ADVANCED DIGITAL TELEVISION

SYSTEM DESCRIPTION

SUBMITTED TO THE FCC/ACATS BY

THE ADVANCED TELEVISION RESEARCH CONSORTIUM:

DAVID SARNOFF RESEARCH CENTER

NBC

NORTH AMERICAN PHILIPS

THOMSON CONSUMER ELECTRONICS

FEBRUARY 27, 1991

EXECUTIVE SUMMARY

Advanced Digital Television (ADTV) is a fully digital system that delivers high-definition television (HDTV) in a 6-MHz channel. The design of ADTV has been driven by the need of the terrestrial Broadcast, Cable Television, and Consumer Electronics industries to provide the American public with a quality simulcast HDTV service that is robust and reliable. To achieve its goals, ADTV has made significant improvements to proven digital compression and transmission techniques, and molded them into a single cohesive system. There are three key elements in the ADTV system.

- First, ADTV's video compression, called MPEG++, is based on a specific implementation of the MPEG[1] (Moving Pictures Expert Group) compression approach. MPEG++ upgrades the standard MPEG approach to HDTV performance level and incorporates a video data prioritization layer that allows the most important video data to be transmitted with the greatest reliability.

- Second, ADTV incorporates a Prioritized Data Transport (PDT) layer. PDT is a cell relay-based data transport layer that supports the prioritized delivery of video data, thus providing the feature of graceful service degradation under impaired channel conditions. PDT also offers service flexibility for a wide mixture of video, audio, and auxiliary data services, and compatibility to broadband ISDN (integrated services digital network).

- Third, ADTV applies spectral-shaping techniques to Quadrature Amplitude Modulation (QAM) to carefully minimize interference *from* and *to* any co-channel NTSC signals. The result is an extremely robust data transmission system, known as the Spectrally-Shaped QAM (SS-QAM).

Although submitted as a 1050-line system, ADTV is designed to provide flexible support of a wide range of services and future media formats. The initial hardware implementation will use the interlaced scan format for video source and display (1050/59.94/2:1), with a 16:9 aspect ratio and more than twice the NTSC resolution. Selection of this initial format was based on current camera and display technologies, and does not preclude a future adoption of other video formats, consistent with the evolution of studio equipment and production standards.

In summary, ADTV is an integrated system that uniquely incorporates elements of digital video data compression (MPEG++), data transport (PDT), and digital transmission (SS-QAM) into a high-quality HDTV system that meets the needs of the broadcasters, cable operators, consumer electronics manufacturers, and the consumers. By adapting proven, widely-accepted techniques, ADTV provides an open door for industry acceptance and future ease of integration with other video and multimedia equipment and services. Together, these important attributes of ADTV form the basis for a successful development of an HDTV industry in the United States.

---

1 MPEG is a committee within the International Standards Organization (ISO) that is currently working toward a standard for digital video storage applications.

TABLE OF CONTENTS

LIST OF FIGURES

1. ADTV System Overview

1.1. System Rationale

It is the consensus within the ATRC that a well designed digital HDTV system for terrestrial broadcast offers significant performance advantages over analog approaches. Digital approaches allow a video compression system to efficiently and adaptively exploit spatial and temporal redundancies in a picture far better than analog approaches can. Digital techniques have the well-known property of being impervious to moderate levels of channel impairment, and avoid accumulation of noise and other artifacts when passed through a series of transmission relay stages. Digital technology is consistent with the accelerating technological trend towards integrated digital processing, storage and transmission of voice, video, and data. Service flexibility afforded by digital approaches far exceeds that of any analog approaches. Over the lifetime of a simulcast HDTV standard, the universality of digital technology will surely lead to novel combinations of applications now considered separate, along with greater compatibility among different types of consumer electronics, telecommunications, and computing equipment.

1.1.1. MPEG++ Compression

Digital simulcast research efforts have been on-going for several years within ATRC laboratories. A rigorous and comprehensive study and evaluation of numerous state-of-the-art video compression systems was an integral part of this effort. Among the different video compression approaches, the MPEG (Moving Pictures Experts Group, a committee within the International Standards Organization) compression approach represents a collection of well-known, well-proven compression methods that have been tightly integrated into a single algorithm. The picture quality of a particular algorithmic implementation of the MPEG approach evaluated by the ATRC was impressive. The ATRC implementation of the MPEG approach outperforms most custom approaches because it incorporates results filtered from a great deal of image coding research conducted around the world during the past years.

MPEG compression, like other video compression, is vulnerable to hostile transmission conditions. To combat the disruptive nature of transmission errors on the video data stream, a channel-specific video prioritization layer was designed into the ADTV compression algorithm. We use the name MPEG++ to indicate that MPEG has been adapted and improved to meet the requirements of the terrestrial broadcast environment. Coupled with prioritized data delivery, MPEG++ represents an extremely robust video compression algorithm.

1.1.2. Prioritized Data Transport (PDT)

The unique requirements of simulcast transmission demand special attention to the delivery and transport of the compressed video information. This motivated an ATRC research effort to design a transport format to provide extremely reliable delivery of the ADTV service to the viewers. The result of this "channel-ruggedization" effort is the "fast packet" cell transport format.

Cell-relay transport format refers to a data communication format in which information bits are carried in *cells* consisting of fixed-size data, header, and trailer. Because of the relative ease of switching and routing a fixed-size cell, even at high-speed, most high-speed data communication networks have adopted a fixed-size cell relay format. A well known example of a cell-relay transport format is the ATM (asynchronous transfer mode) protocol in broadband ISDN (integrated services digital network).

Cell relay provides rugged logical synchronization that is essential for reliable delivery of variable length coded compressed video in the presence of transmission errors. With its link-level asynchronous time division multiplexing features, ADTV offers the advantage of flexible multiplexing of video, audio, and data with bit-rates that do not need to be specified in advanced.

1.1.3. Spectrally-Shaped QAM (SS-QAM)

The modulation used in the ADTV system was developed specifically for the terrestrial simulcast application with the joint design goals of achieving high spectral efficiency and NTSC-robust/NTSC-friendly operation. The modem emits a signal whose spectrum has been specifically shaped to avoid *mutual* interference with co-channel NTSC signals, providing significantly better simulcast properties than a standard wideband QAM.

1.2. SYSTEM ARCHITECTURE

Figure 1.1 illustrates the system architectural view of ADTV. Four principal layers are shown in the figure: the compression algorithm layer, the data prioritization layer, the transport layer, and the physical communication layer. The compression algorithm layer performs the tasks of data pre/post-processing and compression/decompression. The prioritization layer has the tasks of data priority assignment at the encoder and combining the data elements of different priority into coherent data streams for decompression at the decoder. The transport layer is responsible for data multiplexing, cell formatting, and cell error detection. The physical communication layer performs the tasks of channel error correction coding/decoding, modem carrier modulation, channel equalization, and frequency translation to/from i.f./r.f. Table 1.1 is a summary of the key technical features of the ADTV system, and Table 1.2 list the key system parameters.

Services	
Video	Sophisticated MPEG++ video compression producing high-quality HDTV pictures at 20 Mbps. Video data prioritization prepares the compressed video data for transmission over terrestrial broadcast environment. MPEG++ provides the following benefits: • Hardware complexity at the encoder, not the consumer receiver. • Future quality improvement can be achieved by modifications at the encoder only; not the receiver. • Flexible support of video formats.
Audio	ADTV provides for up to four digital audio channels of CD-quality sound commensurate with the picture quality of an HDTV system. The four channels nominally make up two stereo pairs. The audio compression will follow closely the industry standard for broadcast digital audio, e.g., MUSICAM[2].
Data	ADTV supports an auxiliary data channel for program-related information (e.g., encryption) or any other community service information.
Prioritized Data Transport (PDT)	ADTV's layered data transport format supports important service flexibilities and reliable delivery of variable length coded compressed video. The transport layer provides flexible support of multiple video, audio, and data services, without specific constraints on the bit-rate. ADTV's layered data format simplifies the process of transcoding for other delivery media (e.g., satellite, fiber-optic or coax cable, digital recording) by providing several logical "entry points" into the data stream.
Spectrally-Shaped QAM (SS-QAM)	By a process of spectral shaping, mutual-interference between the ADTV signal and any co-channel NTSC signal is minimized. This provides a high degree of robustness to the ADTV signal in the presence of NTSC interference. At the same time, the ADTV signal is designed to be substantially rejected by an NTSC receiver.

Table 1.1: ADTV Technical Features.

2 Masking-Pattern-Adapted Universal Subband Integrated Coding and Multiplexing. MUSICAM has been adopted by the ISO MPEG committee.

Video Characteristics	
Raster Format	1050/2:1 Interlace
Aspect Ratio	16:9
Frame Rate	29.97 frames/sec
Active Video	
Luminance	1440 (H) x 960 (V)
Chrominance	720(H) x 480 (V)
Horizontal Resolution (Static and Dynamic)	810 TVL per Picture Height
Transport Cells	
Cell Size	256 Bytes
Link-Level Overhead	3 Bytes (1.1%)
Payload Size	253 Bytes
Total Data Rate	21.00 Mbps
Video	14.98 Mbps
Audio	1.02 Mbps
Data (max.)	0.04 Mbps
Error Correction and Link-Level Cell Overhead (percentage of total rate)	23.6 %

Table 1.2: ADTV System Parameters.

