HIGH DEFINITION TELEVISION

A high-definition television (HDTV) system is one whose performance is significantly better than a conventional television system. An HDTV system delivers spectacular video and multichannel compact disc-(CD) quality sound. The system also has many features lacking in conventional systems, such as auxiliary data channels and easy interoperability with computers and telecommunications networks.

Conventional television systems were developed during the 1940s and 1950s. Examples of conventional systems are National Television Systems Committee (NTSC), Sequential Couleur a Memoire (SECAM) and Phase Alternating Line (PAL). These systems are comparable in video quality, audio quality, and transmission robustness. The NTSC system, used in North America, is used as a reference for conventional television systems when discussing HDTV in this article.

For many decades, conventional television systems have been quite successful. However, they were developed with the technology available during the 1940s and 1950s. Advances in technologies, such as communications, signal processing, and very large scale integration (VLSI) have enabled a major redesign with substantial improvements over conventional television systems. An HDTV system is one result of this technological revolution.

CHARACTERISTICS OF HDTV

The many characteristics of an HDTV system that markedly differ from a conventional television system are described in this section.

High Resolution

An HDTV system delivers video with spatial resolution much higher than a conventional television system. Typically, video with a spatial resolution of at least four times that of a conventional television system is called high-resolution video. Resolution represents the amount of detail contained within the video, which is also called "definition." This is the basis for high-definition television. An NTSC system delivers video at a resolution of approximately 480 lines in an interlaced format at an approximate rate of 60 fields/s (it is actually 59.94 Hz, but we will not distinguish between 59.94 and 60). Each line contains approximately 420 pixels or picture elements. The number of lines represents the vertical spatial resolution in the picture, and the number of pixels per line represents the horizontal spatial resolution. Interlaced scanning refers to the scanning format. All conventional television systems use this format. Television systems deliver pictures that are snapshots of a scene recorded a certain number of times per second. In interlaced scanning, a single snapshot consists of only odd lines, the next snapshot consists of only even lines, and this sequence repeats. A snapshot in interlaced scanning is called a field. In the NTSC system, 60 fields are used per second. Although only snapshots of a scene are shown, the human visual system perceives this as continuous motion, as long as the snapshots are shown at a sufficiently high rate. In this way, the video provides accurate motion rendition.

More lines and more pixels per line in a field provide more spatial details that the field can retain. An HDTV system may have 1080 lines and 1920 pixels/line resolution in an interlaced format of 60 fields/s. In this case, the spatial resolution of an HDTV system would be almost ten times that of an NTSC system. This high spatial resolution is capable of showing details in the picture much more clearly, and the resultant video appears much sharper. It is particularly useful for sports events, graphic material, written letters, and movies.

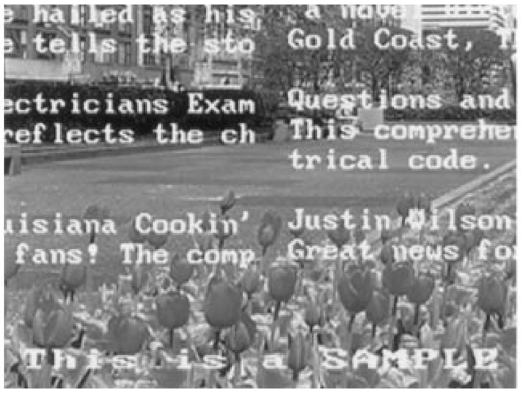
The high spatial resolution in an HDTV system enables a large-screen display and increased realism. For an NTSC system, the spatial resolution is not high. To avoid the visibility of a line structure in an NTSC system, the recommended viewing distance is approximately seven times the picture height. For a two-foot-high display screen, the recommended viewing distance from the screen is 14 feet, seven times the picture height. This makes it difficult to have large-screen television receivers in many homes. Because of the long viewing distance, the viewing angle is approximately 10°, which limits realism. For an HDTV system with more than twice the number of lines, the recommended viewing distance is typically three times the picture height. For a two-foot-high display, the recommended viewing distance is six feet. This can accommodate a large-screen display in many environments. Because of the short viewing distance and wider aspect (width-to-height) ratio, the viewing angle for an HDTV system is approximately 30°, which significantly increases realism.

An HDTV system also delivers higher temporal resolution by using progressive scanning. Unlike interlaced scanning, where a snapshot (field) consists of only even lines or only odd lines, all of the lines in progressive scanning are scanned for each snapshot. The snapshot in progressive scanning is called a frame. Both progressive scanning and interlaced scanning have their own merits. The choice between the two generated much discussion during the digital television standardization process in the United States. An HDTV system can have only interlaced scanning, progressive scanning, or a combination of the two.

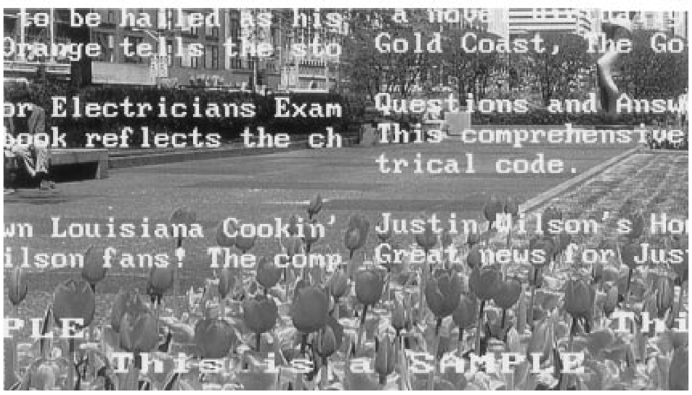
An HDTV system delivers video with substantially higher spatial and temporal resolution than a conventional television system. In addition to its superior resolution, an HDTV system typically has other important features discussed here.

Wide Aspect Ratio

An NTSC television receiver has a display area with an aspect ratio of 4 : 3. The aspect ratio is a ratio of width to height. The 4 : 3 aspect ratio was chosen because movies were made with a 4 : 3 aspect ratio when the NTSC system was first developed. Since then, movies have been made with a wider aspect ratio. To reflect this change, an HDTV system typically has a wider 16 : 9 aspect ratio. The difference in spatial resolution and aspect ratio between an NTSC system and an HDTV system is illustrated in Fig. 1. Figure 1(a) is a frame with an aspect ratio of 4 : 3 and Fig. 1(b) is a frame with an aspect ratio of 16 : 9. The difference in the spatial details be-



(**a**)



(**b**)

Figure 1. Resolution and aspect ratio of a conventional and a high-definition television system: (a) a segment of a conventional television video frame; (b) the corresponding segment of a high-definition television video frame.

tween the two pictures is approximately the difference in spatial resolution between a conventional television and HDTV.

Digital Representation and Transmission

In a conventional television system, the video signal is represented in an analog format, and the analog representation is transmitted. However, the analog representation is highly susceptible to channel transmission degradations, such as multipath effects or random noise. In a conventional television system, video received through the air (terrestrial broadcasting) often has visible degradations, such as ghosts and snowlike noise.

In a digital HDTV system, the video signal is represented digitally and transmitted by modern digital communication technology. The effects of channel degradation manifest themselves differently in a digital transmission system. In an HDTV system broadcast digitally over the air, the video received is essentially perfect within a certain coverage area (within a certain level of channel degradation). Outside that area, the video is not viewable. Unlike an analog NTSC system, where the video degrades gradually as the channel degradation increases, a digital HDTV system delivers either an essentially perfect picture or no picture at all. This is called the *cliff effect* or *digitally clean* video.

Multichannel Digital Audio

An HDTV system can deliver multichannel sound. The number of audio channels that accompany a video program may be as many as one desires. Multiple audio channels can be used to produce the effect of surround sound, often used in movie theaters, and for transmitting different languages in the same video program. In addition to multichannel sound, the reproduced sound has the quality of an audio compact disc (CD).

A television system is often considered primarily a video service. However, audio service is particularly important for HDTV applications. Generally, people will not watch video with poor-quality audio, even when the video quality is similar to HDTV. In addition, high-quality audio enhances our visual experience. The same video, accompanied by higher quality audio gives the impression of higher quality video than when it is accompanied by low-quality audio. An HDTV system delivers multichannel audio with CD-quality sound. In addition to a superb listening experience, it enhances our visual experience beyond what is possible with high-resolution video alone.

Data Channel

A conventional television system is a stand-alone system whose primary objective is entertainment. A digital HDTV system utilizes a data transmission channel. Its data can represent high-resolution video and audio and also any digital data like computer data, newspapers, telephone books, and stock market quotes. The digital HDTV system can be integrated easily to operate with computers and telecommunication networks.

HISTORY OF HDTV IN THE UNITED STATES

Since the development and deployment of the conventional television system, improvements have been carried out. One major improvement was an increase in the picture resolution. An early development effort to increase picture resolution by approximately a factor of 5 was made in the late 1960s by NHK, Japan's government-sponsored television broadcaster. Together with Japan's electronics manufacturers, NHK developed an analog system called multiple sub-Nyquist encoding (MUSE). Although the system performs some signal compression, it was simply an extension of a conventional television system. Originally, the NTSC system was delivered over a 6 MHz channel, the amount of spectrum bandwidth needed to deliver one channel of NTSC video and audio. To increase the amount of information delivered by a factor of 5, the MUSE system required approximately 30 MHz, five times the bandwidth. The MUSE system was considered for satellite transmissions to consumers because of the large bandwidth required.

