming interface is provided to allow the AVC either to be called by another program or to call other programs. The AVC may access data stored in the OS/2 Extended Edition database. For additional computational power, the AVC can communicate with KnowledgeTool™, the IBM host-based expert system product. In this case, KnowledgeTool performs the inference logic using the AVC as a powerful user interface.

The Audio Capture and Playback Adapter and audio digital theory

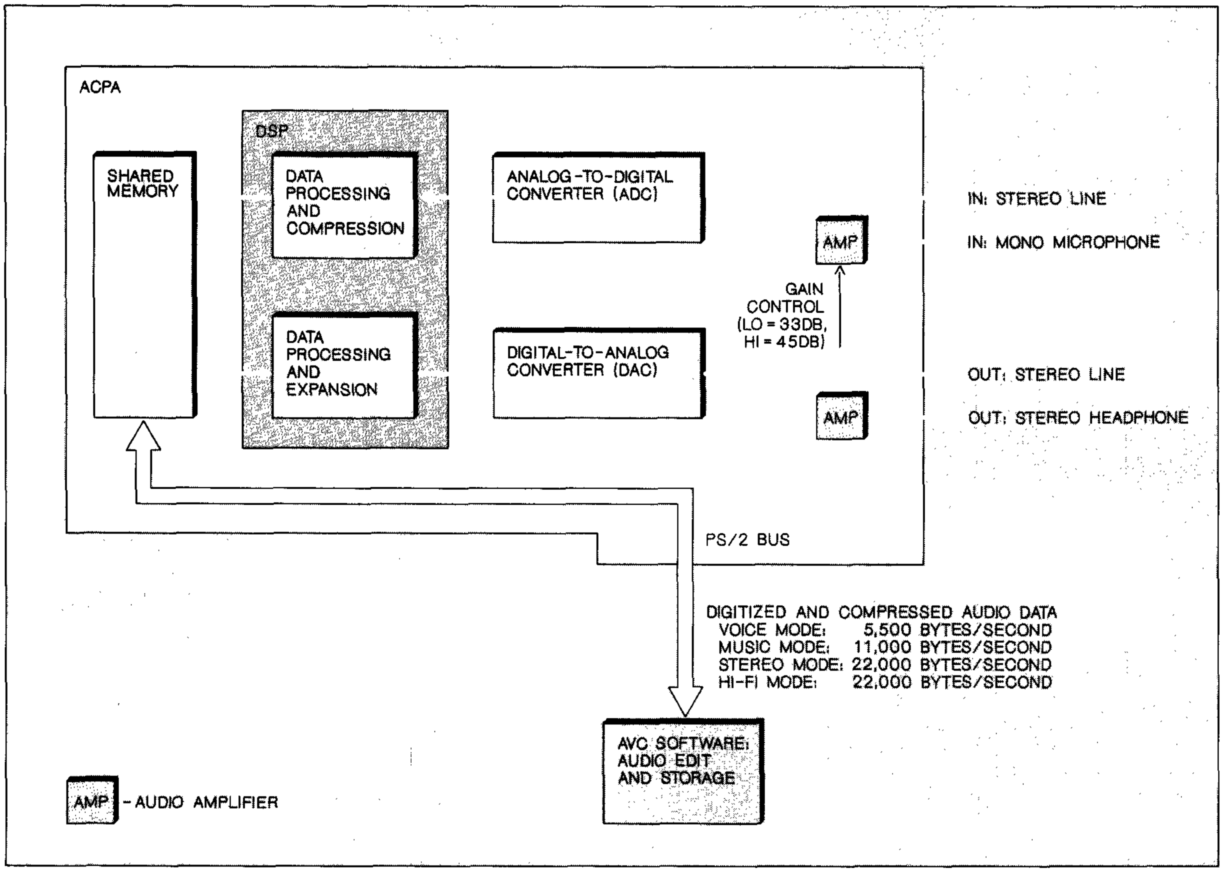
The Audio Capture and Playback Adapter (ACPA) performs both digitization and regeneration of an external audio signal as represented functionally in Figure 1. The digital signal processor (DSP) is a self-contained processor on the ACPA that is configured to perform data compression, expansion, and other audio processing such as mixing and volume control. The ACPA has audio input jacks for both microphone and line; the difference is that the microphone jack is connected to an amplifier to boost the signal. The output of most external audio devices is suitable for the line input. A signal from an analog audio source is digitized by the analog-to-digital converter (ADC). The digitized signal is compressed to save storage and is stored in the ACPA shared memory. The AVC software reads shared memory and stores the audio in PS/2 memory. The audio editing process is described in the section on Audio Visual Connection software in this paper. To regenerate the audio, the AVC software sends the digital data to the ACPA shared memory where the data are expanded. The data are then converted back to analog by the digital-to-analog converter (DAC). The resulting analog signal is sent to output jacks for both headphones and line, the only difference being the power level available at each jack. The line output is suitable for connection to most external audio amplifiers.

The ACPA digital signal processor (DSP) software is loaded from PS/2 memory, thus allowing functional enhancements without changes to the hardware. The DSP software may be changed during operation to redefine the functional capabilities of the DSP. What is now a digital audio processor may become a music generator by the loading of new DSP software. (See the discussion of the Musical Instrument Digital Interface [MIDI] later in this paper for more on this process.)

Audio input. The analog-to-digital conversion is performed by *sampling*. Sampling theory[2] shows that an analog signal can be faithfully converted into a digital representation if the sampling frequency is twice that of the highest analog frequency component. To record the audio spectrum heard by the human ear, the upper frequency component is about 20 000 Hz (hertz or cycles per second). To retain this frequency response, a Nyquist sampling rate of twice the upper frequency component of 40 000 samples per second is required. Frequencies greater than 20 000 Hz must be prevented from entering the digitization process because they may generate false sampling (*aliasing*) audio signals. To accomplish this, a low-pass filter (antialias filter) is used, and a slightly higher sampling rate is used to account for physical limitations in this filter.