Figure 1.2 is a system-level block diagram of ADTV illustrating the implementation of the layered architecture in the form of subsystems corresponding to the four layers shown in Figure 1.1. These are labelled the Compression Encoder/Decoder (CE/CD), the Prioritization Encoder/Decoder (PE/PD), the Transport Encoder/Decoder (TE/TD), and the Modem-FEC Encoder/Decoder (MFE/MFD). Figure 1.3 shows some typical blocks within each encoder/decoder pair. Modules/blocks within the same subsystem are said to be *peer* of one another. In the sections that follow, broad functional descriptions of each subsystem and their associated modules/blocks are presented. Detailed descriptions of modules within each subsystem will be given in subsequent sections in this document.

1.3. COMPRESSION ENCODER/DECODER (CE/CD)

All input and output format transcoding operations are performed by the Pre/Post Processor.

Within the Video Processor lies the MPEG++ compression encoder/decoder. The video compression algorithm is the MPEG++ algorithm. The video compression operations include motion estimation, motion-compensated predictive coding, adaptive DCT quantization, and variable-length coding/decoding (VLC/VLD). Picture frames are classified into one of three types: the intra-frames (I frames), the predicted frames (P frames), and the bidirectional frames (B frames). The I frames are processed independent of all other frames. The P frames are coded by a motion-compensated predictive coder using the previous I or P frames. The B frames are coded by a bidirectional motion-compensated predictive coder using the two adjacent I or P frames. In this document, the I and the P frames are also referred to as *anchor* frames. The coded picture information (pixel values for the I frames and residual error after prediction for the P and the B frames) is transformed by Discrete Cosine Transform (DCT) and then adaptively quantized. Variable length coding (VLC) is applied to the quantized DCT coefficients.

At the receiver, in addition to the appropriate inverse operations of video decoding, the video decoder has the task of error recovery. Error recovery may involve data manipulation at each layer of the decoding process.

1.4. PRIORITIZATION ENCODER/DECODER (PE/PD)

1.4.1. Priority Processor

Surrounding the central compression codec (coder/decoder) is the data prioritization layer. The data prioritization process identifies pieces of information that are more critical (in terms of their impact on the system performance) and therefore require more careful treatment by the transport and communication process. Different *priority* is assigned to different pieces of information or *data elements*.

The main function of the Priority Processor is to identify the channel error protection requirement of each piece of information or data element. A data element is typically a codeword of a specified bit length. The Priority Processor also knows of the "type" of data element. For instance, the types of video data elements include motion information, block size, DCT coefficient, quantization parameters, etc. The Priority Processor then assigns different *priority* to each data element according to an assignment rule. The assignment is made dynamically, reflecting any fluctuation in the channel load. In general, the load on the channel will vary in time because of the variable output rate of the compressed video data.

1.4.2. Rate Controller

The Rate Controller monitors the channel load and instructs the appropriate module in the video encoder to increase or decrease the flow of data into the channel. The Rate Controller has a separate communication into the video encoder to regulate the amount of video compression according to the output data rate. The Rate Controller also communicates with the rate buffers in the Transport Encoder (see next section) to solicit buffer occupancy information. The main purpose of the Rate Controller is to maintain and control the flow of data (variable-rate) into a fixed rate channel. The rate-regulation is referred to as *buffer feedback control* in classical variable-rate video compression systems.

1.5. TRANSPORT ENCODER/DECODER (TE/TD)

A sophisticated transport format has been developed by the ATRC specifically to handle information with different error protection requirements over a simulcast channel. Data encapsulation to support a prioritized transport is designed specifically to provide payload chaining and segmentation capability in order to maximize channel utility and minimize the impact of payload losses on the system.

The Transport Processor asynchronously multiplexes the *payload* data with different priorities into basic transport units called *cells*. Cells are fixed sized and are prioritized according to their prescribed protection levels. Each cell has its own error checking bits (cyclic redundancy codes, CRC), which provide a powerful link-level error detection capability to the system. In addition, using the method of *segmentation, chaining* of the cell payload is maintained while avoiding cell-to-cell error propagation where possible.

The Transport Encoder/Decoder also performs rate buffering for the purpose of variable to fixed bit rate conversion (e.g., variable rate to fixed channel rate buffering, synchronization of video data from the different protection queues, etc.) The buffers at the transmitter and at the receiver work together to ensure that the overall data bits experience a *fixed* delay despite the variable delay caused by the variable length encoding.

1.6. MODEM-FEC ENCODER/DECODER(MFE/MFD)

The outer most ADTV layer is the physical communication layer comprising Reed-Solomon error-correcting coding and the modem. The Spectrally-Shaped QAM (SS-QAM) in the ADTV modem, together with powerful forward error correction channel coding, provides a prioritized transmission service that is extremely reliable even under degraded reception conditions such as strong NTSC co-channel interference.

Reed-Solomon forward error correction (FEC) codes are applied to the data bytes before the carrier-modulation stage. Depending on the priority, different FEC is applied to the data. In addition, data interleaving is performed as part of the operations of channel coding by the FEC module. This ensures that bursts of channel bit errors can be treated as uncorrelated random bit errors which can often be corrected by the Reed-Solomon codes. At the receive side, inspection and error correction operations take place at the peer modules within the Modem-FEC Decoder.

Before information is sent from the transmitter, it must go through the process of modulation. Similarly, a process of demodulation is necessary to translate the information from the received analog waveform into digital bits and bytes. These operations are the classical functions of a modem (modulator/demodulator.) The basic modulation method used is QAM (Quadrature Amplitude Modulation) with specific spectral shaping incorporated into the modulation process so that the transmitted signal affords the ADTV system a duality of protection: low sensitivity to co-channel NTSC interference, and strong rejection by an NTSC receiver.

2. Compression Encoder/Decoder

2.1. Pre/Post Processing

The input/output video format for ADTV is 1050/59.94/2:1. Recognizing the 2:1 relationship with 525-line NTSC, the ATRC believes that the 1050 system is likely to be the production standard of choice in North America. However, being digital with flexible format specification headers, the ADTV system approach is equally applicable to other video formats. For the purpose of the FCC tests, the ADTV prototype hardware is designed for the 1050/59.94/2:1 format. Specifically, the ADTV prototype hardware will have the following video format. (See also Table 1.2.) The input to the encoder video input processor is component Y-U-V. The luminance (Y) signal is sampled to produced picture frames of 1440 active pixels by 960 active lines. The color information (U, V) is sampled at half the Y sampling frequency, producing color difference frames of 720 active pixels by 960 active lines. Further processing is carried out to vertically decimate the U and V components to 480 active lines per frame. The necessary vertical interpolation will be carried out at the receiver video post processor. The ADTV prototype hardware will operate on interlaced frames.

2.1.1. Color Conversion

Color information is carried by the two color difference signals of U and V. The chromaticity of Y, U, and V is in accordance to the SMPTE 240M specification. The U and V sample density is half that of Y in both the horizontal and the vertical directions. The temporal density of U and V is identical to that of Y. In addition, each U and V sample location corresponds to the geometric center of a rectangle formed by the four adjacent Y samples. (See Figures 2.1 and 2.2.)

2.2. Video Compression Processor

The bulk of the video compression is accomplished within the Video Processor block. As mentioned before, the basic compression approach is that of the MPEG compression algorithm. Appendix A provides some useful background information on video compression related to MPEG. Because many modules in a video compression codec are functionally well defined and well understood, the emphasis in this section is on the integration of these functional modules. The MPEG ISO draft provides a detailed description of the compressed data format and of the decoder.

Before going into the detail of the video processor, there are some basic terms that must be defined.

Video Sequence	A video sequence is a collection of picture representations and the associated chronological relationship among the pictures.
Picture	A picture is a two dimensional representation of the scene to be displayed. Hereafter in this document, unless specified otherwise, a picture refers to the set of two-dimensional pixel arrays of Y, U, and V representing a particular video raster frame.
Group of Pictures	A Group of Pictures (GOP) is a series of pictures with a particular structure. A GOP comprises up to three types of pictures: the I-frames, the P-frames, and the B-frames. Each type of picture is subjected to (at the encoder side) and reconstructed from (at the decoder side) different coding and decoding processes. An example of a GOP is shown in Figure 2.3.
I-Frames (Intra-frames)	I-frames are processed by intra-frame operations. They serve as anchors to the temporal prediction and/or interpolation processes.
P-Frames (Predicted-frames)	The compression of P-frames rely on temporal prediction from previous anchor pictures (either I- or P-frames.) The motion compensation is incorporated in the temporal prediction, and because the motion estimation is always from the past into the future, the motion compensation is called *forward* motion compensation.
B-Frames (Bidirectionally-predicted frames)	B-frames are always temporally predicted from two adjacent anchor pictures. The anchor pictures can be either I- or P-frames. The B-frame temporal prediction uses motion compensation in both *forward* and *backward* directions, and hence the name Bidirectional prediction. B-frames are not used in prediction. Figure 2.3 shows the temporal relationship of the three types of pictures.
Basic Block	For the purpose of performing the Discrete Cosine Transform (DCT), a picture is divided into basic blocks. Each basic block is an 8x8 array of values (pixels or DCT coefficients.) In a picture, the pixels for a particular block are spatially adjacent to each other just as they are in the block, i.e., there is no rearrangement of pixels from the picture to the basic block structure. Figure 2.1 shows a basic block and its relationship to a picture.
Macro Block	A macro block comprises four basic Y blocks and one basic U block and one basic V block. The Y, U, and V blocks corresponds the same spatial area in a picture. Figure 2.2 shows a macro block and its relationship to a picture.
Slice	A slice is a set of integer number of adjacent macro blocks from a picture.

Table 2.1: Definition of Terms.