The NTSC system was developed for terrestrial transmission of television signals. Because the NTSC system requires 6 MHz of bandwidth, the available very high-frequency (VHF) and ultrahigh-frequency (UHF) bands, suitable for terrestrial broadcasting of television signals, were divided into 6 MHz channels. Initially, there was plenty of spectrum. The NTSC system, however, utilizes its given 6 MHz of spectrum quite inefficiently. This inefficiency generates interference among the different NTSC signals. As the number of NTSC signals broadcast terrestrially increased, the interference problem became serious. The solution was not to use some channels. These unused channels are known as taboo channels. In a typical highly populated geographical location in the United States, only one of two VHF channels is used and only one of six UHF channels is used. In addition, in the 1980s, other services, such as mobile radio, requested the use of the UHF band spectrum. As a result, an HDTV system that requires a large amount of bandwidth, such as Japan's MUSE system, was not an acceptable solution for terrestrial broadcasting in the United States.

At the request of the broadcast organizations, the United States Federal Communications Commission (FCC) created the Advisory Committee on Advanced Television Service (ACATS) in September 1987. ACATS was chartered to advise the FCC on matters related to standardizing advanced television service in the United States, including establishment of a technical standard. At the request of ACATS, industries, universities, and research laboratories submitted proposals for the advanced television (ATV) technical standard in 1988.

While the ACATS screened the proposals and prepared testing laboratories for their formal technical evaluation, the FCC made a key decision. In March 1990, the FCC selected the simulcast approach for advanced television service rather than the receiver-compatible approach. In the receiver-compatible approach, existing NTSC television receivers can receive an HDTV signal and generate a viewable picture. This was the approach taken when the NTSC introduced color. A black-and-white television receiver can receive a color television signal and display it as a viewable black-and-white picture. In this way, the existing black-and-white television receivers would not become obsolete. It was possible to use the receiver-compatible approach for color introduction because color information did not require a large amount of bandwidth and a small portion of the 6-MHz channel used for a black-and-white picture could be used to insert the color information without seriously affecting the black-and-white picture.

In HDTV, the additional information needed was much more than the original NTSC signal, and the receiver-compatibility requirement would require an additional spectrum to carry the HDTV. Among the proposals received, the receiver-compatible approaches typically required an additional 6 MHz augmentation channel that carried the enhancement information, which was the difference between the HDTV signal and the NTSC signal. Even though the augmentation approach solves the receiver-compatibility requirement, it has several major problems. The approach requires an NTSC channel to transmit an HDTV signal. This means that the highly spectrum-inefficient NTSC system cannot be converted into a more efficient technical system. In addition, the introduction of HDTV would permanently require a new channel for each existing NTSC channel. The FCC rejected this spectrum-inefficient augmentation channel approach.

Although the FCC's decision did not require receiver compatibility, it did require transmission of an entire HDTV signal within a single 6 MHz channel. In the simulcast approach adopted by the FCC, an HDTV signal that can be transmitted in a single 6 MHz channel can be designed independently of the NTSC signal. An NTSC television receiver cannot receive an HDTV signal. To receive an HDTV signal, a new television receiver would be needed. To ensure that existing television receivers do not become obsolete when HDTV service is introduced, the FCC would give one new channel for HDTV service to each NTSC station requesting it. During the transition period, both NTSC and HDTV services will coexist. After sufficient penetration of HDTV service, NTSC service will be discontinued. The spectrum previously occupied by NTSC services will be used for additional HDTV channels or for other services. Initially, the FCC envisioned that the new HDTV channel and the existing NTSC channel would carry the same programs, so as not to disadvantage NTSC receivers during the transition period. This is the basis for the term *simulcasting*. Later, this requirement was removed. The simulcast approach is illustrated in Fig. 2.

The simulcast approach provides several major advantages. It presents the possibility of designing a new spectrum-efficient HDTV signal that requires significantly less power and does not seriously interfere with other signals, including the NTSC signal. This allows the use of the taboo channels, which could not be used for additional NTSC service because of the strong interference characteristics of the NTSC signals. Without the taboo channels, it would not have been possible to give an additional channel to each existing NTSC broadcaster for HDTV service. In addition, it eliminates the spectrum-inefficient NTSC channels following the transition period. The elimination of NTSC broadcasting vacates the spectrum that it occupied. Furthermore, by removing the NTSC signals that have strong interference characteristics, other channels could be used more efficiently. The 1990 FCC ruling was a key decision in the process to standardize the HDTV system in the United States.

The 1990 decision also created several technical challenges. The HDTV signal had to be transmitted in a single, 6 MHz channel. In addition, the signal was required to produce minimal interference with NTSC signals and other HDTV signals. At the time of the FCC's decision in 1990, it was not

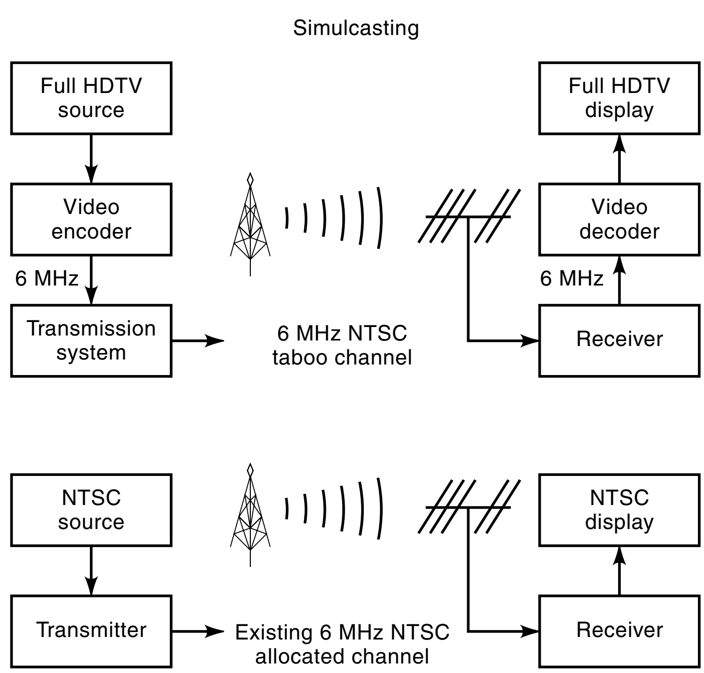


Figure 2. Illustration of simulcasting approach for transition from an NTSC system to a digital high-definition television system.

clear that such a system could be developed within a reasonable time. Later events proved that developing such a system at a reasonable cost to broadcasters and consumers was possible using modern communications, signal processing, and VLSI technologies.

Before the formal technical evaluation of the initial HDTV proposals began, some were eliminated, others were substantially modified, and still others combined their efforts. Five HDTV system proposals were ultimately approved for formal evaluation. One proposed an analog system whereas four others proposed all-digital systems. The five systems were evaluated in laboratory tests at the Advanced Television Testing Center (ATTC) in Alexandria, Virginia. Subjective evaluation of picture quality was performed at the Advanced Television Evaluation Laboratory (ATEL) in Ottawa, Canada. In February 1993, a special panel of experts reviewed the test results of the five HDTV system proposals and made a recommendation to the ACATS. The panel concluded that the four digital systems performed substantially better than the analog system. The panel also concluded that each of the four digital systems excelled in different aspects. Therefore, the panel could not recommend one particular system. The panel recommended that each digital system be retested after improvements were made by the proponents. The four digital proponents had stated earlier that substantial improvements could be made to their respective system. The ACATS accepted the panel's recommendation and decided to retest the four systems after improvements were made. As an alternative to the retest, the ACATS encouraged the four proponents to combine the best elements of the different systems and submit one single system for evaluation.

The four digital system proponents evaluated their options and decided to submit a single system. In May 1993, they formed a consortium called the Grand Alliance to design and construct an HDTV prototype system. The Grand Alliance was composed of seven organizations (the organizations listed are those who were members at the inception of the Grand Alliance. Later, some member organizations changed their names): General Instrument (first proposed digital transmission of an HDTV signal and submitted one of the four initial systems); Massachusetts Institute of Technology (submitted a system together with General Instrument); AT&T and Zenith (submitted one system together); and Philips, the David Sarnoff Research Center, and Thomson Consumer Electronics (submitted one system together). Between 1993 and 1994, the Grand Alliance chose the best technical elements from the four systems and made further improvements on them. The Grand Alliance HDTV system was submitted to the ATTC and ATEL for performance verification. Test results verified that the Grand Alliance system performed better than the previous four digital systems. A technical standard based on the Grand Alliance HDTV prototype system was documented by the Advanced Television System Committee (ATSC), an industry consortium.

The HDTV prototype proposed by the Grand Alliance was a flexible system that carried approximately 19.4 million bits per second (19.4 Mbits/s). Even though it used the available bit capacity to transmit one HDTV program, the bit capacity could also be used to transmit several programs of standard-definition television (SDTV) or other digital data, such as stock quotes. SDTV resolution is comparable to that of the NTSC, but it is substantially less than the HDTV. The documented technical standard (known as the ATSC standard) allowed the transmission of SDTV programs and HDTV programs.