Figure 2 illustrates the sampling process. Each sample produces a number that represents the amplitude

Figure 1 Audio Capture and Playback Adapter functional diagram



of the signal at a given point in time. An 8-bit sample or *quantization* divides a signal into 256 (that is, 2^8) discrete levels. By using more bits per sample, a lower noise and more accurate representation of the original signal may be obtained. The Audio Capture and Playback Adapter, like the compact disc, records each of its stereo channels at 44 100 samples per second with 16-bit quantization (that is, 2^{16}) to yield 65 536 discrete levels. The amplitude of the analog signal is measured every 22.7 microsecond (1/44 100 of a second). The measured amplitude is assigned to one of the 65 536 levels and is stored as a 2-byte value. Every second, 44 100 2-byte samples are generated for each of the two stereo channels. That is, data are generated at a rate of 176 400 bytes per second or 10.5 megabytes of data per minute.

Data compression is built into the digital signal processor (DSP) program to reduce the data transmission demand. Table 1 summarizes the four compression modes: hi-fi, stereo, music, and voice. Compression is achieved by *downsampling* (i.e., using every *n*th sample) and by the use of compression algorithms. A twofold downsampling is performed in most modes, thereby producing a two-to-one compression. The compression algorithm involves *adaptive differential pulse code modulation* (ADPCM), which stores the differences between sample values, rather than the actual value of each sample. ADPCM achieves a four-to-one compression.

Audio output. To reduce digital-to-analog conversion errors and to simplify the hardware design, the DSP code uses an *oversampling* technique, whereby audio data from the computer are sampled multiple times to produce 88 200 samples per second. An analog voltage is generated by a digital-to-analog converter from the value of each digital sample. This analog