2.2.1. *Group of Pictures, GOP*

As defined in Table 2.1, a Group of Pictures comprises up to three types of frames, the I, P, and B frames. An example of a Group of Pictures is shown in Figure 2.3. The temporal coding relationship of the three types of frames is also shown in Figure 2.3. Briefly, the I frames are processed using only intra-frame DCT adaptive quantization; the P frames are processed using a hybrid temporal predictive DCT coder with adaptive quantization and *forward motion compensation*; and the B frames are processed using a hybrid temporal predictive DCT coder with adaptive quantization and *bidirectional motion compensation*. The I and the P frames are also referred to as the *anchor* frames because of their roles in the bidirectional motion compensation of the B frames.

The Group-of-Pictures structure offers a good tradeoff between the high efficiency of temporal predictive coding, good error-containment features of periodic intra-only processing, and fast picture acquisition times (half the Group-of-Pictures time span, on the average,) whenever viewers hop from one channel to another.

Figures 2.5–2.7 show block diagrams of the video compression processor. Figure 2.6 is a more detailed view of the shaded blocks in Figure 2.5. The description of the I, P, and B processing will refer to these figures. There are, however, a few modules shared by all three processes. Specifically, the Input Sequencer and the Raster Line to Block/Macroblock converter modules in Figure 2.5 are common to the whole video encoder processor.

2.2.2. *Input Sequencer*

The Group-of-Pictures data structure requires some unique sequencing of the input video frames. Because of the backward motion compensation in B frame processing, the *anchor frames* must be processed before the B frames associated with the two anchors. Suppose X_n denotes a frame of type X (i.e., X can be I, P, or B) and the subscript *n* refers to its time index, i.e., X_1 is chronologically before X_2 which in turn comes before X_3. For a sequence of frames that is to be displayed in the order as shown in Figure 2.4(a), the frames are processed in the order as shown in Figure 2.4(b). The re-ordering of the input sequence from the normal chronological frame order to frame order required by the video compression processor is carried out by the Input Sequencer. The frames are transmitted in the same order as they are processed.

2.2.3. *Raster Line to Block/Macroblock Converter*

The basic DCT *transform unit* is an 8 by 8 pixel block called a *basic block* or simply a *block*. The basic *quantization unit* is four adjacent *blocks* of Y, and one U and one V *blocks*. Such a quantization unit is called a *macroblock*, as shown in Figure 2.1. The conversion from raster lines format to the *block* and the *macroblock* format is done by the Raster Line to Block/Macroblock Converter.

The next few sections describe the I, P, and B processing in detail. The pictures are assumed to be in the proper process sequence and in block or macroblock format. For all three frame types, the quantization is adjusted to achieve an equitable relationship between the data rate and the image quality. Zigzag scanning (see Figure A.3) is used to order all DCT coefficients.

2.2.4. *I-frame Processing*

An I frame is processed by an intra-frame adaptive DCT coder. No motion compensation is performed since the I-frame compression does not involve any temporal prediction. However, I-frames are used in motion estimation and compensation of other P and B frames. Therefore an I frame is stored in one of the two anchor frame stores indicated in Figures 2.5, and 2.7.

The compression encoding proceeds a macroblock at a time. The six blocks in a macroblock are each processed by a DCT processor which manipulates the pixel values to generate a macroblock of DCT coefficients. (See Appendix A for further details.) These coefficients are then quantized according to a bit-allocation scheme.

The first coefficient of a DCT transform block represents the average value of the block (Y, U, or V) and is referred to as the DC value. A fixed quantizer is applied to this DC value. The remaining coefficients represent the higher frequency (in the DCT sense) and are referred to as the AC values. The AC values are first weighted by a down-loadable quantization matrix before a uniform adaptive quantization. The quantization step size for the AC coefficients is controlled by the Rate Controller. The Rate Controller is discussed elsewhere in this document.

The inverse operations are done at the receivers. The quantization parameters for the AC coefficients are sent to the receiver as macroblock header information. The I-frame reconstruction also occurs at the transmitter so that both the transmitter and the receiver will have the exact same copy to use in the P- and B-frame processing. As mentioned, a reconstruted I frame is stored in one of the two anchor frame stores.

2.2.5. *P-frame Processing*

The predicted frames (P frames) comprise two basic data components: the motion compensation information, and the prediction residue or intra-frame information. The P-frame motion compensation is always in the *forward* direction, i.e., motion is always referenced to the *past* anchor frame. P-frame motion estimation uses exhaustive search over a search region, with half-pixel accuracy as described in Appendix A. The search region is proportional to the number of B frames between two consecutive anchor frames.

The P-frame macroblock may be in one of several modes; each mode relates to the use of temporal predictive coding with or without motion compensation, and intra-frame coding. Typical criteria for selecting a particular mode for a macroblock are minimum absolute temporal difference and minimum variance.

The macroblocks, whether in the form of motion-compensated prediction residue or intra-frame pixel values, are transformed by a DCT processor and then quantized. For intra-macroblocks (i.e., macroblocks coded using intra-frame process), the DCT coefficient quantization is identical to that used for the I frames. For motion-compensated macroblocks, the DCT coefficients are quantized using a uniform quantizer. The DC and the AC coefficients have equal quantization. The quantizer step size is also controlled by the Rate Controller.

The inverse operations are done at the receivers. The quantization parameters are sent to the receiver in the slice header. P-frame reconstruction occurs at the transmitter so that both the transmitter and the receiver will have the exact same copy to use in the B-frame processing. As with the I frames, a reconstruted P frame is stored in one of the two anchor frame stores.

2.2.6. B-frame Processing

The main difference between the P-frame and the B-frame processing is in the motion processing. Unlike the P-frames, the B-frames are subjected to *bidirectional* motion estimation/compensation. The motion references are the two anchor frames sandwiching the B frames. In the Group-of-Pictures structure shown in Figure 2.3, there are two B frames sandwiched between every pair of I and/or P anchor frames.

Like P-frame macroblocks, the B-frame macroblocks have a number of modes. In addition to all the modes for a P-frame macroblock, the B-frame macroblock modes further include a *bidirectional* interpolative mode in which both forward *and* backward motion compensation are used, and a *unidirectional* mode, i.e., either forward-only motion compensation, or backward-only motion compensation. In the interpolative mode, an average of the forward and the backward motion-compensated macroblocks is used as the prediction macroblock. As in the P-frame motion estimation, the B-frame motion estimation uses the method of exhaustive search. The search regions are proportional to the temporal distance between the B frame and the two anchor frames.

The process of DCT transformation and the subsequent adaptive quantization for the B-frame macroblocks is identical to that for the P-frame macroblocks, i.e., zigzag scanning, and uniform quantization.

The B frames are not used as prediction for other frames, so the encoder does not perform B-frame reconstruction.

2.2.7. *Differential, Run-Length, and Variable-Length Coding (VLC)*

Figure 2.6 shows the processing of the quantized DCT coefficients and other non-pixel information. Appendix A contains a detailed description of runlength coding and VLC.

The quantized DC coefficients of all the I-frame macroblocks are differentially coded with a DPCM (differential pulse code modulation) coder. P- and B-frame macroblocks in intra-frame mode are also coded in similar fashion.

For the quantized AC coefficients after the zigzag scan ordering, the zero runs and following non-zero values are coded with VLC. For the less likely combinations of zero runs and non-zero values, fixed-length coding is used. These fixed-length codes are preceded by a special code. In addition, an end-of-block (EOB) code is used to indicate when all the remaining coefficients in a block are zero.

Motion vectors of the same type are differentially coded.

Macroblock addresses are referenced to the respective positions of the macroblocks within a slice. A macroblock whose motion vector data is zero and whose quantized coefficients are also zero is not transmitted; these macroblocks are called *fixed*. Only macroblocks with data are transmitted; these macroblocks are called *non-fixed*. The address of the first *non-fixed* macroblock in a slice is its absolute address within that slice. Subsequent *non-fixed* macroblocks are addressed relative to previous *non-fixed* macroblock. Another way to view this is to consider the macroblock addresses as the number of *fixed* macroblocks between any two consecutive *non-fixed* macroblocks within a slice.

In addition, variable-length coding (see section on VLC in Appendix A) is applied to all the coded information: motion vectors, macroblock addresses, block types, etc.

2.3. RECEIVER ERROR RECOVERY

In the event of any channel errors, the unique situation of the broadcast environment prevents any re-transmission effort from the transmitter to any particular receiver or sets of receivers. Instead, some receiver error recovery measures must be carried out by the individual receivers. Although the ADTV system specification supports error recovery, as in most aspects of television receiver design, the actual level of recovery to be carried out in a receiver is at the discretion of the receiver manufacturers.

3. PRIORITIZATION ENCODER/DECODER

Figure 3.1 is a functional block diagram of the Prioritization Encoder which comprises the Priority Processor and the Rate Controller. In addition to communication with the video encoder, the Rate Controller passes dynamic priority allocation information to the Priority Processor. This information is updated at the buffer control intervals (e.g., at start of Groups of Pictures, start of frames, start of slices, and/or start of macroblocks). More will be said on this in a later section on the Rate Controller. The Priority Processor uses the information from the Rate Controller in its own algorithm for a dynamic priority assignment to the various data elements.

3.1. PRIORITY PROCESSOR

Every data element gets a priority assignment from the Priority Processor. The approach of priority assignment is equivalent to an asynchronous *codeword* multiplex scheme in which each codeword or data element is multiplexed to one of the priority classes according to the assigned priority for that data element.