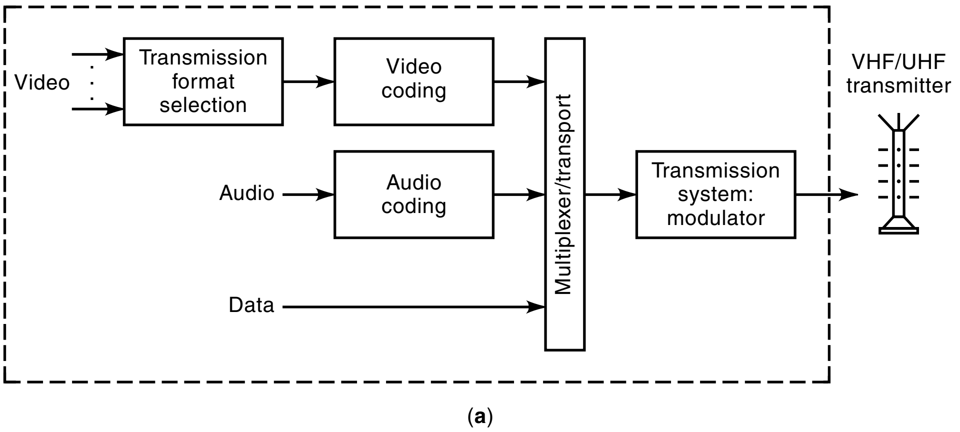
In November 1995, the ACATS recommended the ATSC standard as the U.S. advanced television standard to the FCC. The ATSC standard had allowed a set of only 18 video resolution formats for HDTV and SDTV programs. The FCC eased this restriction in December 1996 and decided that the ATSC standard with a relaxation of the requirements for video resolution format would be the US digital television standard. In early 1997, the FCC made additional rulings to support the new technical standard, such as channel allocation for digital television service.

GRAND ALLIANCE HDTV SYSTEM

A block diagram of a typical HDTV system is shown in Fig. 3. The information transmitted includes video, audio, and other auxiliary data, such as stock quotes. The input video source may have a format (spatial resolution, temporal resolution, or scanning format) different from the formats used or preferred by the video encoder. In this case, the input video format is converted to a format used or preferred by the video encoder. Then the video is compressed by a video encoder. Compression is needed because the bit rate supported by the modulation system is typically much less than the bit rate needed for digital video without compression. The audio, which may be multichannel for one video program, is also compressed. Because the bit rate required for audio is much less than that for video, the need to compress the audio is not as crucial. Any bit-rate savings, however, can be used for additional bits for video or other auxiliary data. The data may represent any digital data, including additional information for video and audio. The compressed video data, compressed audio data, and any other data are multiplexed by a transport system. The resulting bit stream is modulated. Then the modulated signal is transmitted over a communication channel.

At the receiver, the received signal is demodulated to generate a bit stream, which is demultiplexed to produce compressed video, compressed audio, and other data. Then the compressed video is decompressed. The video format received may not be the same as the format used in the display. In this case, the received video format is converted to the proper display format. The compressed audio, which may be multichannel, is decompressed and distributed to different speakers. The use of the data received depends on the type of information that the data contains. The communication channel in Fig. 3 may also represent a storage device, such as digital video disc. If the available bit rate can support more than one video program, multiple video programs can be transmitted.

There are many different possibilities for the design of an HDTV system. For example, there are various methods used for video compression, audio compression, and modulation. Some modulation methods may be more suitable for terrestrial transmission, whereas others may be more suitable for satellite transmission. Among the many possibilities, this article focuses on the Grand Alliance's HDTV system. This system was designed over many years of industry competition and cooperation. The system's performance was carefully evaluated by laboratory and field tests and was judged acceptable for its intended application. The system was the basis for



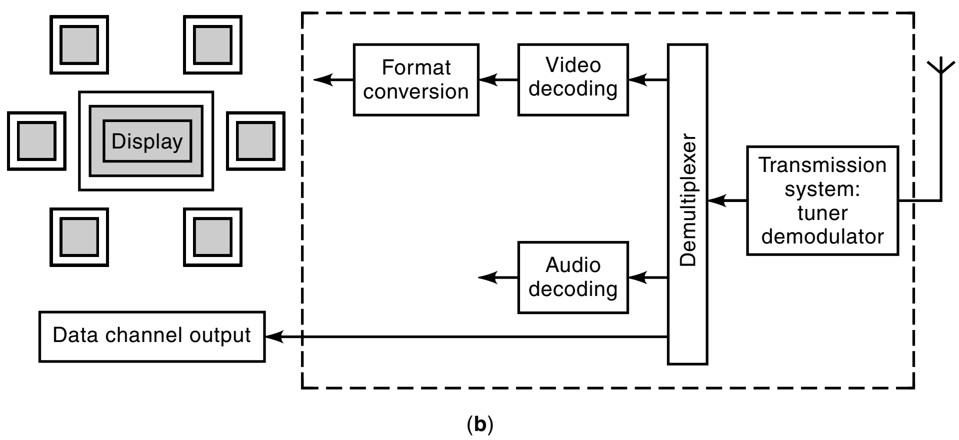


Figure 3. A block diagram of a typical HDTV system: (a) transmitter; (b) receiver. Reprinted with permission from IEEE (© 1995 IEEE).

the US digital television standard. Even though this article focuses on one system, many issues and design considerations encountered in the Grand Alliance HDTV system could be applied to any HDTV system.

The overall Grand Alliance HDTV system consists of five elements: transmission format selection, video coding, audio coding, multiplexing, and modulation. These are described in the following sections.

Transmission Format Selection

A television system accommodates many video input sources, such as video cameras, film, magnetic and optical media, and synthetic imagery. Even though these different input sources have different video formats, a conventional television system, such as the NTSC, uses only one single transmission format. This means that the various input sources are converted to one format and then transmitted. Using one format simplifies the receiver design because a receiver can eliminate format conversion by designating that the display format is the same as the transmission format. This is shown in Fig. 4. When the NTSC system was standardized in the 1940s and 1950s, format conversion would have been costly.

The disadvantage of using one transmission format is the inefficient use of the available spectrum, because all video input sources must be converted to one format and then transmitted in that format. For example, in the NTSC system, film (whose native format is 24 frames/s with progressive scanning) is converted to 60 fields/s with interlaced scanning. Then it is transmitted in the NTSC format. Transmission of video in a format other than its native format is an inefficient use of the spectrum.

The Grand Alliance HDTV system utilizes multiple transmission formats. This allows using of a video transmission

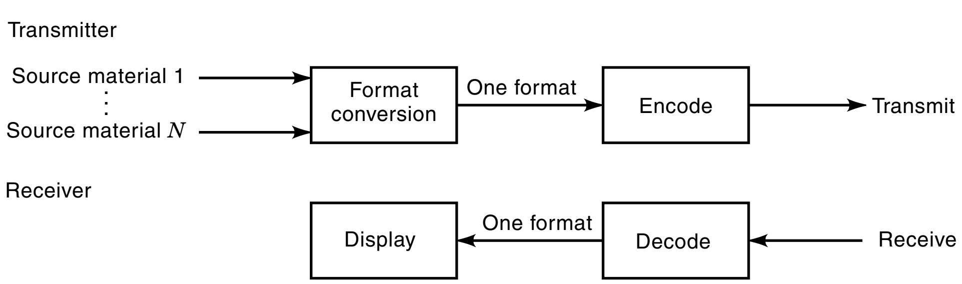


Figure 4. A television system with one single video transmission format.

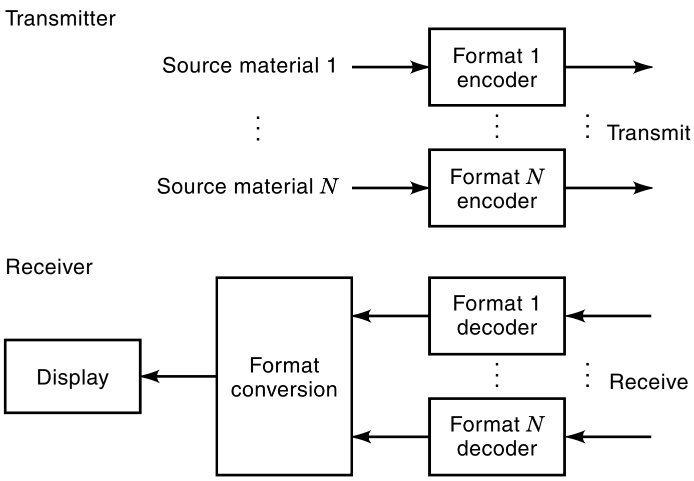


Figure 5. A television system with multiple video transmission formats.

format identical to or approximating the native video source format. In addition, the system allows the use of different formats for various applications. From the viewpoint of spectrum efficiency, allowing all possible video formats would be ideal. Because a display (such as a cathode ray tube) has typically one display format, the different formats received must be converted to one display format, as shown in Fig. 5. Allowing for too many formats complicates format conversion. In addition, most of the benefits derived from multiple formats are obtained by carefully selecting a small set of formats. For HDTV applications, the Grand Alliance system utilizes six transmission formats as shown in Table 1. In the table, the spatial resolution of C × D equals C lines of vertical resolution with D pixels of horizontal resolution. The scanning format may be either a progressive scan or interlaced scan. The frame/field rate refers to the number of frames/s for progressive scan and the number of fields/s for interlaced scan.