Figure 2 Audio sampling

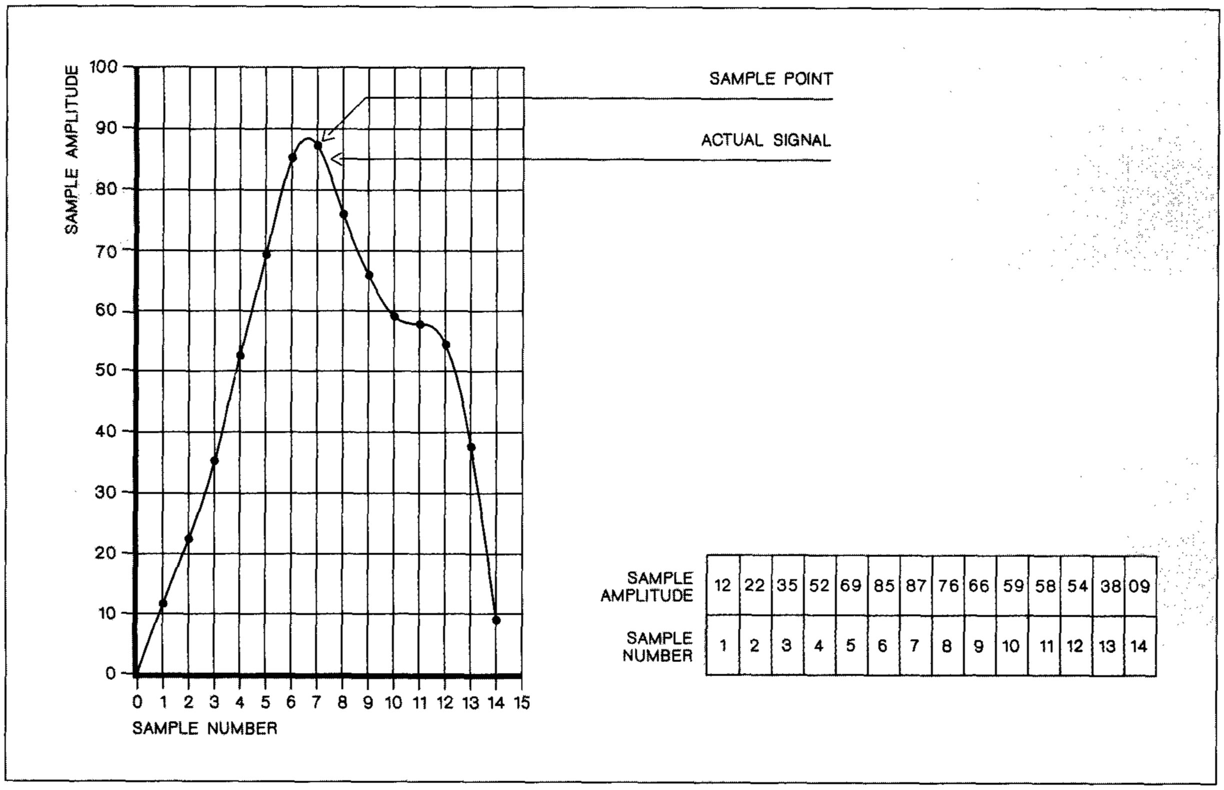


Table 1 Audio mode characteristics

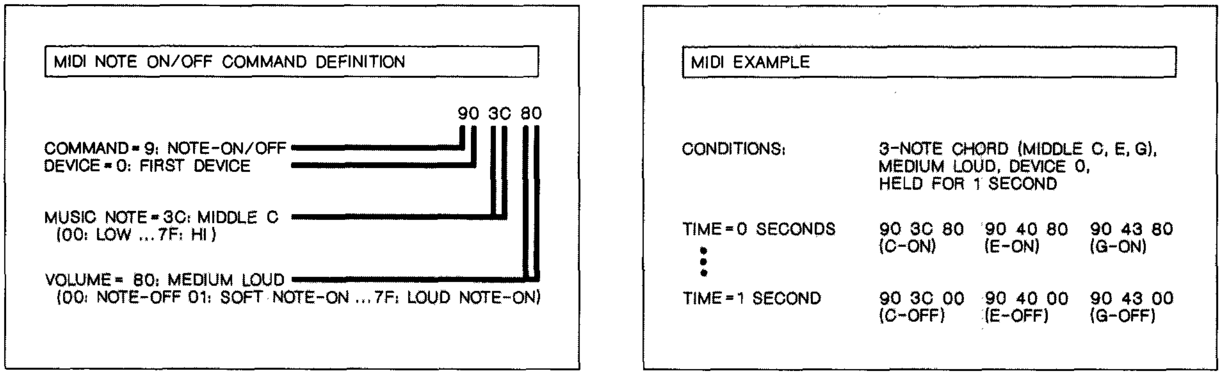
Mode	Mono/Stereo	Data Rate	Frequency Range
Hi-fi	Monophonic	22,050 bytes/second (1 channel × 44,100 samples/second × 4 bits/sample)	20 to 14,000 Hz
Stereo	Stereophonic	22,050 bytes/second (2 channel × 22,050 samples/second × 4 bits/sample)	20 to 7,000 Hz
Music	Monophonic	11,025 bytes/second (1 channel × 22,050 samples/second × 4 bits/sample)	20 to 7,000 Hz
Voice	Monophonic	5,512 bytes/second (1 channel × 11,025 samples/second × 4 bits/sample)	20 to 3,500 Hz

voltage is passed through a digital low-pass filter to remove any high-frequency component and thus produce a smooth output voltage.

MIDI. The Musical Instrument Digital Interface (MIDI) is a data protocol used extensively in the electronic music industry. The ACPA digital signal processor may be reprogrammed by the AVC software to respond to a MIDI data stream and generate its own music. In this mode, the ACPA may play up to eight concurrent musical tones from the MIDI data stream, while playing sampled audio in voice mode. For example, MIDI mode may play three notes as a piano and one note each as a string, flute, horn, bass, and drum along with the sampled voice.

MIDI data provide music capability with far less data storage requirements than sampled music. MIDI com-

Figure 3 MIDI data definition with example



mands denote musical events such as a NOTE-ON (i.e., the start of a note) or a NOTE-OFF (i.e., the end of a note) but not the actual sound of the music. Most MIDI commands have parameters. A NOTE-ON command, for example, defines which note is played, its volume, and the musical device or type of sound being produced. MIDI provides for up to 16 musical devices, but only up to eight are used with the ACPA to denote the eight possible concurrent tones. Figure 3 illustrates the format of a NOTE-ON/OFF command and shows the MIDI data stream for a three-note chord held for one second. In MIDI mode, the MIDI sequence requires 18 bytes of data plus timing information; in the stereo mode, 22 000 bytes of sampled sound are required.

To use MIDI mode, an external MIDI composing program is used to create a song from a music keyboard or from the display of a music score on the computer screen. Once edited and completed, a computer file of that song is generated in standard MIDI format, a standardized file format in the electronic music industry. MIDI song files may also be purchased in a completed form from music publishers. The AVC accepts this MIDI file as input, allows the selection of musical tones, configures the ACPA DSP to generate music, and sends the MIDI data to the DSP for music generation.

The Video Capture Adapter and video digital theory

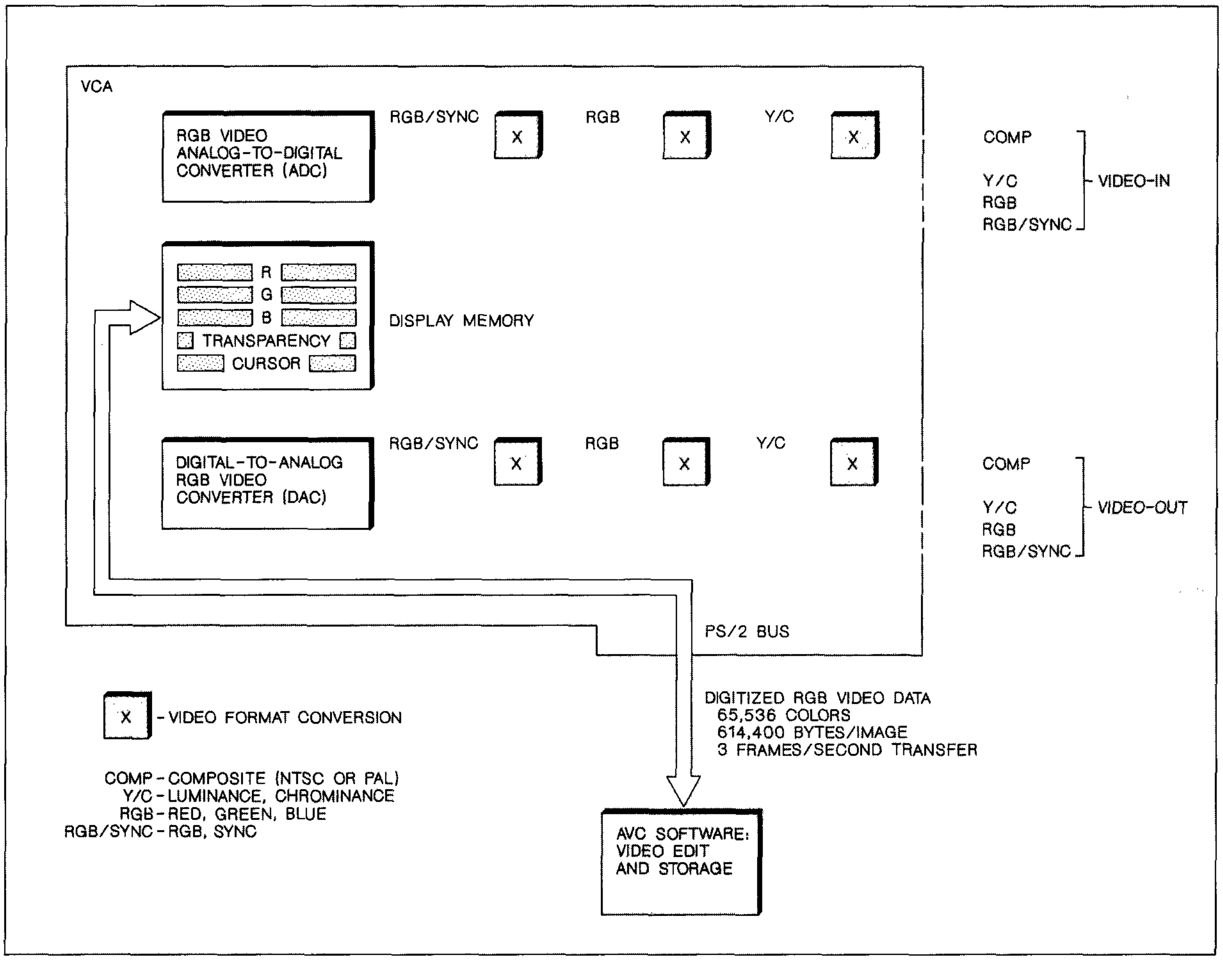
The Video Capture Adapter (VCA) performs both digitization and regeneration of an external video signal.[3] The VCA responds to four types of video signal:

- RGB/Sync—The Red, Green, Blue, Sync mode uses four data paths, one for synchronization and one each for the primary colors: red, green, and blue.
- RGB—The Red, Green, Blue mode uses three data paths where sync is combined with the green signal.

  Because RGB and RGB/Sync modes provide a separate signal for each of the primary colors, each produces a very high-quality image but requires a higher-priced camera and monitor than the other modes.
- Y/C—The luminance/chrominance mode uses two data paths, one for luminance (Y, the monochrome image) and one for chrominance (C, the color content). Y/C is commonly called Super VHS and is becoming increasingly popular in the marketplace. Although the output image is lower in quality than RGB, Y/C hardware is less expensive.
- Composite—The Composite mode combines the entire video signal into one signal but with a corresponding lower quality and lower cost. Composite is the signal produced by the commercial broadcast television industry, consumer video camera, and video cassette recorder.

There are two Video Capture Adapter (VCA) models, one for the National Television Systems Committee (NTSC) composite signal, which is used in North America and Japan, and the other for the Phase Alternating Line (PAL) composite signal, which is used in much of Europe. Composite video signals are *interlaced*, so that the transmission data rate is

Figure 4 Video capture adapter functional diagram



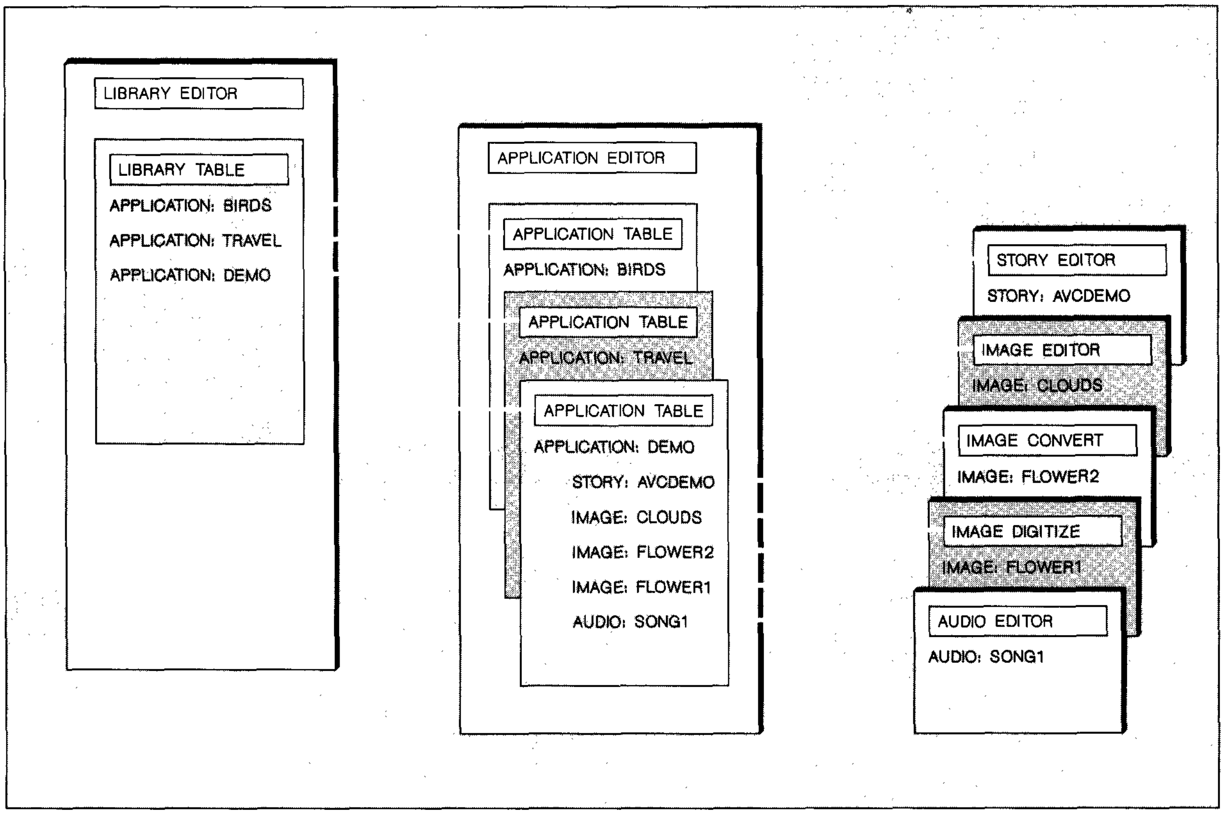
halved by refreshing first the odd-numbered scan lines and then the even-numbered scan lines. NTSC systems refresh alternate halves of the 525 horizontal line display at 60 fields per second, and PAL systems refresh alternate halves of the 625 horizontal line display at 50 fields per second.

Figure 4 is a functional representation of the VCA. A signal from an analog video source is converted to RGB/Sync, if not originally in that format, and digitized by the video analog-to-digital converter (ADC). The resulting image data are sent to the video digital-to-analog converter (DAC) where they are converted into the various types of video signals and sent to the video-out connectors. A monitor is attached to the VCA video-out connector to display the live video from the video source. When a video image is captured, the image is stored in the VCA display memory, and the AVC software reads the VCA display memory and stores the image in computer memory. The video editing process is given in the discussion of AVC software later in this paper.