Figure 3.2 is a flow diagram showing the logical sequencing of operations to bring about the desirable allocation in the data into two major priority classes: high priority (HP) and low priority (LP).

The Priority Processor has a number of tasks. Each task can be thought of as a functional module. The two primary tasks/modules are:

1) Target-allocation of HP/LP traffic per Group of Pictures
Based on bits allowed by the Rate Controller for each frame of a particular Group of Pictures, a *pre-allocation* is made of the HP and the LP bits for every frame in that Group of Pictures. The HP/LP fractional allocations may vary with the frame type.

2) HP/LP Priority Assignment
The HP/LP allocation of the coded bits from each frame is carried out causally, i.e., after the bits to be transmitted have been generated by the encoder and stored in a HP/LP decision buffer.

To provide a dynamic HP/LP traffic allocation, the Priority Processor also provides the following functionalities:

3) Update of the HP/LP allocation within a Group of Pictures
The functionality is similar to task 1 discussed above. This module updates the fractional rate allocations before coding a frame in a Group of Pictures. Based on information from all

previously coded frames within a Group of Pictures (such as how close the total bit allocation for the coded frames has been met by the rate controller, the current HP/LP allocations, etc.) the initial setting for the HP/LP allocation for a new frame within a Group of Pictures is modified. Recall from task 1 that pre-allocation of every frame in a Group of Pictures precedes the priority-allocation process of that Group of Pictures.

4) HP/LP allocation updates within each frame

The HP and LP bit counts are updated after each portion of a frame corresponding to a *priority analysis interval* (PAI) has been processed and its data passed to the next processor. The PAI depends on the size of the HP/LP decision buffer and may comprise multiple slices. The updated count is used to fine-tune the fractional allocations for the next portion of the frame to be processed. This is necessary to compensate for the transport overhead that is not included in the decision process of task 2.

3.1.1. Codeword Ranking

The data within the HP/LP decision buffer consists of information from many coded DCT blocks. These data elements are ranked in terms of their relative importance within the context of a particular frame, namely the I, P, and B frames. In general, data elements for each block are assigned High/Low Priority based on their positions in the rank list. Within a macroblock, the codeword order is assumed to represent the priority rank order. For example, the codewords for all the DC coefficients for a macroblock come before those for the AC coefficients. Table 3.1 shows the rank list for the different types of frames. Only the video data elements are ranked and allocated into the HP and LP classes as described above. All other data elements (audio, and auxiliary data) are always assigned High priority.

I-frames	1. Headers 2. DC values 3. Low frequency coefficients 4. High frequency coefficients
P- and B-frames	1. Headers 2. Motion vectors 3. DC values 4. Low frequency coefficients 5. High frequency coefficients

Table 3.1: Priority Rank List for Different Frames.

3.1.2. Receiver Operations

The receiver Priority Processor functions mainly as an asynchronous codeword demultiplexer, with some additional functionalities related to error tracing of data elements. The receiver Transport Processor provides data elements as well as other control information such as the priority-allocation decision to the Priority Processor. With the help of the control information, the Priority Processor reconstructs the priority rank list from the received data elements.

Information regarding data elements that are missing or have been flagged by the error indicator coming from the receiver Transport Processor will be passed on to the appropriate processors in the next layer. Because data elements are transported in *data groups*, when data elements within a data group have been contaminated by channel errors, it is the responsibility of the receiver Priority Processor to provide appropriate error indicators for these data elements to their respective Compression Decoder processor. While the same kind of information for all data elements can be provided to their respective Compression Decoder processors, the actions taken by these processors will be different for different type of data elements. For example, an error on, say motion vectors for a particular macroblock or slice in a B frame, may trigger the receiver error concealment processor to perform some specific video-level error concealment tasks.

3.2. RATE CONTROLLER

Another important functionality provided by the Priority Encoder is the channel rate regulation. This is done by the Rate Controller. Figure 3.1 shows the functional modules within the Rate Controller. As mentioned before, the Rate Controller monitors the status of the rate buffers in the Transport Encoder. It uses the buffer occupancy information to compute the necessary compression requirement and feeds the results in the form of appropriate quantization parameters to the Video Processor in the Compression Encoder. The Video Processor will use the information when it processes the next picture *slice* or *macroblock*. The Rate Controller also provides input to the Priority Processor regarding the initial allocation of HP/LP rate for the next Group of Pictures.

3.2.1. HP/LP Rate Allocation

With knowledge of the buffer status (i.e., how close to fullness, how fast it has been changing, etc.) the HP/LP target allocation is computed. Typically, the allocation specifies the ratio of bit rates to be used for the HP and LP data by means of an HP:LP target rate ratio. An algorithm computes the ratio under the constraint of maximum image quality *in the presence of channel losses*, for a given HP:LP target rate ratio.

3.2.2. Quantization Control

Similar to the HP/LP rate allocation process, the Rate Controller dictates to the video compression processor an appropriate set of compression parameters to be used for the next slice or macroblock of picture. Based on the actual data rates used for the I, P, and B frames in previous and current Groups of Pictures, the Rate Controller derives an initial data rate allocation for the I, P, and B frames in the next Group of Pictures. The initial I:P:B rate allocation is then computed. The initial rate allocation is computed with the objective of achieving approximately equal image quality within a Group of Pictures while keeping the rate allocation within some defined range.

4. TRANSPORT ENCODER/DECODER

The Transport Encoder/Decoder (TE/TD) comprise the cell relay functionalities and the rate buffers.

4.1. TRANSPORT PROCESSOR

4.1.1. Transmitter Operations

Data elements are supplied to (from) the Transport Encoder (Decoder) from (to) the Prioritization Encoder (Decoder). Communication regarding the size of the a data element, its type, and its priority, is via separate control data path. Additional slice- and/or macroblock-level information that is needed for the *segmentation and chaining* of data cells is also passed between the Transport Processor and the Priority Processor.

The Transport Processor generates appropriate header fields for *data groups*. A *data group* is a set of data elements that are to be transported together. For example, video *data groups* comprise data elements belonging to the same *slice*. Similarly, audio *data groups* comprise data elements of the same MUSICAM *audio data frame*.

The header fields (e.g., Service Type, Transport Header, etc.,) are used in the construction of a basic transport unit called a *cell*. A cell resembles a data packet in conventional packet networks in modern data communication. It has a header and a trailer enclosing a payload area. Each cell has a fixed size, although the actual size of the header fields varies depending on what is being carried in the payload. The trailer field contains error-checking information bits, and is fixed-size for all cells.

The transport header contains chaining and segmentation information which allows *data groups* to be segmented across cells. This feature provides a high degree of efficiency in the use of fixed-size cells while limiting the propagation of channel error from one cell to the next. For instance, when a cell is found to be in error, its payload is typically unusable. With the header information, however, any adjacent *data groups* that are not transported by the erroneous cell are not affected by the loss of that particular cell.

4.1.2. Receiver Operations

Because of their fixed size, cells are easily identified once receiver synchronization is established. The following discussion assumes cell synchronization.

Every received cell is first processed by a CRC (cyclic redundancy code) processor, which provides reliable detection of most errors that may occur in a cell. If a cell is found to contain bit error(s), its associated error indicator is set so that the Transport Processor and/or any subsequent processors can take appropriate actions. Cell sequencing is always maintained at the receiver, regardless of the error status of the cells. The Transport Processor at the receiver will make use of the cell Transport Header information to achieve *data group* synchronization immediately following an erroneous cell. In this way, the Transport Processor knows which data group is in error, and passes such information to the receiver Priority Processor.

4.1.3. Data Format

Figure 4.1 shows the logical structure of the data format. The link-level description of a cell applies to *all* data types, i.e., video, audio, and auxiliary data. The remaining cell structure descriptions apply only to the video cells.

Current ADTV design supports fixed-size cells of 256 bytes long. At the link-level, a cell is logically divided into three parts: the service-type (ST) header, the frame-check-sequence (FCS) trailer, and of course the *transport payload* [3] itself.

The ST header may contain information regarding the priority (High/Low) associated with the cell, and the type of service/application for which a particular cell is used. The ADTV prototype hardware supports three service types: video, audio, and data.

The FCS trailer is a 16-bit error-checking sequence for a particular cell. An example of such a sequence is the standard CRC as defined by CCITT. Frame checking is applied over all bits in a particular cell.

The transport payload carries service-specific data. For instance, a teletext data payload may contain data that only a teletext decoder can use, and an audio payload contains audio data.

There are two components within a video transport payload: the Transport Header (TH), and the *service payload.* The information contained in the TH provides the *chaining and segmentation* capability of ADTV transport cells, while providing an error indication to the next level of the data format.

The video service payload contains primarily compressed video codewords. At the start of a *slice*, the video service payload of the corresponding HP or LP cells will begin with a record header (RH) before the video codewords. The video service record header appears at the start of every *slice*. Record headers carry explicit information for identifying and associating video codewords in a cell to the overall Group-of-Pictures structure.

3 At different level within the cell structure, the concept of payload applies. To make the distinction among payloads, we refer to the link-level payload as *transport payload*, the service-level payload as *service payload.*

4.2. RATE BUFFERING

4.2.1. Transmitter Buffering

Although some minimal buffering is required during the process of putting the *data groups* into fixed-size transport cells, these cells are generated asynchronously in that over any unit of time, the number of cells generated is not constant. The channel coding module interfaces with the Transport Processor at a fixed clock rate. Therefore the Transport Processor must be able to completely smooth out any rate variation as a result of any preceding asynchronous processes. Elastic rate buffers are used for the purpose of smoothing out the rate variation. One important consequence of having elastic buffers is that there will be an associated maximum end-to-end video delay. The maximum delay is dependent on the size of the buffers.