The Grand Alliance system utilizes both 720 lines and 1080 lines. The number of pixels per line was chosen so that the aspect ratio (width-to-height ratio) is 16 × 9 with square pixels. When the spatial vertical dimension that corresponds to one line equals the spatial horizontal dimension that corresponds to one pixel, it is called *square pixel.* For 720 lines, the scanning format is progressive. The highest frame rate is 60 frames/s. The pixel rate is approximately 55 Mpixels/s. For the video compression and modulation technologies used, a substantial increase in the pixel rate above 60 to 70 Mpixels/s may result in a noticeable degradation in video quality. At 60 frames/s with progressive scanning, the temporal resolution is very high, and smooth motion rendition results. This format is useful for sports events and commercials. The 720-line format also allows the temporal resolution of 30 frames/s and 24 frames/s. These frame rates were chosen to accommodate film and graphics. For film, whose native format is 24 frames/s, conversion to 60 frames/s rate and then compressing it results in substantial inefficiency in spectrum utilization. For 720 lines at 24 frames/s, it is possible to simultaneously transmit two high-resolution video programs within a 6 MHz channel because of the lower pixel rate (approximately 22 Mpixels/s each).

Table 1. HDTV Transmission Formats Used in the Grand Alliance System

Spatial Resolution	Scanning Format	Frame/Field Rate
720 × 1280	Progressive scanning	60 frames/s
720 × 1280	Progressive scanning	30 frames/s
720 × 1280	Progressive scanning	24 frames/s
1080 × 1920	Progressive scanning	30 frames/s
1080 × 1920	Progressive scanning	24 frames/s
1080 × 1920	Interlaced scanning	60 fields/s

In the 1080-line format, two temporal rates in progressive scan are 30 frames/s and 24 frames/s. These temporal rates were chosen for film and graphics with the highest spatial resolution. Another temporal rate used is the 1080-line interlaced scan at 60 fields/s. This is the only interlaced scan HDTV format used in the Grand Alliance system. It is useful for scenes obtained with a 1080-line, interlaced-scan camera. The pixel rate for 1080-line progressive scan at 60 frames/s would be more than 120 Mpixels/s. The video encoded for such a high pixel rate can result in substantial degradation in video quality for the compression and modulation technologies used in the Grand Alliance system. Therefore, there is no 1080-line, 60 frames/s progressive scan format in the system.

All conventional television systems, such as the NTSC, utilize only interlaced scanning. In such systems, a display format is matched to the single transmission format. The display requires at least a 50 Hz to 60 Hz rate with a reasonable amount of spatial resolution (approximately 480 active lines for the NTSC system). An alternative strategy in the NTSC system would be to preserve 480 lines with progressive scan, but at 30 frames/s. To avoid display flicker, each frame can be repeated twice at the display, making the display rate 60 Hz. Repetition of a frame at the receiver would require a frame memory, which was not possible with the technologies available when the NTSC system was standardized. Because of the exclusive use of interlaced scanning in conventional television systems, early HDTV video equipment, such as video cameras, was developed for interlaced scanning.

An interlaced display has video artifacts like interline flicker. Consider a sharp horizontal line that is in the odd field, but not in the even field. Even though the overall large area flicker rate is 60 Hz, the flicker rate for the sharp horizontal line is only 30 Hz. As a result, the line flickers in a phenomenon called interline flicker. Interline flickers are particularly troublesome for computer graphics or written material that contains many sharp lines. Partly for this reason, almost all computer monitors use progressive scanning.

When a television system is used as a stand-alone entertainment device, its interoperability with computers is not a serious issue. For a digital HDTV system, however, it is no longer a stand-alone entertainment device, and its interoperability with computers and telecommunications networks is useful. When a display device uses progressive scan with an interlaced transmission format, a conversion process called deinterlacing must convert the interlaced scan format to a progressive scan format before it is displayed. A high-performance deinterlacer requires complex signal processing. Even when a high-performance deinterlacer is used, a progressive transmission format yields better performance than an interlaced transmission format for graphics, animation, and writ-

Table 2. Applications of HDTV Transmission Formats Used in the Grand Alliance System

Format	Applications
720 × 1280, PS, 60 frames/s	Sports, concerts, animation, graphics, up-converted NTSC, commercials
720 × 1280, PS, 24 frames/s or 30 frames/s	Complex film scenes, graphics, animation
1080 × 1920, PS, 24 frames/s or 30 frames/s	Films with highest spatial resolution
1080 × 1920, IS, 60 fields/s	Scenes shot with an interlaced-scan camera

Source: Reprinted with permission from IEEE (© 1995 IEEE).

ten material. For this and other reasons like simple processing, the computer industry preferred only the progressive transmission format for television. Other industries, such as television manufacturers and broadcasters, preferred interlaced scanning, which is used in all conventional television systems including the NTSC. Interlaced scanning worked well for entertainment material and HDTV video equipment, including a video camera already developed for interlaced scanning. This disagreement between the industries generated considerable discussion and delayed the digital television standardization process in the United States. The Grand Alliance HDTV system used five progressive scan formats and one interlaced scan format. The FCC decision in December 1996 removed most of the restrictions on transmission formats and allowed both progressive and interlaced formats. This decision left the choice of the transmission format to free market forces.

A multiple transmission format system utilizes the available spectrum more efficiently than a single transmission format system by better accommodating video source materials with different native formats. In addition, multiple transmission formats can be used for various applications. Table 2 shows possible applications for the six Grand Alliance HDTV formats. For a multiple transmission format system, one of the allowed transmission formats is chosen for a given video program before video encoding. The specific choice depends on the native format of the input video material and its intended application. The same video program with a given format may be assigned to a different transmission format, depending on the time of broadcast. If the transmission format chosen differs from the native format of the video material, format conversion occurs. An example of format conversion from an input video with a spatial resolution of 1080 × 1920 progressive scan at 60 frames/s to a transmission format of 720 × 1280 progressive scan at 60 frames/s is shown in Fig. 6. The signal processing operations used in this conversion process, spatial low-pass filtering and subsampling, reduce the spatial resolution.

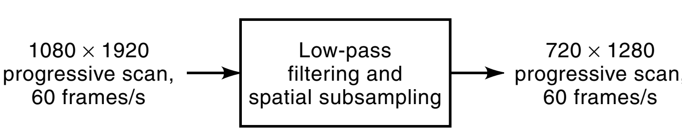


Figure 6. An example of video format conversion for transmission.

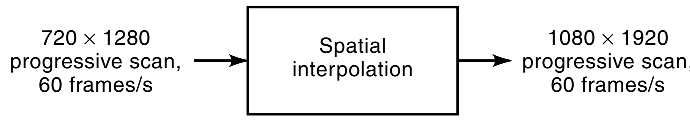


Figure 7. An example of video format conversion for display.

In a multiple transmission format television system, format conversion is also typically required at the receiver. For display devices, such as a cathode ray tube (CRT), only one display format is typically used. To accommodate received video with different formats, conversion is required from the received format to the display format. An example of conversion from a received video format of 720 × 1280 progressive scan at 60 frames/s to the display format of 1080 × 1920 progressive scan at 60 frames/s is shown in Fig. 7. The signal processing operation is spatial interpolation.

Video Coding

The Grand Alliance HDTV system transmits at least one HDTV program in a single 6 MHz channel. For the modulation technology used in the Grand Alliance system, the maximum bit rate available for video is approximately 19 Mbits/sec. For a typical HDTV video input, the bit rate is on the order of 1 Gbits/s. This means that the input video must be compressed by a factor of more than 50.

For example, consider an HDTV video input of 720 × 1280 with progressive scan at 60 frames/s. The pixel rate is 55.296 Mpixels/s. A color picture consists of three monochrome images: red, green, and blue. Red, green, and blue are the three primary colors of an additive color system. By mixing the appropriate amounts of red, green, and blue lights, many different color lights are generated. By mixing a red light and a green light, for example, a yellow light is generated. A pixel of a color picture consists of the red, green, and blue components. Each component is typically represented by eight bits (256 levels) of quantization. For many video applications, such as television, eight bits of quantization are considered sufficient to avoid video quality degradation by quantization. Then each pixel is represented by 24 bits. The bit rate for the video input of 720 × 1280 with progressive scan at 60 frames/s is approximately 1.3 Gbits/s. In this example, reducing the data rate to 19 Mbits/s requires video compression by a factor of 70.

Video compression is achieved by exploiting the redundancy in the video data and the limitations of the human visual system. For a typical video, there is a considerable amount of redundancy. For example, much of the change between two consecutive frames is due to the motion of an object or the camera. Therefore, a considerable amount of similarity exists between the two consecutive frames. Even within the same frame, the pixels in a neighborhood region typically do not vary randomly. By removing the redundancy, the same (redundant) information is not transmitted.

For television applications, the video is displayed for human viewers. Even though the human visual system has enormous capabilities, it has many limitations. For example, the human visual system does not perceive well the spatial details of fast-changing regions. The high spatial resolution in such cases does not need to be preserved. By removing the redundancy in the data and exploiting the limitations of the

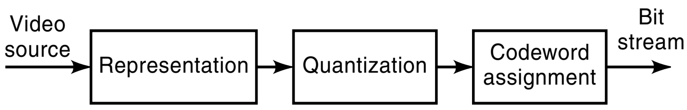


Figure 8. Three basic elements of a video encoder.

human visual system, many methods of digital video compression were developed. A digital video encoder usually consists of the three basic elements shown in Fig. 8. The first element is representation. This element maps the input video to a domain more suitable for subsequent quantization and codeword assignment. The quantization element assigns reconstruction (quantization) levels to the output of the representation element. The codeword assignment assigns specific codewords (a string of zeros and ones) to the reconstruction levels. The three elements work together to reduce the required bit rate by removing the redundancy in the data and exploiting the limitations of the human visual system.