To generate video, the AVC software sends the digital video data to the VCA display memory, where it is converted back into analog by the digital-to-analog converter (DAC). The resulting analog signal is converted into the various types of video signals and sent to the video-out connectors. A video cassette recorder can be attached to the video-out connector of the VCA to record the AVC presentation.

Video input. If the input signal is composite video, the luminance and chrominance components are

Figure 5 AVC software functional diagram



computed to form a Y/C signal. The luminance signal, with the sync removed, is mixed with a demodulated form of the chrominance signal to produce RGB. The sync signal for RGB is found within the green signal of the RGB and is separated to form RGB/Sync. For input signal other than composite video, the input is sent to the appropriate point in this chain.

The RGB/Sync signal is sampled by the video analog-to-digital converter (ADC) into three digital signals per *pel* (i.e., *p*icture *el*ement or pixel, which is a single dot of the image). Each picture element has three six-bit components, for red, green, and blue. The sampling process is functionally identical to audio sampling, except that the video sampling rate is approximately 12.2 megahertz. If the input is live composite video, a single video frame may be sampled in 1/30th of a second for NTSC and 1/25th of a second for PAL. For the AVC, the three signals are packed into one 16-bit word in the format: 5 bits for red, 6 bits for green, and 5 bits for blue. Each pel can distinguish 65 536 (i.e., 2^{16}) different hues; thus each video frame requires 614 400 bytes.

Video output. There are three sources for video output: live video, still video captured in the VCA display memory, and still video from the computer stored in the VCA display memory. In each case, the video is converted to analog by the video digital-to-analog converter (DAC) in the VCA, much as the audio signal was converted. Once in analog, the video output image is regenerated in the various formats: RGB/Sync, RGB, Y/C, and composite. This output is transmitted to a video device, such as a monitor or video cassette recorder.

Before capture, the three 6-bits/pel of the RGB video-in digital signals pass directly to the DAC for conversion to analog for display on a monitor connected to video-out. The multimedia author views the input video and makes framing and color balance adjustments, as required.

After capture, the video-in image is stored in the display memory as a 16-bits/pel RGB stored image. The monitor connected to video-out now shows the captured image.

During playback, the AVC software may convert the display modes to RGB, output the RGB data to the VCA display memory, and then to video-out. The user can then record the multimedia presentation on a video cassette recorder.

The Audio Visual Connection software

The AVC software[4] coordinates the multimedia process, providing an integration of audio and image with the authoring process. Figure 5 is a functional representation of the major AVC software components.

Library and application editors. An AVC *library* is a list of AVC file groups called *applications*, which may be thought of as special-purpose directories. As shown in Figure 5, the library table lists applications (BIRDS, TRAVEL, DEMO) and the application table lists *objects*, a generic term referring to a computer file that represents a physical object. The AVC application editor handles three types of objects: story, image, and audio. In this example, the application DEMO has one story file, three image files, and one audio file.

Audio editor. To record voice, the multimedia author connects a microphone to the ACPA and talks. Similarly, to record from a compact disc, the user connects a compact disc player to the ACPA and plays the compact disc. As audio is recorded, the data are displayed as shown in Figure 6. Time in seconds is shown under the column headed Play Time, and the small squares under the heading Sound indicate audio volume. The technician adjusts the volume because, if the volume is too high, the magnitude exceeds the ACPA analog-to-digital capacity and distortion occurs; if the volume is too low, distortion may occur because few bits are used to represent the signal. A poor recording can be redone because the audio file can be deleted as with any computer file. During recording, the remaining columns are blank.