4.2.2. Receiver Buffering

The receiver contains complementary buffers to the encoder buffers so that the overall delay introduced by the transmitter-receiver buffers to the end-to-end system is *fixed*. This fixed delay equals the maximum buffer delay.

4.2.3. Receiver Group-of-Pictures Synchronization

It is important that the receiver buffers operate at fullness levels complementary to those of the transmitter buffers. For each Group of Pictures, the transmitter buffer states are among the overhead information sent to the receivers. The Sync-and-Timing-Setup module at the receiver extracts the information from the received cells and informs the receiver rate buffers so that the complementary relationship between the transmit buffers and the receive buffers is maintained. Receiver Group-of-Pictures synchronization is established when the receiver first powers up, and every time it switches to a new channel.

5. MODEM-FEC ENCODER/DECODER

5.1. FORWARD ERROR CORRECTION (FEC)

The channel coding module has the primary responsibility of reducing the data transmission error rate to an acceptable level. The FEC module interfaces with the Transport Processor on the one hand, and the Modem on the other.

The FEC module is responsible for segmenting the data in blocks which are amenable for the Reed-Solomon encoding. To mitigate the effects of burst errors, interleaving is done over multiple coded blocks at the encoder, and de-interleaving at the decoder. The de-interleaved coded blocks then contain random bit errors even when the channel errors are bursty. In general, random bit errors are more amenable to Reed-Solomon error correction.

5.2. R.F. MODEM

The transmitted signal of the ADTV modem is designed specifically for the transmission environment in the television broadcast bands. A particular feature of the ADTV signal is its exceptional robustness towards interference by NTSC signals. Another feature of the transmitted ADTV signal is that its spectrum is shaped to minimize apparent interference in NTSC receivers picking up the ADTV signal as co-channel interference. This affords the ADTV system a duality of protection in the presence of co-channel NTSC signals: both the NTSC-into-ADTV and the ADTV-into-NTSC interference is minimized.

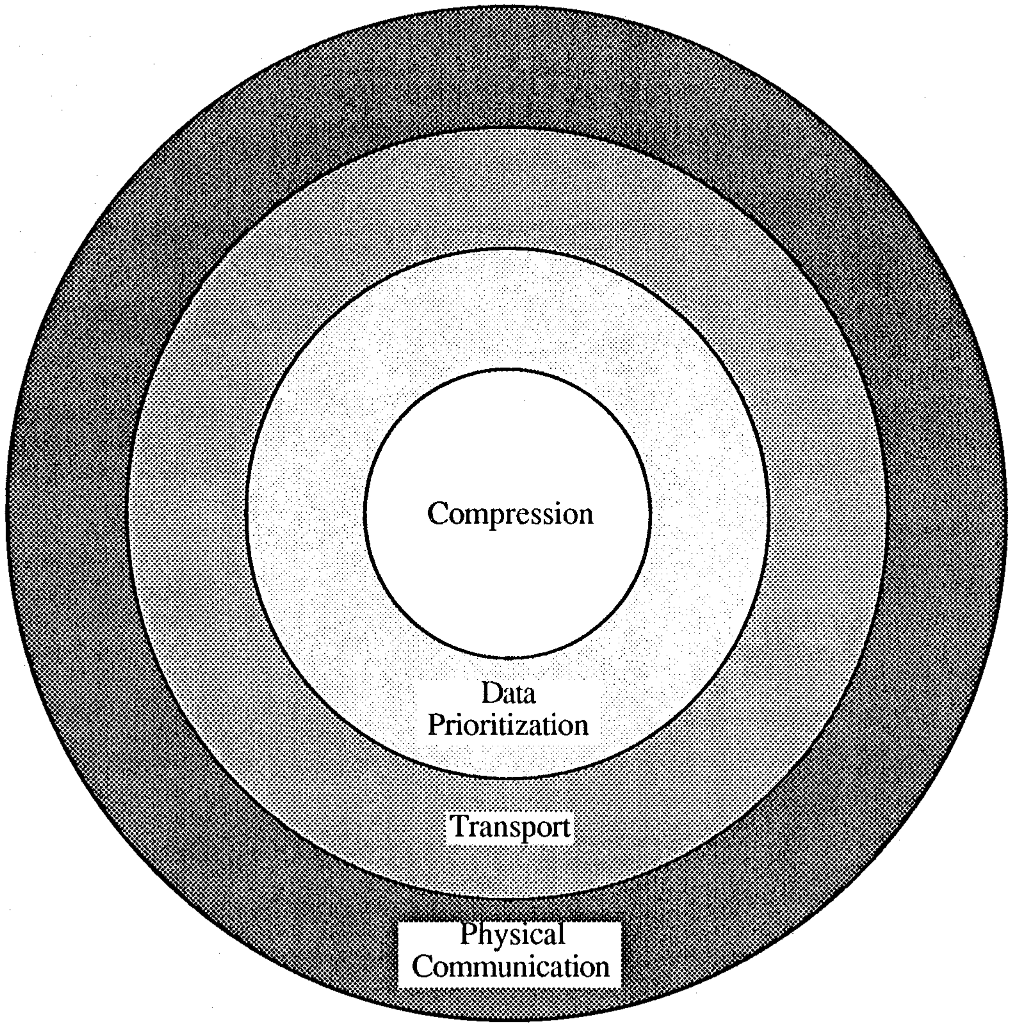
The ADTV modem is based on conventional Quadrature Amplitude Modulation (QAM) technology extensively developed for telephone and digital microwave applications. It uses spectrum shaping techniques to meet the communication requirements of the compression algorithm while minimizing cross-interference between ADTV and existing NTSC services. The ADTV receiver modem is equipped with an adaptive channel equalizer to combat the effects of multipath propagation and ghosts. The modem processing is implemented entirely in digital signal processing circuits that lend themselves to large-scale integration.

6. SUMMARY

Although essential to a broadcast digital HDTV system, data compression alone will not create an entire HDTV transmission system. Once the data have been compressed by the source coder, they must be "packaged" and delivered (transmitted) over the hostile terrestrial broadcast channels. It is a mandatory design requirement by the FCC that any new HDTV service be NTSC co-channel compatible in order to allow the new HDTV signal to be broadcast without requiring additional transmission spectrum. Currently-defined "taboo" channels are available in the sense that there are presently no signals which are transmitted at these taboo frequencies. However, these channels are taboo because a typical NTSC broadcast signal transmitted at these frequencies would interfere with and be interfered by the broadcasts in existence today. This poses an added design requirement on the HDTV system designer to not only reduce the bandwidth requirement of the HDTV signals with efficient source coding, but also design a clever signal packaging technique which will be robust in an NTSC co-channel environment. This low interference requirement applies in two directions: (i) minimal interference from the new HDTV signal *into* NTSC so as not to disrupt existing NTSC services; and (ii) minimal interference *from* NTSC into the HDTV signal so that a high-quality HDTV service can be provided.

A practical solution toward meeting the challenges of the digital simulcast environment is ADTV. The efficient MPEG++ compression, the Prioritized Data Transport format, and the Spectrally-Shaped QAM have been melded into a powerful system approach for digital simulcast of HDTV. The successful integration of the three key layers outlined in the Executive Summary means that ADTV is well poised to provide robust high-quality broadcast digital HDTV service to the American public.

ADTV is a proposal that is built upon a solid foundation of widely-accepted technologies and important emerging standards. This will lead to rapid industry acceptance, but perhaps even more importantly, this approach will establish an important common technology base across many business sectors vital to successful development of a globally-competitive HDTV industry in the United States.



Compression Algorithmic Layer
- Input/Output Pre-/post-processing.
- Data Compression/Decompression

Prioritization Layer
- Separate data into different priority queues.
- Combine data from different priority queues into coherent data streams for decompression.
- Receiver error control.

Transport Layer
- Data multiplexing/demultiplexing.
- Cell-relay formatting.
- Error detection.
- Delay-compensating buffering.

Physical Communication Layer
- Channel coding (FEC).
- Physical modem modulation.
- Channel adaptive data equalization.
- Frequency translation to/from r.f.

Figure 1.1: ADTV Layered Architecture.