Many different methods exist for each of the three elements in the image coder. The Grand Alliance system utilizes a combination of video compression techniques that conform to the specifications of the Moving Pictures Expert Group (MPEG-2) video compression standard. This is one of many possible approaches to video compression.

MPEG-2 Standard. The International Standard Organization (ISO) established the Moving Pictures Expert Group (MPEG) in 1988. Its mission was to develop video coding standards for moving pictures and associated audio. In 1991, the group developed the ISO standard 11172, called *coding of moving pictures and associated audio.* This standard, known as MPEG-1, is used for digital storage media at up to about 1.5 Mbits/s.

In 1996, the MPEG group developed the ISO standard 13818 called *Generic coding of moving pictures and associated audio.* This standard is known as MPEG-2 and is an extension of MPEG-1 that allows flexibility in input format and bit rates. The MPEG-2 standard specifies only the syntax of the coded bit stream and the decoding process. This means that there is some flexibility in the design of an encoder. As long as the encoder generates a bit stream consistent with the MPEG-2 bit-stream syntax and the MPEG-2 decoding process, it is considered a *valid* encoder. Because there are many methods to generate the coded bit stream that are consistent with the syntax and the decoding process, some optimizations and improvements can be made without changing the standard. The Grand Alliance HDTV system uses some compression methods included in MPEG-2 to generate a bit stream that conforms to the MPEG-2 syntax. An MPEG-2 decoder can decode a video bit stream generated by the Grand Alliance video coder.

Grand Alliance Video Compression System. A block diagram of the Grand Alliance video encoder is shown in Fig. 9. It is assumed that the input video consists of a succession of progressively scanned frames. The system also accommodates interlaced fields. The basic principles discussed in this section apply to both frames and fields.

Color Matrix Conversion. Each pixel in a frame is represented by three numbers: red (R), green (G), and blue (B). For each pixel, the Grand Alliance system converts RGB components to a set of three numbers using a 3×3 matrix operation. Each of the three components Y, C_b, and C_r can be viewed as a monochrome image. The Y component is called luminance, and the C_b and C_r components are called chrominance. For typical video frames, most of the high-frequency information is contained in the luminance component. In addition, the human visual system appears to be less sensitive to high-frequency information in the chrominance components. To exploit this, the Grand Alliance system uses lowpass filters and subsamples by a factor of 2 along each of the horizontal and vertical directions for each of the two chrominance components. This reduces the data rate for each of the two chrominance components by a factor of 4. The number of pixels in each chrominance component is 25% of the luminance component.

Motion Compensation. Much of the variation in intensity from one frame to the next is caused by the motion of objects or the camera. For a typical video sequence, the previous frame can often be used to predict the current frame, with appropriate compensation for object and camera motion. To the extent that a prediction can be made, the information that must be transmitted to reconstruct the current frame is the portion of the current frame that cannot be predicted from the previous frame. Video coding encodes the difference between the current frame and the prediction of the current frame from the previously reconstructed frame accounting for motion. This is known as motion-compensated encoding. The difference encoded is called the motion-compensated residual. Substantial bit savings can be obtained by predicting the current frame and encoding only the error in the prediction residual.

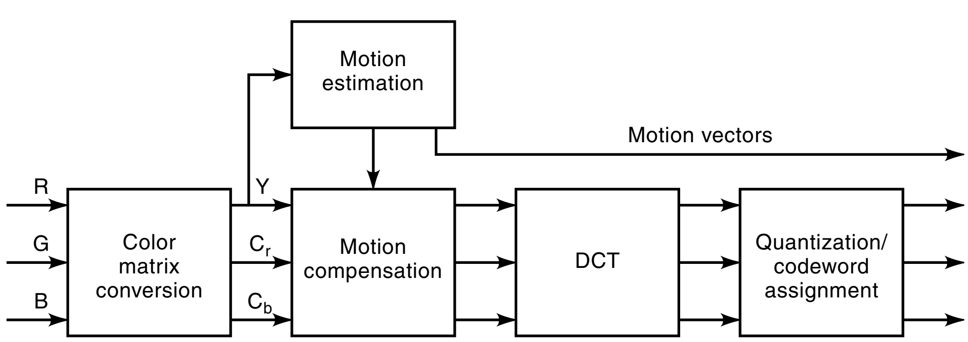


Figure 9. A block diagram representation of a video encoder.

Frames are classified into three groups. Frames in the first group are called intraframes or I-frames. These I-frames are encoded without motion compensation and without reference to other frames. When the previous frame cannot sufficiently predict the current frame, (e.g., scene changes), motion compensation is not useful and actually degrades compression performance. In addition, I-frames are useful for operations, such as random access, searches, and fast-forward operations because they do not depend on other frames. The I-frames also add robustness to bit errors. When bit errors occur and a portion of the current frame is in error, the errors can propagate in motion-compensated encoding until they are resynchronized. The I-frames provide such synchronization. They also provide an entry point when a television channel is changed.

The other two groups can utilize motion compensation. They are called predicted frames (P-frames) and bidirectionally predicted frames (B-frames). P-frames use only a past frame for prediction. The B-frames use a past frame, a future frame, or a combination of a past frame and a future frame. When a future frame is used to predict the current frame, the future frame must be reconstructed first. This is possible, for example, when the future frame is an I-frame and does not depend on the current frame. The future frame could then be coded, transmitted, or decoded first. Then it can be used to predict (code) the current frame. In terms of compression efficiency, B-frames are usually the best. The I-frames that do not use motion compensation are typically the least efficient in bit-rate reduction.

Discrete Cosine Transform. Motion compensation reduces the redundancy in video data along the temporal dimension. For the frame intensities or motion-compensated residuals, redundancy also exists in the spatial dimension. To exploit spatial redundancy, the intensities or motion-compensated residual are mapped to a frequency-domain representation called the discrete cosine transform (DCT). Among many possible frequency-domain representations, the DCT is chosen because it performs well when decorrelating the spatial correlation in typical video data and because fast algorithms exist to compute it.

The luminance Y component is divided into many blocks called macroblocks. One macroblock consists of 16×16 pixels. Each macroblock is further subdivided into four 8×8 blocks. For the chrominance C_b and C_r components, a block of 8×8 pixels corresponds to a macroblock of the luminance Y component because of the subsampling discussed earlier. The DCT is computed for each 8×8-pixel block for both the luminance and chrominance components.

For an I-frame, each block represents the intensities. For a P-frame or a B-fame, a block may represent the intensities of the motion-compensated residual or the frame intensities because some blocks within a P-frame or a B-frame may be encoded without motion compensation. For example, in spatial regions that correspond to newly exposed areas as a result of object movement, prediction by motion compensation is not effective and actually degrades the compression efficiency. Therefore, in these situations, it is often better to turn off the motion-compensated prediction for the block in question and code the pixel intensities instead.

There are several reasons for dividing a frame into blocks. The most important reason is its capability to adapt to local signal characteristics. For example, it may be better to encode some blocks without motion compensation and others with motion compensation. As another example, some local spatial regions of a frame may have nearly uniform intensities. For such regions, only the DC component (zero frequency component) of the DCT may need to be transmitted. For other local spatial regions where the signal varies more rapidly, additional frequency components may need to be transmitted. Another reason for block processing is the potential efficiency in computations, because it allows multiple processors to simultaneously process different blocks.

Quantization and Codeword Assignment. DCT coefficients are quantized by uniform quantization, where the quantization step size depends on luminance versus chrominance, motion-compensation versus nonmotion-compensation, the DCT frequency, and so on. The quantization step size is chosen to reflect the importance of a particular coefficient with respect to the human visual system. DCT coefficients of lower frequency typically have more impact than those of higher frequency. Therefore, a smaller quantization step size is usually used.

The DCT coefficients of the frame intensities or the motion-compensated residuals often have small amplitudes. For the blocks that use motion compensation, the motion-compensated residual is typically small and the corresponding DCT coefficients have low amplitudes. Even for blocks that do not use motion compensation, DCT coefficients (except for very low frequencies) are typically small. Video intensities often do not vary rapidly within a small local region, and the amplitudes of many high-frequency DCT coefficients are typically small. As a result, a large portion of the DCT coefficients are quantized to zero. This is one reason why the DCT is used. Instead of transmitting the value of zero many times, the locations of the DCT coefficients not quantized to zero are represented by run-length coding. Only those DCT coefficients with nonzero quantized amplitudes are transmitted. Substantial bit-rate reduction occurs because a large portion of the DCT coefficients have zero amplitude and thus do not require many bits.

The quantization of DCT coefficients and other parameters corresponds to assigning each parameter to an appropriate reconstruction level. When transmitting to the receiver which of the possible reconstruction levels has been selected, a specific codeword (a string of 0s and 1s) is assigned to each possible reconstruction level. Upon receiving the codeword, the receiver identifies the reconstruction levels by matching the codeword to an appropriate entry in the codebook. To ensure that the receiver identifies the reconstruction level uniquely, a different codeword is assigned to each reconstruction level. In addition, many reconstruction levels are transmitted sequentially and codewords must be assigned so that they are decoded correctly when the receiver receives them sequentially. The codewords with these characteristics are described as uniquely decodable.

The Grand Alliance system uses Huffman coding to assign codewords. This method assigns shorter codewords to more likely reconstruction levels and longer codewords to less likely reconstruction levels to reduce the average bit rate. The method is simple to use and generates uniquely decodable codewords that require the lowest possible average bit rate.