Figure 6 AVC audio editor screen with explanation

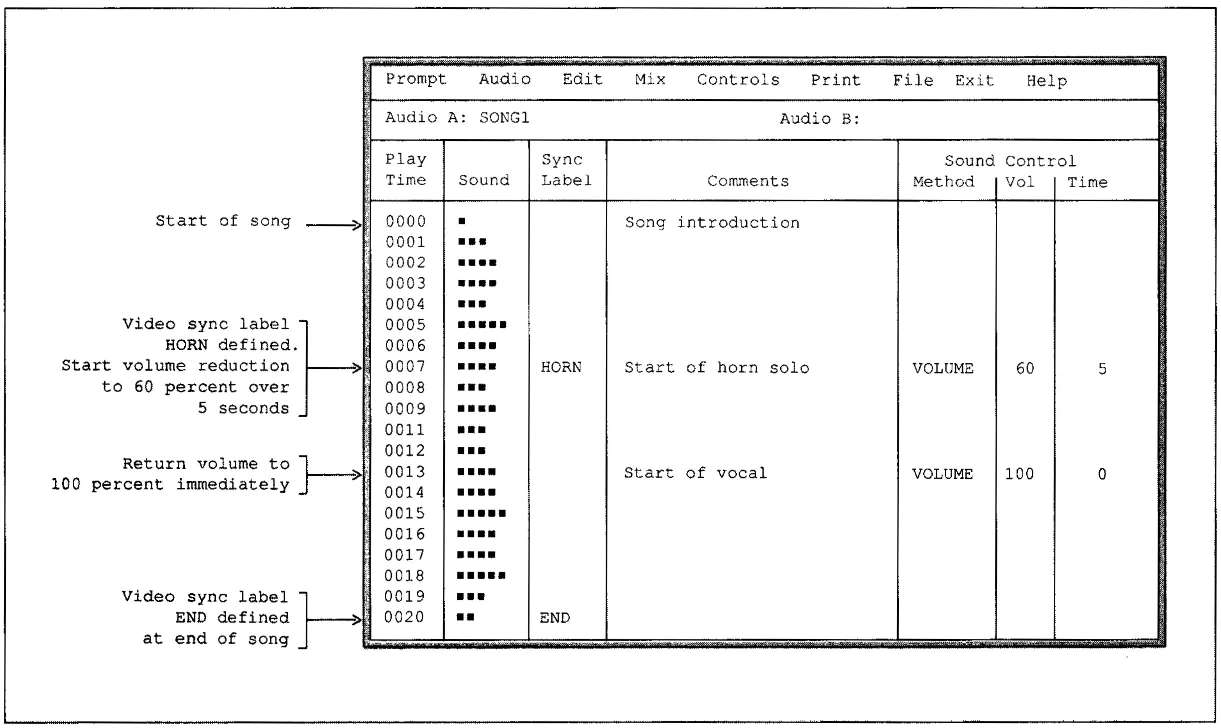


Table 2 Image storage modes and full screen image storage requirements

Attribute	RGB Mode	8514/A Mode	VGA Mode	VGA-8 Mode	MCGA Mode
Image resolution	640 pels × 480 lines	640 pels × 480 lines	640 pels × 480 lines	360 pels × 480 lines	320 pels × 200 lines
Color resolution	16 bit (65,536 colors)	8 bit (256 colors)	4 bit (16 colors or grey shades)	8 bit (256 colors)	8 bit (256 colors)
Maximum storage	615,000 bytes	310,000 bytes	155,000 bytes	155,000 bytes	65,000 bytes
Compressed storage	No compression	180,000 bytes	100,000 bytes	100,000 bytes	50,000 bytes

After the audio has been recorded, the multimedia author has the option of displaying time in 1-second or 0.1-second increments. One or more of these time increments of sound can be copied, moved, inserted, and deleted in a manner analogous to a word in a word processor. If a recording has a click, the sound increment containing the click may be deleted. If a speaker mispronounces a word, the proper pronunciation may be recorded later and overlaid on the original recording. If a speech is too long, portions may be deleted to bring it within the allotted time. If there is an audio file of music and one of speech, the two may be mixed to form a new file.

In Figure 6, the multimedia author entered the two names in the Sync Label column as video synchronization labels to sync video to the audio. The HORN label is used as an example later in the section on the AVC software story editor. Audio synchronization may be defined to allow two audio files to be aligned during playback. One audio file might contain music while the other contains speech. In this way, multiple language versions of a presentation may be produced by substituting the speech file for one in a different language. The multimedia author uses the fourth column for comments and the remaining columns for sound control, as shown.

Image digitize. To capture an image, the multimedia author connects a video camera and a video monitor to the VCA. The developer then watches the image transmitted to the monitor by the camera, and, with the press of the computer's key, snaps the picture when ready. Once an image has been digitized by the VCA, it can be stored in one of five modes within the computer: RGB, 8514/A, VGA, VGA-8, or MCGA. Table 2 shows the characteristics of the image storage modes. RGB is the format of the incoming data from the VCA, which may be saved as an *image master*, because it contains the maximum resolution and color. RGB cannot be displayed on a PS/2 screen. To be used in a multimedia application, RGB must be converted to one of the PS/2 display modes.

The non-RGB modes in Table 2 involve data compression, where the compression rate varies from image to image. The compressed-storage values shown in Table 2 are typical, but the actual storage is larger for very complex images and smaller for partial images or images with large areas of solid color. The 8514/A mode produces the best resolution and color, but requires an 8514/A adapter within the PS/2. The remaining modes provide tradeoffs among storage requirements, image resolution, and number of colors.

A computer display is limited in the number of distinct colors that can be displayed concurrently. For PS/2 color displays, up to 256 different colors may be displayed concurrently, although the colors selected may come from a possible selection of 65 536. Because the 16-bit RGB data can produce 65 536 (i.e., 2^{16}) colors, the AVC must carefully select its colors to produce a pleasing rendition of the original image and yet not exceed the color limit. This selected set of colors is known as a *palette* and is a fundamental concern with the display of images. There are three types of AVC palette: standard, custom, and fixed. The task of palette definition is more difficult when several images are displayed on a screen at one time, because the number of colors selected for all the displayed images must still adhere to the 256 color limit. In addition, one color is set aside as the *transparency* color to control the mixing of images and a set of colors is often reserved for text. (See the section on the story editor.)

A *standard palette* is a predetermined set of colors that provide a best fit to a wide range of images. Although seldom optimum for any given image, the standard palette achieves an excellent rendering of many images and, because all images use the same set of colors, multiple images may be displayed concurrently. With the AVC, one useful 256 color standard palette is composed of 198 colors for image, 56 colors for text, and one color as the transparency color. One color is not used.

A *custom palette* is an AVC-selected set of colors for a specific image. Because a custom palette is optimized for each image, an excellent rendition is achieved for almost all images. One concern exists when multiple images are displayed concurrently. If the number of colors in all the displayed images exceeds the color limit of the screen, colors in certain images may be changed. This color change is called a *palette shift* and produces incorrect colors in some of the images. To avoid a palette shift, the user restricts the number of colors available for each image and allows the AVC to select a unique custom palette for each image. One useful 256 color custom palette allows two images to be concurrently displayed with 99 colors per image plus 56 colors for text, and one color as the transparency color.

A *fixed palette* is similar to a custom palette except that the multimedia developer may exert control over the selection of colors for an image. To achieve the best possible rendering of an image, a developer may initially use the AVC to select either a standard or custom palette, convert this palette to a fixed palette, and then hand-select certain colors. To achieve a special effect, such as an image composed of hues of one color, a developer selects a set of colors from the 65 536 possible colors, and the AVC generates an image using only these colors.

Image convert. An image stored in a given image mode may be converted to any other image mode. The multimedia developer must realize that when an image is converted to a mode with higher resolution, the lost resolution in the first image cannot be recouped. A prudent style would be to save images in the maximum resolution, the RGB mode, and then convert to the particular display modes as needed.

Besides the image editor and VCA, there are two other methods to obtain AVC images, and they do not require a VCA. The first method involves the conversion of files into AVC format from selected image scanner or non-IBM video capture or image generating systems. The second method captures a screen image produced by another software application.