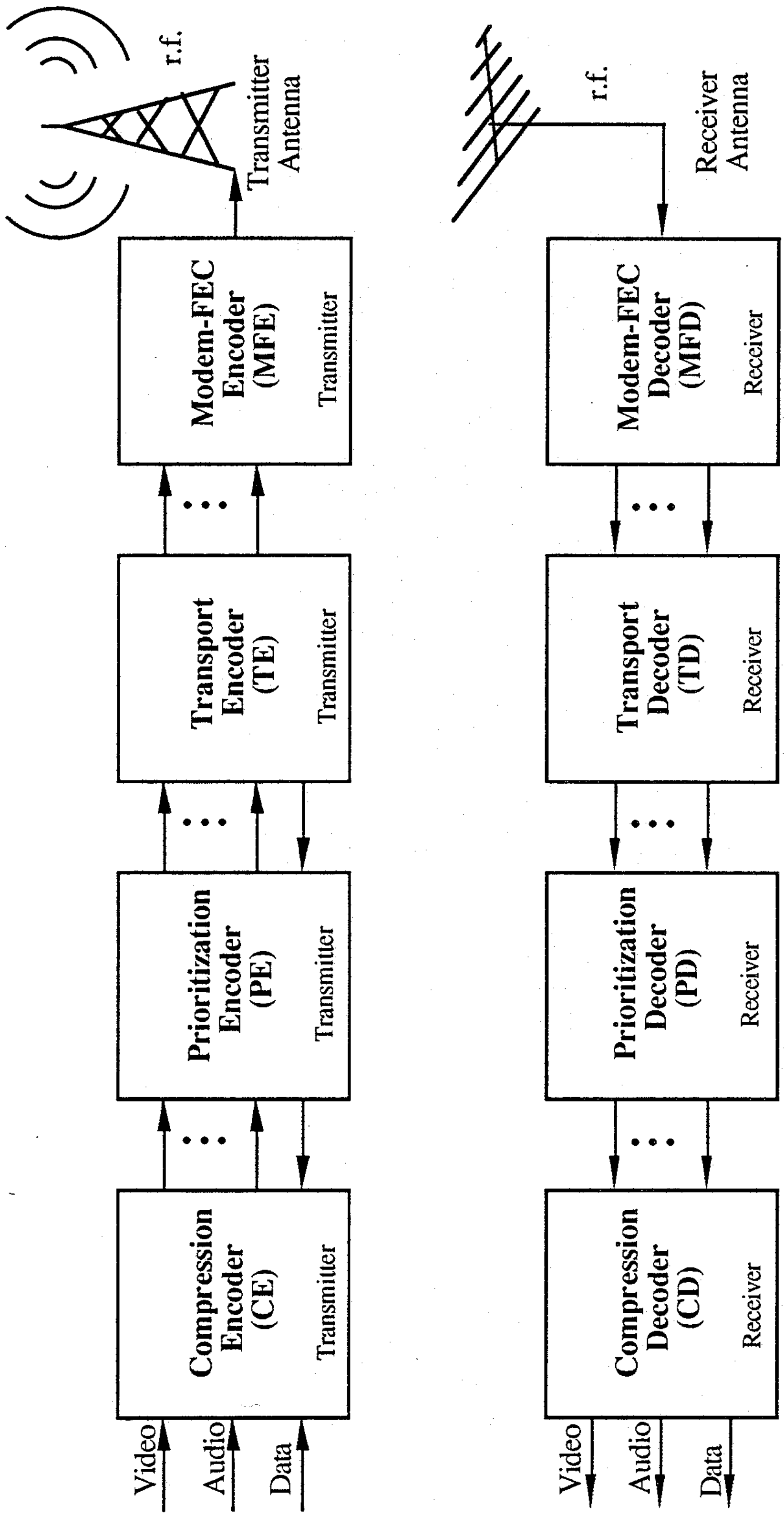


Figure 1.2: Overall System (Transmitter to Receiver).

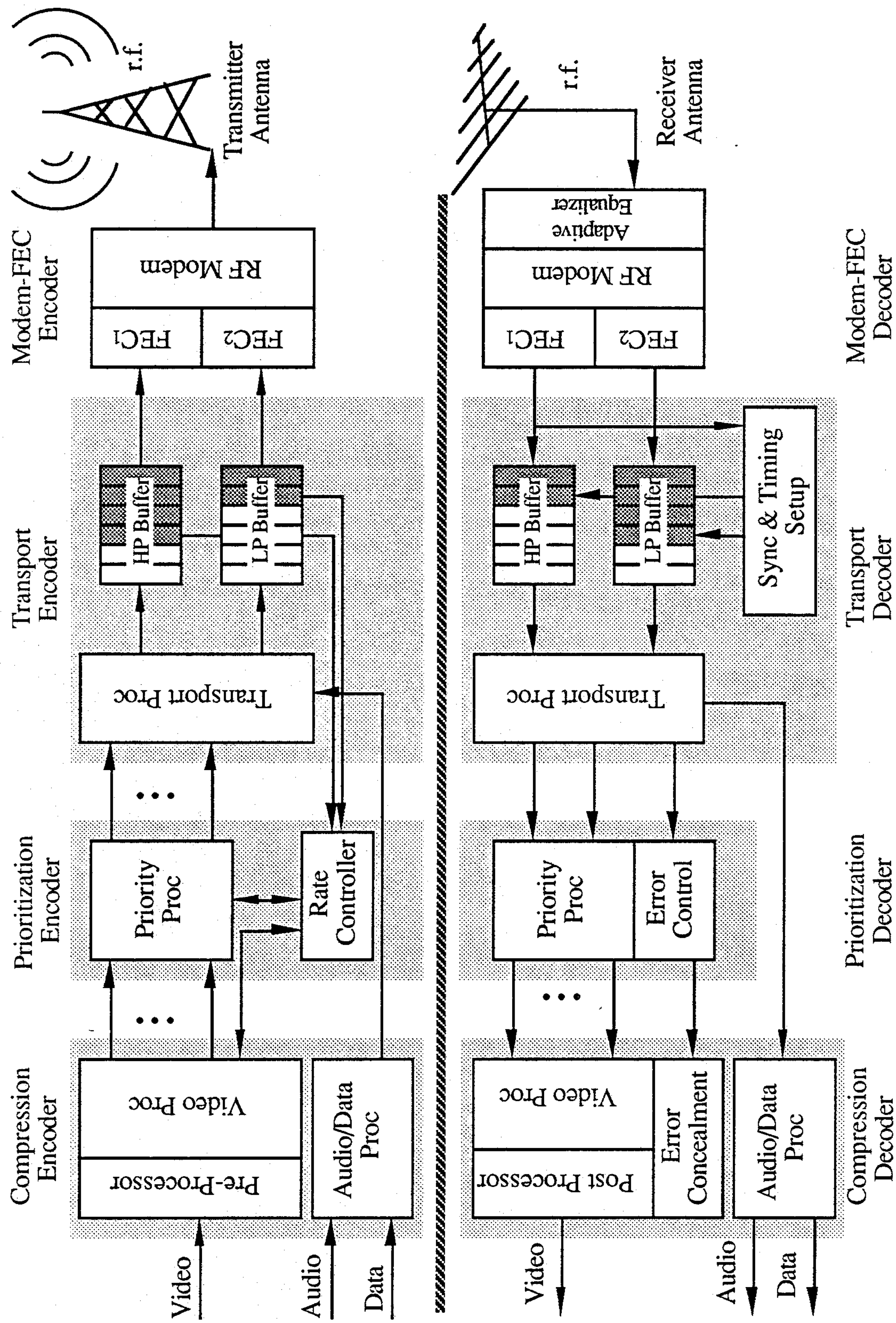


Figure 1.3: Typical Blocks/Modules of Each Layer.

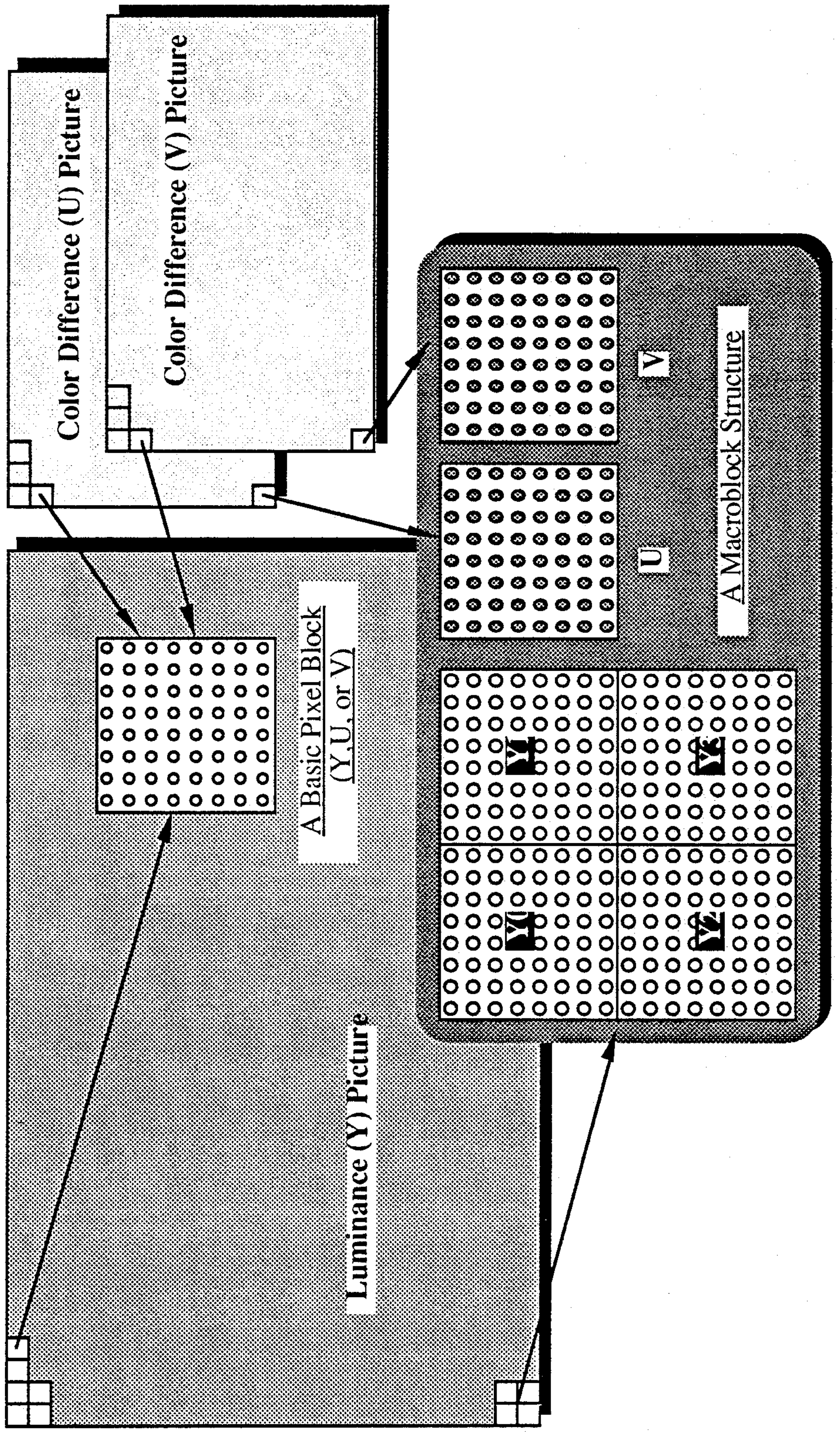


Figure 2.1: Picture Basic Block and Macroblock.