Video Decoder. The decoder conversely performs the operations of the encoder. The string of bits is decoded to generate the quantized parameters needed for video reconstruction. The quantized parameters are inverse-transformed and motion-compensated to reconstruct the luminance components. The chrominance components are interpolated to account for the subsampling operation performed at the encoder. Then the luminance and chrominance values are multiplied by a 3×3 matrix to generate the corresponding R, G, and B values, used by a display.

Using these video compression methods, the Grand Alliance system can compress an HDTV video within a bit rate of 19 Mbits/s with very good video quality, except for extremely difficult scenes.

Audio Processing

The Grand Alliance HDTV system compresses the audio signal to efficiently use the available spectrum. To reconstruct CD-quality audio after compression, the compression factor for audio is substantially less than that of the video. High-quality audio is important for HDTV viewing. The data rate for audio is inherently much lower than that for video, and the additional bit-rate efficiency achieved at the expense of audio quality is not worthwhile for HDTV applications.

Consider one channel of audio. The human auditory system is not sensitive to frequencies above 20 kHz. The audio signal sampled at a 48 kHz rate is sufficient to ensure that the audio information up to 20 kHz is preserved. Each audio sample is typically quantized at 16 bits/sample. The total bit rate for one channel of audio input is 0.768 Mbits/s. Exploiting the limitations of the human auditory system, the bit-rate requirement is reduced to 0.128 Mbits/s, and the reproduced audio quality is almost indistinguishable from that of the input audio. The compression factor achieved is six, which is substantially less than the video's compression factor of more than 50. In the case of video, it is necessary to obtain a very high compression factor, even at the expense of some noticeable quality degradation for difficult scenes because of the very high-input video bit rate (more than 1 Gbps). In the case of audio, additional bit-rate savings from 0.128 Mbits/s at the expense of possible audio quality degradation is not considered worthwhile for HDTV applications.

The Grand Alliance HDTV system uses a modular approach to the overall system design. Various technologies needed for a complete HDTV system can be chosen independently from each other. The audio compression method, for example, can be chosen independently of the video compression method. Even though the MPEG-2 standard used for video compression in the Grand Alliance system includes an audio compression method, the Grand Alliance selected the Audio Coder 3 (AC-3) standard on the basis of several factors, including performance, bit-rate requirement, and cost.

The compression system for one audio channel is shown in Fig. 10. The input is digital audio, which is obtained at a sampling rate of 48 kHz. The data is segmented by overlapping smooth windows. The length of one window is 512 points, and each window is overlapped by 50%. Overlapping data blocks is needed to avoid blocking artifacts. Because signal characteristics change over time, block processing allows adapting the processing method to local signal characteristics. For short transient regions, the window length can be reduced to 256 data points.

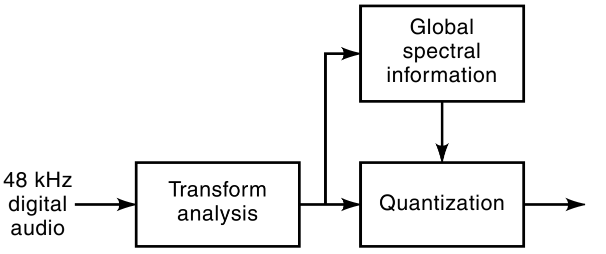


Figure 10. A block diagram representation of an audio encoder.

For each block, the temporal samples are transformed to the frequency domain to obtain spectral coefficients. The frequency-domain analysis used is the time-domain aliasing cancellation (TDAC) method. Because each block is overlapped by 50%, each sample is processed in two separate blocks. The TDAC method generates 256 spectral coefficients from a block of 512 audio samples. Despite the aliasing created by undersampling by a factor of 2, the TDAC method cancels aliasing and enables perfect reconstruction of the original signal by exploiting the fact that each audio sample is used in two different blocks. The spectral coefficients can also be computed by using a fast Fourier transform type (FFT) algorithm in the TDAC method.

Spectral envelope information can be obtained and encoded from the spectral coefficients that result from each block of data. The spectral envelope contains information that represents the global characteristics of the spectral coefficients. It is used to ascertain the masking level, which is useful in determining the level of quantization coarseness needed for different spectral coefficients. Spectral coefficients that do not require fine quantization are quantized coarsely, thus contributing to the bit-rate savings.

One feature of the human auditory system exploited in audio compression is the masking effect. A tone at a certain frequency can mask other tones within a neighboring frequency band called the critical band. If one spectral component is masked by other spectral components, the masked component does not have to be quantized finely. Spectral envelope information can be used to obtain a masking curve based on a psychoacoustic model of the human auditory system. As long as the noise introduced by quantization remains below the masking curve, it is not heard. Quantization of spectral components is carried out to ensure that the quantization noise cannot be perceived. At a bit rate of 128 kbits/s, one audio channel can be reproduced at the receiver with audio quality almost indistinguishable from that of the original input audio.

The Grand Alliance system can encode a maximum of six audio channels per audio program. The channelization follows the International Telecommunication Union (ITU-R) recommendation BS-775: "Multichannel stereophonic sound system with and without accompanying picture." The six audio channels are left, center, right, left surround, right surround, and low-frequency enhancement. The bandwidth of the low-frequency enhancement channel extends to 120 Hz, whereas the other five channels extend to 20 kHz. The six audio channels are also called 5.1 channels. Because the six channels are not completely independent for a given audio program, this de-

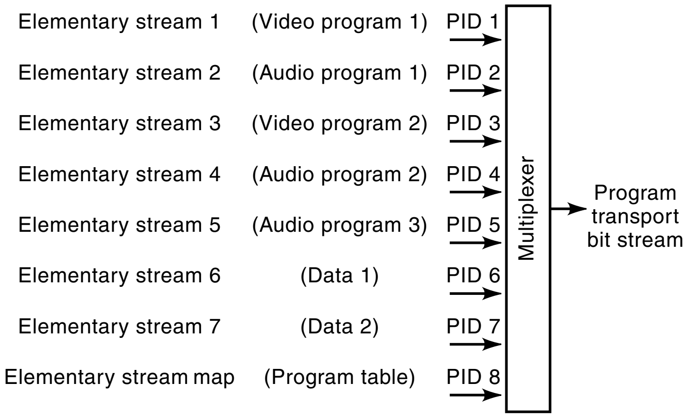


Figure 11. An example of the multiplex function used to form a program transport bit stream.

pendency can be exploited to reduce the bit rate. The Grand Alliance system encodes 5.1 channel audio at a bit rate of 384 kbits/s with audio quality essentially the same as that of the original.

Transport System. The bit streams generated by the video and audio encoders and the data channel must be multiplexed in an organized manner so that the receiver can demultiplex them efficiently. This is a main function of the transport system. The Grand Alliance system uses a transport format that conforms to the MPEG-2 system standard, but it imposes some constraints that allow more rapid channel change, which is important for a television system. This means that the Grand Alliance decoder cannot decode an arbitrary MPEG-2 system bit stream, but all MPEG-2 decoders can decode the Grand Alliance bit stream.

The bit stream that results from a particular application, such as video, audio, or data, is called an elementary bit stream. The elementary bit streams transmitted in a 6 MHz channel are multiplexed to form the program transport bit stream. Each elementary bit stream has a unique packet identification (PID) number, and all of the elementary bit streams within a program transport bit stream have a common time base.

An example of the multiplex function used to form a program transport stream is shown in Fig. 11. The first two elementary streams are from one television program. The next two elementary streams are from another television program. As long as the available bit rate for the channel can accommodate more than one television program, the transport system will accommodate them. The fifth elementary stream is only an audio program without the corresponding video. The next two elementary streams are from two different data streams. The last elementary stream contains the control information, which includes a program table listing all elementary bit streams, their PIDs, and the applications, such as video, audio, or data. All eight elementary streams that form the program transport bit stream have the same time base.

The Grand Alliance transport system uses the fixed-length packet structure shown in Fig. 12. Each packet consists of 188 bytes, divided into a header field and payload. The header field contains overhead information, and the payload contains the actual data that must be transmitted. The size of the packet is chosen to ensure that the actual payload-to-overhead ratio is sufficiently high and that a packet lost during transmission will not seriously affect the received video, audio, and data. The payload of each packet contains bits from only one particular elementary bit stream. For example, a packet cannot have bits from both a video elementary bit stream and an audio elementary bit stream.

Information that is not actual data, but is important or useful for decoding the received bit stream, is contained in the header. The header of each packet includes a four-byte link header and may also include a variable-length adaptation header when needed. The link header includes one-byte synchronization to indicate the beginning of each packet, a 13-bit PID to identify which elementary stream is contained in the packet, and information about whether or not the adaptation header is included in the packet. The adaptation header contains timing information to synchronize decoding and presentation of applications, such as video and audio. This information can be inserted into a selected set of packets. The adaptation header also contains information that facilitates random entry into application bit streams to support functions like program acquisition and change.

On the transmitter side, the bits from each elementary stream are divided into packets with the same PID. The packets are multiplexed to form the program transport bit stream. An example is shown in Fig. 13. At the receiver, the synchronization byte, which is the first byte in each packet, is used to identify the beginning of each packet. From the program table in the control packet, information can be obtained on which elementary streams are in the received program transport bit stream. This information, together with the PID in each packet, is used to separate the packets into different elementary bit streams. The information in the adaptation header in a selected set of packets is used to time and synchronize decoding and presentation of different applications (video, audio, data, etc.).