Image editor. It is possible to improvise with the image. The multimedia author can copy any part of an image and paste it elsewhere, either as it was copied or with variations. Since the image is in digital form, an image copy is an exact duplicate of the original image without any loss of clarity. The AVC can resize the image, either keeping the original height-to-width ratio in a larger or smaller image or changing this aspect ratio. The *MIRROR* function may change the direction someone appears to be facing. The *ROTATE* function displays an image at various rotations. To create a drawing or modify an image, tools are available to draw a box or ellipse, sketch a line or modify a pel, and fill or erase an area. For detailed editing, a zoom magnification of up to 64 times is available.

As mentioned in the discussion of image capture palette, a large number of colors may be allocated to text; 56 colors are used in the example given. This allocation allows the AVC to generate very high-resolution text by a technique called *antialiasing* (AA). Video alias sampling errors resemble stair steps imposed on edges at certain angles. With video, the outline of a text character has an infinite resolution, but a video display resolution is limited to its pel size. In one 56 text-color mode, the AVC allows eight sets of seven text colors. A text character uses any one of these sets where the body of the character is one color and the remaining six colors form a gradual series of hues from the character's body color to the color of the image behind the text. The result is an apparent reduction of the jaggedness (i.e., aliasing) resulting from the diagonal or curved edges of a text character. The text appears crisper than is usual for a computer screen.

Text may be outlined and shadowed. Outlining produces a highlighted outline around the character for emphasis. Shadowing produces a shadow for the character to give it a three-dimensional effect.

Two special types of area on the screen may be defined by the multimedia author to provide an interactive environment for the user.

Trigger fields denote an area of the screen that, when selected by the user, cause the AVC story to perform a specific task. For example, a real-estate application screen may show the images of nine homes with each home defined as a trigger field. A user who

Figure 7 AVC story editor screen

Prompt Tell Edit View Print File Exit Help

Story: AVCDEMO

AVA Statement	Screen--> / Sound--->	Method / Method	Direction / Channel	Line / Volume	Time / Time	Wait / Wait	Size / Size	Image X,Y / Sound Beg	Width, Height / Sound End	Screen X,Y
PLAY SONG1		FADEIN	A	100	1	0	FULL			
SHOW FLOWER1		EXPLODE	OUT_H	NONE	1.4	0.4	PART	0, 0	30, 20	0, 0
SHOW FLOWER1		EXPLODE	OUT_H	NONE	1.4	KEY	PART	0, 0	30, 20	0, 20
PLAY SONG2		NONE	A	100	0	0	FULL			
SHOW CLOUDS		CHECKER	VERT	NONE	0.5	HORN	FULL			
PASTE BIRDS		FADEIN			0	0	FULL			
CLEAR BLUE		FADEOUT			2	END	FULL			

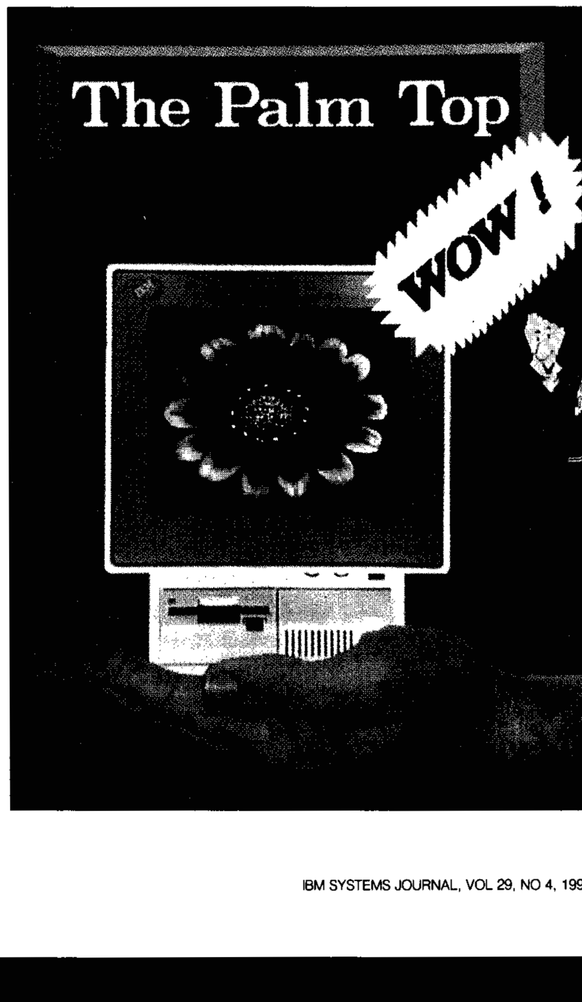
wants more information on one of the homes moves the cursor to where the image of the desired house is located. The story detects that a specific trigger field is selected and initiates further displays of the interior of the home, overlays current price and purchasing data, and plays an audio file describing the attributes of the home. The multimedia developer can use an almost unlimited number of trigger fields throughout the presentation, thereby giving the user a hypertext-like capability.

Data fields denote area of the screen where the user may enter data or the AVC may display data. For example, a self-teaching application may pose a question. The user enters an answer in a data field of the screen. The AVC reads the answer and responds accordingly.

Story editor. The AVC story is written in the audio visual authoring (AVA) language,[5] which is a variation of REXX, the IBM procedural language for Systems Application Architecture™ (SAA™). Figure 7 is an example of a story. The AVA Statement column lists the AVA command and the file that it affects. The PLAY command plays an audio file. The example uses the audio sync label HORN of Figure 6. The SHOW command displays one image while the PASTE command displays a second image over the first image. Most AVA commands have parameters such as those that define the type of transition between images (e.g., fade, explode, dissolve), the part of an image to be displayed (in percent of screen), the speed at which a command is to be processed, and the time or condition at which to continue on to the next command. Standard decision logic (e.g., IF, THEN, ELSE, DO, WHILE), mathematical calculations (e.g., ADD, SUBTRACT, MULTIPLY, DIVIDE, MAXIMUM, MINIMUM), and file management (e.g., LOAD, SAVE, CALL, RETURN) are available.