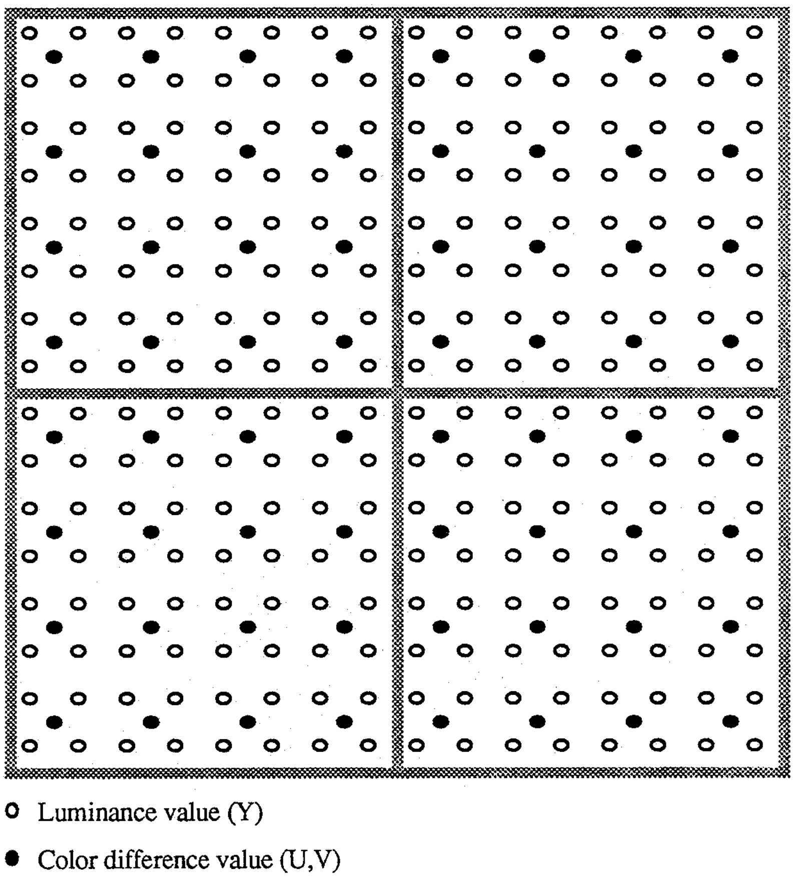


Figure 2.2: Spatial Relationship of Y, U, V Values in a Macroblock.

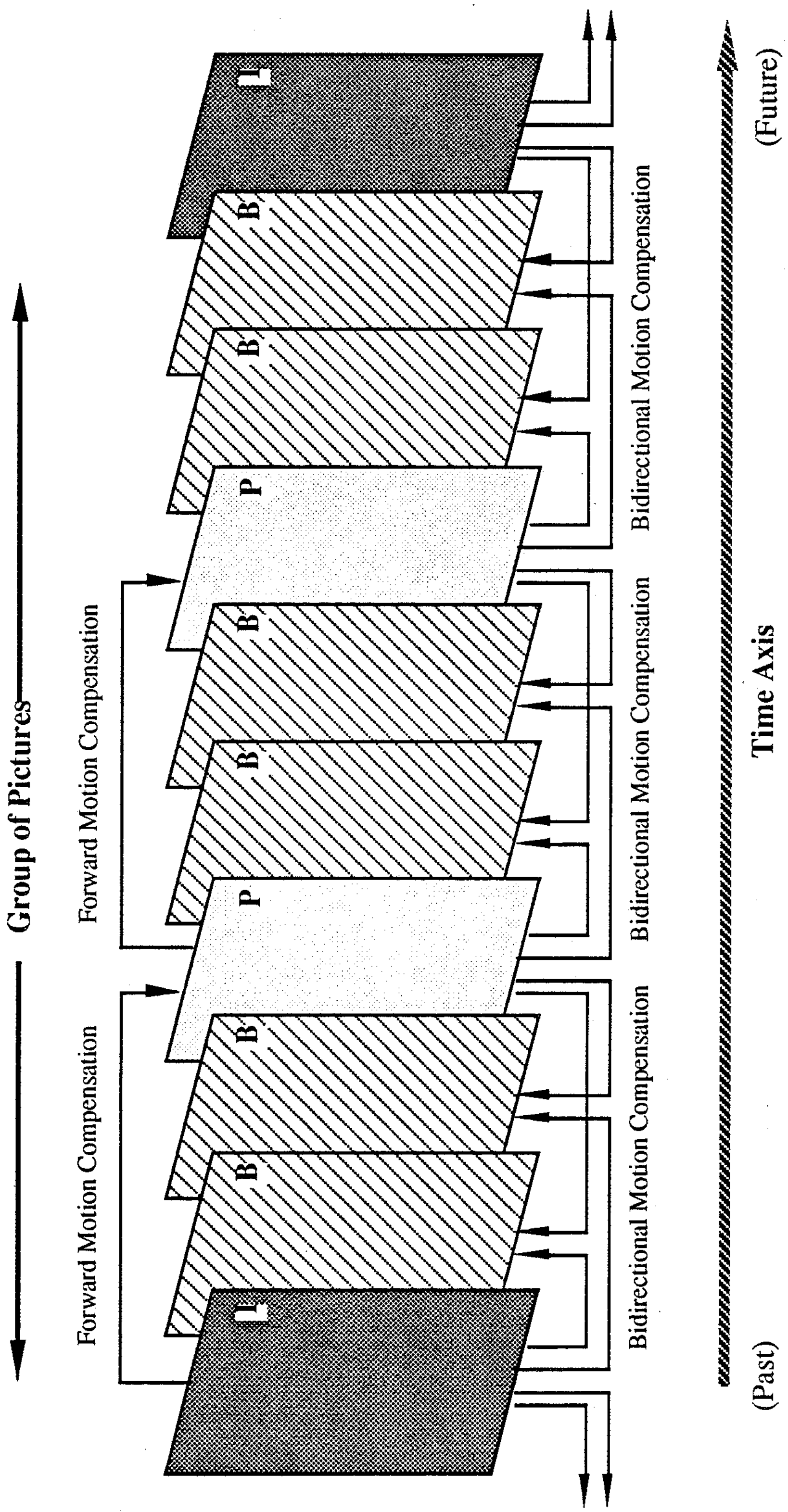


Figure 2.3: An Example of a Group of Pictures.

$$\cdots\ P_{-3}\ B_{-2}\ B_{-1}\ I_0\ B_1\ B_2\ P_3\ B_4\ B_5\ P_6\ B_7\ B_8\ I_9\ \cdots$$

(a) Frame Display Order

$$\cdots\ P_{-3}\ B_{-5}\ B_{-4}\ I_0\ B_{-2}\ B_{-1}\ P_3\ B_1\ B_2\ P_6\ B_4\ B_5\ I_9\ \cdots$$