The transport system used in the Grand Alliance system has many advantages. The system is very flexible in dynamically allocating the available channel capacity to video, audio, and data, as shown in Fig. 13. The system can devote all available bits to video, audio, data, or any combination thereof. The system also can allocate available bits to more than one television program. If video resolution is not high,

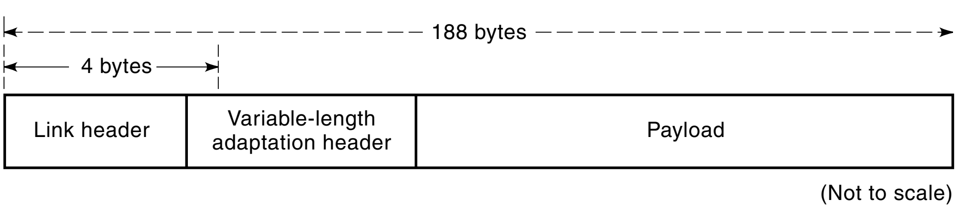


Figure 12. Fixed-length packet structure used in the Grand Alliance transport system.

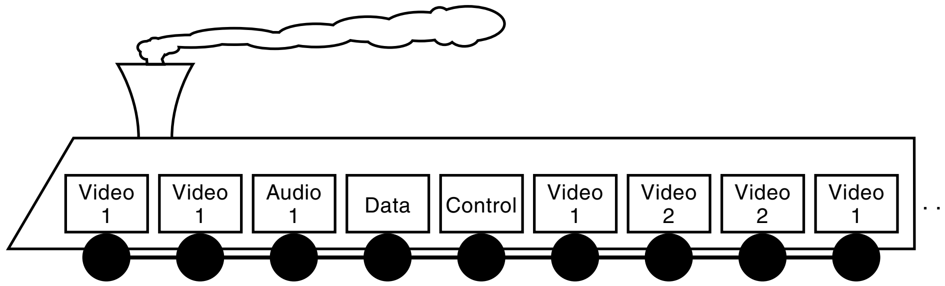


Figure 13. Example of packet multiplexing to form the program transport bit stream.

several standard-definition television programs (comparable to the NTSC resolution) can be transmitted. This is in sharp contrast to the NTSC system, where a fixed bandwidth is allocated to one video program and a fixed bandwidth is allocated to audio. The capability to dynamically allocate bits as the need arises is a major feature of the transport system.

The transport system is also scalable. If a higher bit-rate channel is available, the same transport system can be used by simply adding elementary bit streams. The system is also extensible. If future services become available, like 3-D television, they can be added as new data streams with new PIDs. Existing receivers that do not recognize the new PID will ignore the new data stream. New receivers will recognize the new PID.

The transport system is also robust in terms of transmission errors and is amenable to cost-effective implementation. The detection and correction of transmission errors can be synchronized easily because of the fixed-length packet structure. The fixed-length packet structure also facilitates simple demultiplex designs for low-cost, high-speed implementation.

Transmission System

The bit stream generated by the transport system must be processed to prepare for modulation and then modulated for transmission. Choosing from the many modulation methods depends on several factors including the transmission medium and the specific application. The best method for terrestrial broadcasting may not be the best for satellite broadcasting. Even for terrestrial broadcasting, the use of taboo channels means that interference with existing NTSC channels must be considered to determine the specific modulation technology.

Considering several factors, such as coverage area, available bit rate, and complexity, the Grand Alliance system uses an 8-VSB system for terrestrial broadcasting. A block diagram of the 8-VSB system is shown in Fig. 14.

Data Processing. The data processing part of the 8-VSB (vestigial sideband modulation with 8 discrete amplitude levels) system consists of a data randomizer, a Reed–Solomon encoder, a data interleaver, and a Trellis encoder. Data packets of 188 bytes/packet are received from the transport system and randomized. Portions of the bit stream from the transport system may have some pattern and may not be completely random. Randomization ensures that the spectrum of the transmitted signal is flat and is used efficiently. In addition, the receiver exhibits optimal performance when the spectrum is flat.

The transmission channel introduces various types of noise, such as random noise and multipath, which manifest themselves as random and bursty bit errors. To handle these bit errors, two forward error-correction (FEC) schemes are used. These FEC schemes add redundancy to the bit stream. Errors are then detected and corrected by exploiting this added redundancy. Even though error-correction bits use some of the available bit rate, they increase overall performance of the system by detecting and correcting errors.

The first-error correction method is the Reed–Solomon code, known for its burst noise correction capability and efficiency in overhead bits. The burst noise correction capability is particularly useful when the Reed–Solomon code is used with the Trellis code. The Trellis code is effective in combating random and short impulsive noise, but it tends to generate burst errors in the presence of strong interference and burst noise. The Reed–Solomon code used in the 8-VSB system adds approximately 10% of overhead to each packet of data. The result is a data segment consisting of a packet from the transport system and the Reed–Solomon code bits. The resulting bytes are convolutionally interleaved over many data segments. The convolutional interleaving useful in combating the effects of burst noise and interference is part of the Trellis encoding. The Trellis encoder, which is a powerful technique for correcting random and short-burst bit errors,

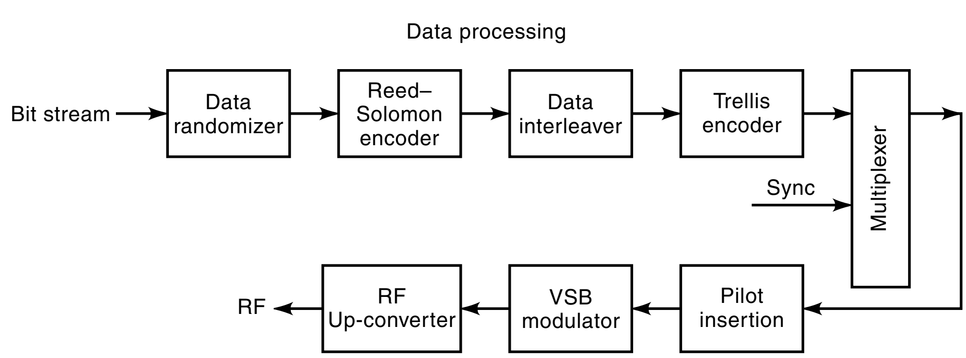


Figure 14. A block diagram of an 8-VSB system used in the Grand Alliance HDTV system.

adds additional redundancy. In the 8-VSB system, the Trellis coder creates a three-bit symbol (eight levels) from a two-bit data symbol.

At the transmitter, the Reed–Solomon encoder precedes the Trellis encoder. At the receiver, the Trellis decoder precedes the Reed–Solomon decoder. The Trellis decoder is effective in combating the random and short-burst bit errors, but it can create long-burst bit errors in the presence of strong interference and bursts. The Reed–Solomon code is effective in combating long-burst bit errors.

The 8-VSB system transmits approximately 10.76 million symbols per second. Each symbol represents three bits (eight levels). When accounting for the overhead associated with the Trellis coder, the Reed–Solomon coder, and additional synchronization bytes, the bit rate available to the transport system's decoder, is approximately 19.4 Mbits/s. The bit rate is for applications, such as video, audio, and data, and also for overheard information (link header, etc.) in the 188-byte packets. The actual bit rate available for the applications is less than 19.4 Mbits/s. The VSB system can deliver a higher bit rate, for example, by reducing the redundancy in the error correction codes and by increasing the number of levels per symbol. However, the result is loss of performance in other aspects, such as the coverage area. The 8-VSB system delivers the 19.4 Mbits/s bit rate to ensure that an HDTV program can be delivered within a 6 MHz channel with a coverage area at least as large as an NTSC system and with an average power level below the NTSC power level to reduce interference with the NTSC.

The Reed–Solomon and Trellis encoding results in data segments. At the beginning of each set of 312 data segments, a data segment is inserted that contains the synchronization information for the set of 312 data segments. This data segment also contains a training sequence for channel equalization at the receiver. Linear distortion in the channel can be accounted for by an equalizer at the receiver. A four-symbol synchronization signal is inserted at the beginning of each data segment. The data segment sync and the 312-segment set sync are not affected by the Trellis encoder and can provide synchronization independent of the data.

Pilot Insertion. Before modulation, a small pilot carrier is inserted in the lower band within the 6 MHz band. The location of the pilot is on the Nyquist slope of NTSC receivers. This ensures that the pilot does not seriously impair existing NTSC service. The channel assigned for the HDTV service may be a taboo channel currently unused because of the cochannel interference with an existing NTSC service located some distance away. The HDTV signal must be designed to ensure that its effect on existing service is minimal.

The NTSC system is rugged and reliable. The main reason for this is the use of additional signals for synchronization that do not depend on the video signals. The NTSC receiver reliably synchronizes at noise levels well below the loss of pictures. In the 8-VSB system, a similar approach is taken. The pilot signal that does not depend on the data is used for carrier acquisition. In addition, the data segment sync is used to synchronize the data clock for both frequency and phase. The 312-segment-set sync is used to synchronize the 312-segment set and equalizer training. Reliance on additional signals for carrier acquisition and clock recovery is very useful. Even when occasional noise in the field causes a temporary loss of data, quick recovery is possible as long as the carrier acquisition and clock recovery remain locked during the data loss. The 8-VSB system ensures that carrier acquisition and clock recovery remain intact well below the threshold level of data loss by relying on additional signals for such functions.