Following is the explanation of the AVC story editor screen shown in Figure 7.

Figure 8 An AVC-created image



- Line 1. Start audio of SONG1, fade volume in from silence to 100 percent in 1 second, go to line 2 immediately.
- Line 2. Display partial image of FLOWER1 by exploding image outward horizontally in 1.4 seconds, wait 0.4 seconds until going to line 3. Partial image is 0 percent to 30 percent of horizontal image and 0 percent to 20 percent of vertical image positioned at the top-left corner (0 percent, 0 percent) of screen.
- Line 3. Display partial image of FLOWER1 by exploding image outward horizontally in 1.4 seconds, wait until any key or mouse button is depressed before going to line 4. The same partial image of line 2 (0 percent to 30 percent horizontal, 0 percent to 20 percent vertical) is displayed below the first image (0 percent horizontal, 20 percent vertical).
- Line 4. Start audio of SONG2, volume is immediately at 100 percent, go to line 5 immediately.
- Line 5. Display full image of CLOUDS by fading in with a checkerboard pattern in 0.5 seconds, wait until SONG2 audio sync label, HORN, is reached before continuing with line 6.
- Line 6. Display image of BIRDS only where this image is not the transparency color. The screen now shows an image of birds superimposed on the image of the clouds.
- Line 7. Fade CLOUDS/BIRDS image to a blue screen in 2 seconds when the SONG2 audio sync label, END, is reached.

Two methods may be used to produce animation. The first method implies motion by such a mechanism as that of gradually displaying an arrow from its tail to its head in several seconds or successively redisplaying the image of a bird across the screen. The second method rapidly displays a series of overlapping partial images resulting in a motion picture effect.

Figure 8 illustrates a picture of an imaginary product that was created using several AVC techniques. Refer to the sidebar for a description of how the AVC was used to create this image. Do not believe the old adage that the camera does not lie; it can be made to lie!

Concluding remarks

Multimedia offers computer users new and expanded opportunities to participate in the learning experience. As with any medium, the quality of the data is of utmost importance. Useless information, no matter how attractively presented, is still useless information. The multimedia author must ensure that a presentation is pertinent to the needs and level of understanding of the user, and the information must be accurate and up-to-date. Multimedia provides a tool to improve communication. As with a conventional presentation, the multimedia author retains responsibility for the actual quality of that communication.

Just as text written by others may not be used without permission (and possibly payment), images and audio produced by others must be given the same respect. Generally speaking, any audio or image in a copyrighted form (e.g., record, compact disc, magazine, television) may not be used without permission. Multimedia authors should inform themselves of sources of permission (e.g., American Society of Composers and Publishers [ASCAP]) and legal requirements for obtaining permission, acknowledgment, and payment.

HOW FIGURE 8 WAS MADE

The three-dimensional sign at the top of the picture was created with Audio Visual Connection (AVC) area drawing tools and filled with contrasting colors. The text was created with a black shadow for emphasis.

The cartoon character was copied from the AVC Help presentation, which had been created with AVC sketching tools. The cartoon bubble was also created with these tools and filled with the color white. The text for the bubble was created using AVC text tools and rotated to match the angle of the bubble.

The image of the flower was copied from a previous AVC presentation, but the image had to be stretched horizontally to fit the proportions of the computer screen. Editing at the pel level improved the definition of the flower's highlights.

The hand and the PS/2 Model 25 were individually photographed with a video camera connected to the Video Capture Adapter. Once captured, each AVC image was edited to substitute the transparency color for the background image. The hand image was then superimposed over the image of the computer, creating the impression of a palm-sized computer.

Finally, the remaining transparency-colored pels were converted to a blue background and the entire image was processed by the AVC to enhance clarity.

Acknowledgments

Early investigation of audio and video concepts by Glenn Henry paved the way toward the development of many of the principles described in this paper. Actual product development of the Audio Visual Connection was led by Brad Beitel and Wayne Galella, and the quality of the product is a direct reflection of their teams of people. I wish to express my appreciation to everyone who provided assistance in writing this paper, especially Christy Conway, Ron Lisle, and Scott McDonald.

Audio Visual Connection, InfoWindow, Personal System/2, PS/2, OS/2, and Micro Channel are registered trademarks, and AVC, Storyboard Plus, LinkWay, Operating System/2, KnowledgeTool, Systems Application Architecture, and SAA are trademarks, of International Business Machines Corporation.

80286 is a trademark of Intel Corporation.

Cited references

1. V. Bush, "As We May Think," *The Atlantic Monthly*, 101–108 (July 1945).
2. *Electronics Engineers' Handbook*, D. G. Fink and D. Christiansen, Editors, McGraw-Hill Book Company, New York, NY (1989), pp. 2–32.
3. A. C. Luther, *Digital Video in the PC Environment*, McGraw-Hill Book Company, New York, NY (1989).
4. *Audio Visual Connection User's Guide*, S15F-7133, IBM Corporation; available through IBM branch offices.
5. *Audio Visual Authoring Language Reference for the Audio Visual Connection*, S15F-7134, IBM Corporation; available through IBM branch offices.

Daniel J. Moore *IBM Entry Systems Division, Austin Laboratory, 11400 Burnet Road, Austin, Texas 78758.* Mr. Moore is an advisory engineer in the Multimedia Systems group at the Austin Laboratory. He joined IBM in 1962 at the Federal Systems Division facility in Owego, New York, where he worked on military applications for the federal government. Since then, he has been involved with personal computer applications at the Austin Laboratory in the fields of telecommunication and electronic music. Mr. Moore was part of the development team for the IBM Music Feature, the first IBM electronic music product, and has represented IBM at numerous music forums. He is the author of one previous paper in the *IBM Systems Journal* on Teletex telecommunication and is currently investigating future audio needs. Mr. Moore received his B.S. degree in electrical engineering from Clarkson College of Technology, Potsdam, New York, in 1962.

Reprint Order No. G321-5414.