Modulation. To transmit the prepared bit stream (message) over the air, the bit stream must be mapped to a band-pass signal that occupies a 6 MHz channel allocated to a station's HDTV service. A modulator modulates a carrier wave according to the prepared bit stream. The result is a band-pass signal that occupies a given 6 MHz channel.

Any band-pass signal can be represented by the expression

$$x(t) = A \cdot [I(t) \cdot \cos(\omega_c t + \Theta) - Q(t) \cdot \sin(\omega_c t + \Theta)] \qquad (1)$$

The amplitude A, the carrier angular frequency ω_c, and the phase are constant. The modulating signal that carries the message is the time-varying I (in-phase) and Q (quadrature) components. In the 8-VSB system, the modulating signal changes once for each symbol at approximately 10.76 million times per second. Each symbol has three bits or eight levels.

The values of I and Q corresponding to a symbol can be represented on an I/Q diagram. The constellation of I and Q for a received VSB signal on an I/Q diagram is shown in Fig. 15. Each symbol received has a particular value of I and a particular value of Q. The constellation that shows the possible I and Q values received in a noise-free environment has eight discrete values of I. Each corresponds to one of the eight values of a symbol. The Q channel is continuous, and the amplitude of a Q value is a sample from an approximately Gaussian probability density function with a mean of zero. Therefore, the constellation shows eight horizontal lines on the I/Q diagram. In the case of quadrature amplitude modulation (QAM), yet another well-known method of digital modulation, the constellation has discrete points on the I/Q diagram.

The 8-VSB system modulates the signal onto an intermediate frequency (IF) carrier, which is the same frequency for all channels. It is followed by an up-conversion to the desired HDTV channel.

Receiver. At the receiver, the signal is processed to obtain the data segments. One feature of the receiver is the NTSC rejection filter, which is useful because of the way HDTV service is being introduced in the United States. In some locations, an HDTV channel occupies the same frequency band as

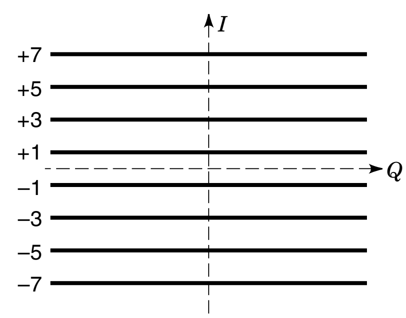


Figure 15. Constellation of I and Q for a received VSB signal.

an existing NTSC channel located some distance away. The interference of the HDTV signal with the NTSC channel is minimized by a very low-power level of the HDTV signal. The low-power level was made possible for HDTV service because it efficiently uses the spectrum. The 8-VSB receiver contains an NTSC rejection filter to reduce interference between the NTSC signal and the HDTV channel. This is a simple comb filter whose rejection null frequencies are close to the video carrier, chroma carrier, and audio carrier frequencies of the NTSC signal. The comb filter, which reduces other aspects of system performance, can be activated only when a strong cochannel NTSC signal interferes with the HDTV channel.

Cliff Effect. Although additive white Gaussian noise does not represent typical channel noise, it is often used to characterize the robustness of a digital communication system. The segment error probability for the 8-VSB system in the presence of additive white Gaussian noise is shown in Fig. 16. At the signal-to-noise ratio (SNR) of 14.9 dB, the segment error probability is 1.93×10^{-4} or 2.5 segment errors/s. The segment errors become visible at this threshold. Thus, up to an SNR of 14.9 dB, the system is perfect. At an SNR of 14 dB, which is just 0.9 dB less than the threshold of visibility (TOV), practically all segments are in error. This means that a system operating perfectly above the threshold of visibility becomes unusable when the signal level decreases by 1 dB or when the noise level increases by 1 dB. This is known as the *cliff effect* in a digital communication system. In an NTSC system, the picture quality decreases gradually as the SNR decreases. To avoid operating near the cliff region, a digital system is designed to operate well above the threshold region within the intended coverage area. Both laboratory and field tests have demonstrated that the coverage area of the HDTV channel is equal to or greater than the NTSC channel, despite the substantially lower power level used for the HDTV signal.

Cable Mode. For a variety of reasons, a cable environment introduces substantially less channel impairments than terrestrial broadcasting. In the case of cable, for example, cochannel interference from an existing NTSC station is not an issue. This can be exploited to deliver a higher bit rate for applications in a 6 MHz cable channel. Although the Grand Alliance HDTV system was designed for terrestrial broadcasting, the system includes a cable mode that doubles the available bit rate with a small modification. Doubling the bit rate means the ability to transmit two HDTV programs within a single 6 MHz cable channel.

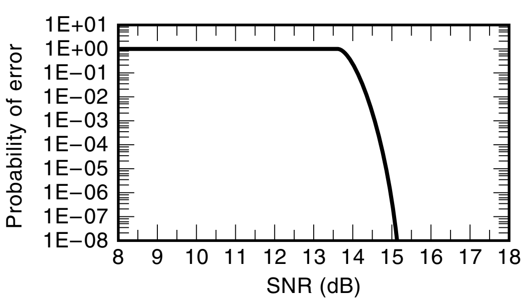


Figure 16. Segment error probability for the 8-VSB system in the presence of additive white Gaussian noise.

To double the bit rate for the cable mode, the Grand Alliance System uses 16-VSB rather than 8-VSB. All other aspects of the system, such as video compression, audio compression, and transport, remain the same. In the 8-VSB system, a symbol is represented by eight levels or three bits. One of the three bits is due to the redundancy created by Trellis encoding. For a cable environment, error correction from the powerful Trellis coding is no longer needed. In addition, the higher available SNR for cable means that a symbol can be represented by sixteen levels or four bits. Because the symbol rate remains the same between the 8-VSB and 16-VSB systems, the available bit rate for the 16-VSB system doubles compared with the 8-VSB system. For the 16-VSB system without Trellis coding, the SNR ratio for the segment error rate that corresponds to the threshold of visibility in the environment of additive white Gaussian noise is approximately 28 dB. This is 13 dB higher than the 8-VSB system with the Trellis coding. This increase in the SNR is acceptable in a cable environment that has substantially less channel impairments than a typical terrestrial environment.

HDTV AND INTEROPERABILITY

The Grand Alliance HDTV system has served as the basis for the digital television standard in the United States. The standard itself defines a significant number of technical elements. The technologies involved, however, will continue to develop without the need to modify the standard. For example, the video compression system adopted defines syntax only for the decoder. There is much room for improvement and advances in the encoder. Technical elements can also be added in a backward-compatible manner. The transmission of a very high-definition television format, not provided in the initial standard, can be accomplished in a backward-compatible manner by standardizing a method to transmit enhancement data. In turn, this can be combined with an allowed video transmission format to deliver the very high-definition format.

The Grand Alliance system was designed for HDTV delivery in terrestrial environments in the United States. For other delivery environments, such as satellite, cable, and environments in other parts of the world, other standards will emerge. Depending on the degree to which elements are common, interoperability among different standards will become an important issue. In terms of program exchange for different delivery media and throughout the world, technologies that convert one standard to another will continue to be developed. Efforts to facilitate the conversion process by adopting common elements among the different standards also will continue.

Interoperability will be an issue among the different HDTV standards, and also among telecommunication services and computers. A traditional television system has been a stand-alone device whose primary purpose was entertainment. Although an HDTV system is used for entertainment, it can be an integral part of a home center for entertainment, telecommunications, and information. The HDTV display can be used as a videophone, a newspaper service, or a computer display. Interoperability between an HDTV system and other services that will be integrated in the future is an important consideration.

BIBLIOGRAPHY

ATSC, *Digital Audio Compression AC-3,* December 20, 1995.

ATSC, *Digital Television Standard,* September 16, 1995.

ATSC, *Guide to the Use of the ATSC Digital Television Standard,* October 4, 1995.

V. Bhaskaran and K. Konstaninides, *Image and Video Compression Standards and Architectures,* Norwell, MA: Kluwer, 1995.

W. Bretl, G. Sgrignoli, and P. Snopko, *VSB Modem Subsystem Design for Grand Alliance Digital Television Receivers, IEEE Trans. Consumer Electron.,* **41**: 773–786, 1995.

A. B. Carlson, *Communication Systems,* 3rd ed., New York: McGraw-Hill, 1986.

The Grand Alliance Members, The US HDTV Standard, The Grand Alliance, *IEEE Spectrum,* **32** (4): 36–45, 1995.

ISO/IEC JTC1 CD 11172. *Coding of moving pictures and associated audio for digital storage media up to 1.5 Mbits/s,* Geneva: International Organization for Standardization (ISO), 1992.

ISO/IEC JTC1 CD 13818. *Generic coding of moving pictures and associated audio,* Geneva: International Organization for Standardization (ISO), 1994.

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

J. S. Lim, *Two-Dimensional Signal and Image Processing,* Englewood Cliffs, NJ: Prentice-Hall, 1990.

A. N. Netravali and B. G. Haskell, *Digital Pictures: Representation and Compression,* New York: Plenum, 1988.

JAE S. LIM
Massachusetts Institute of Technology

HIGH ELECTRON MOBILITY TRANSISTORS. See MODULATION DOPED FETS.

HIGH ENERGY LASERS. See CHEMICAL LASERS.

HIGH-ENERGY-PARTICLE DETECTORS. See PARTICLE SPECTROMETERS.

HIGH ENERGY PHYSICS. See PARTICLE SPECTROMETERS.

HIGH-ENERGY PHYSICS PARTICLE DETECTOR MAGNETS