(b) Frame Process and Transmission Order

Figure 2.4: Display and Process/Transmit Frame Order

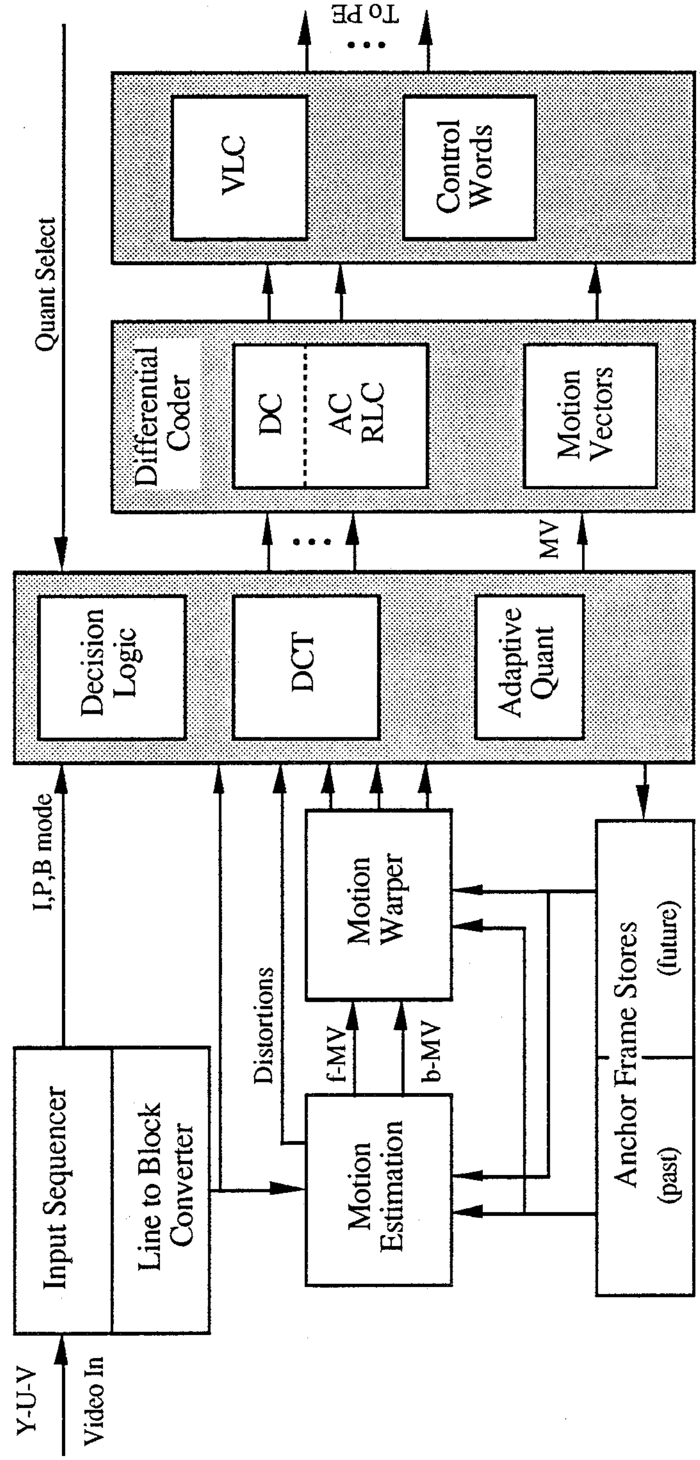


Figure 2.5: Video Compression Encoder.

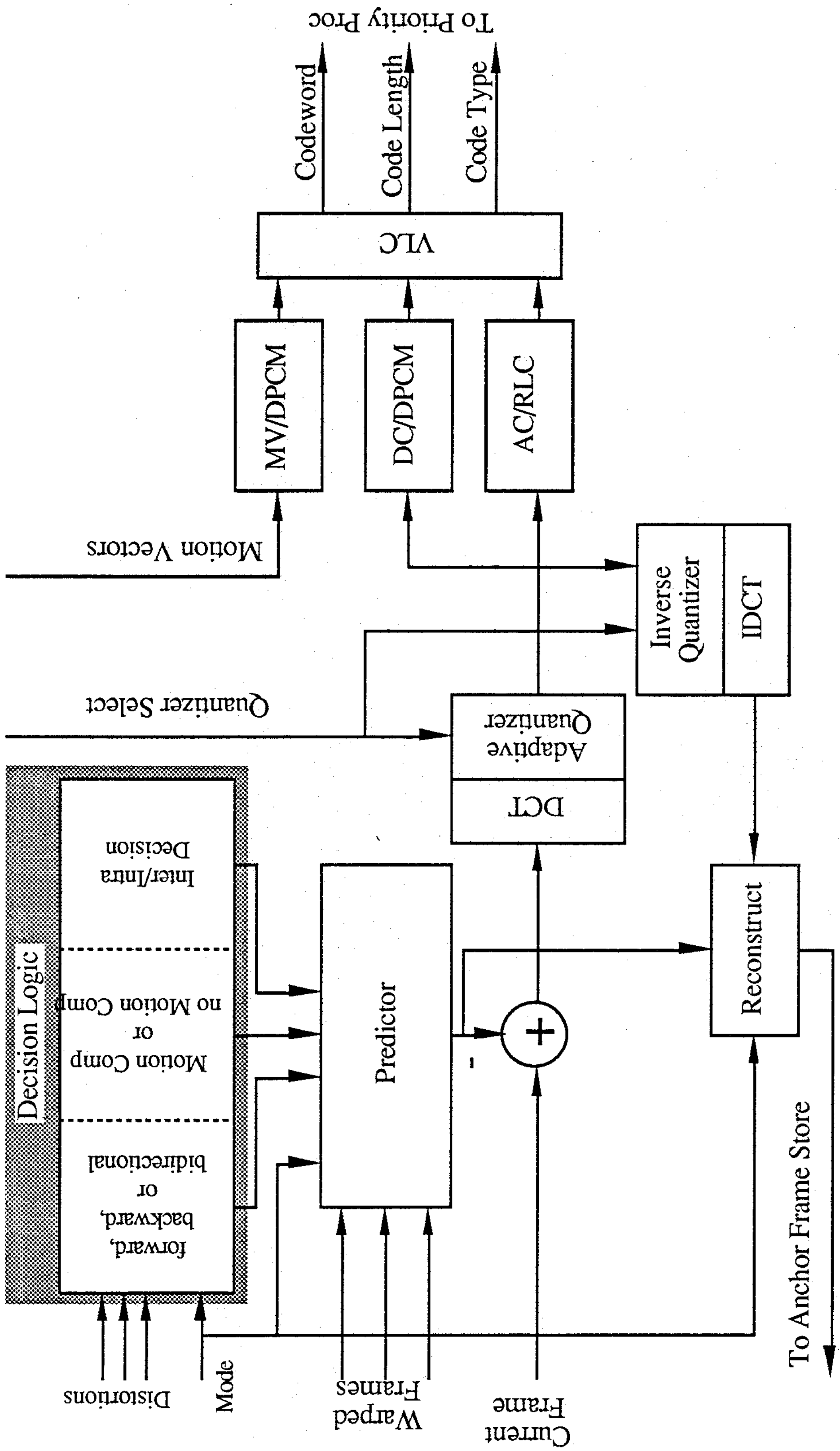


Figure 2.6: Encoder DCT Block.

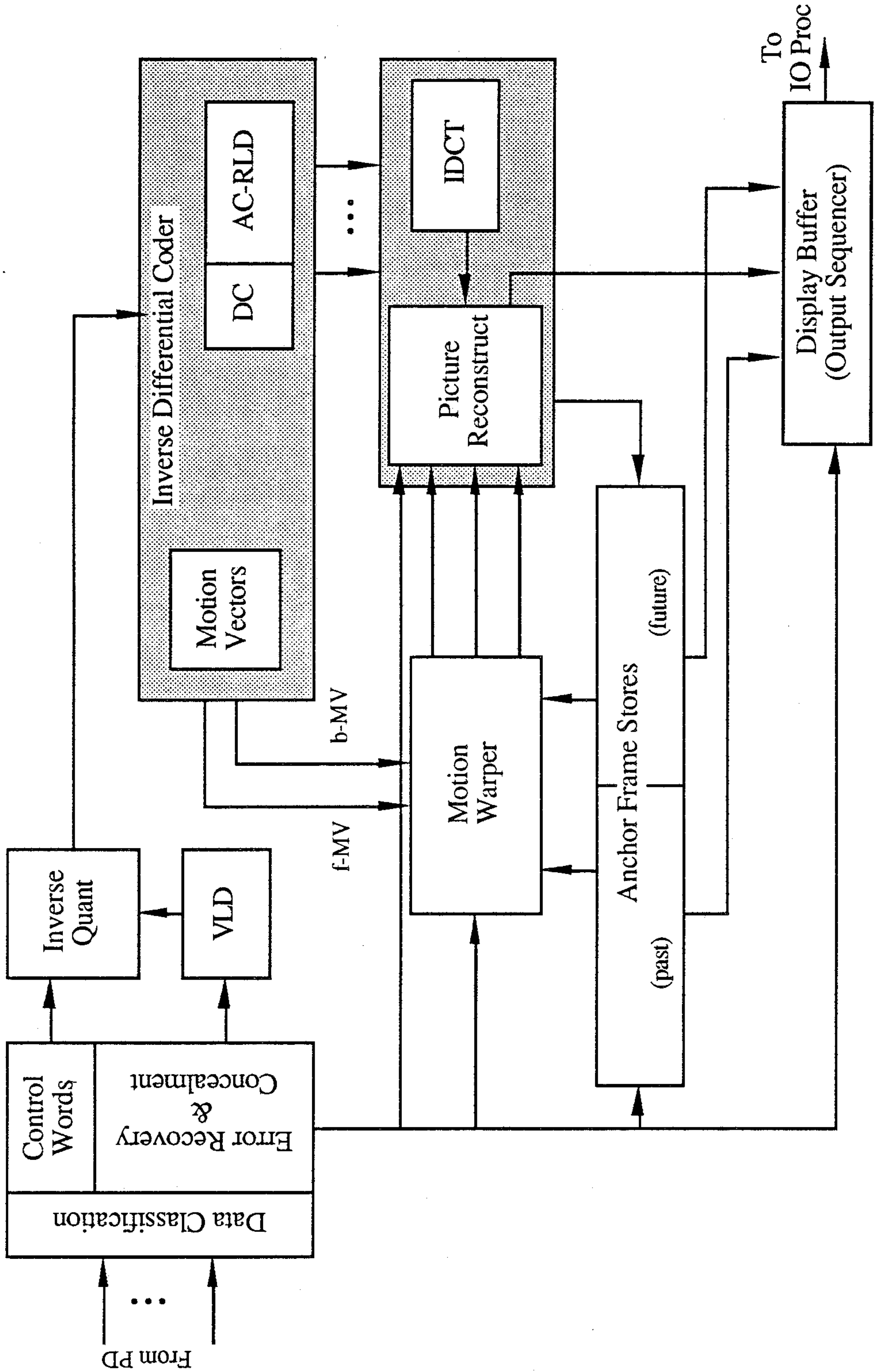


Figure 2.7: Video Compression Decoder.

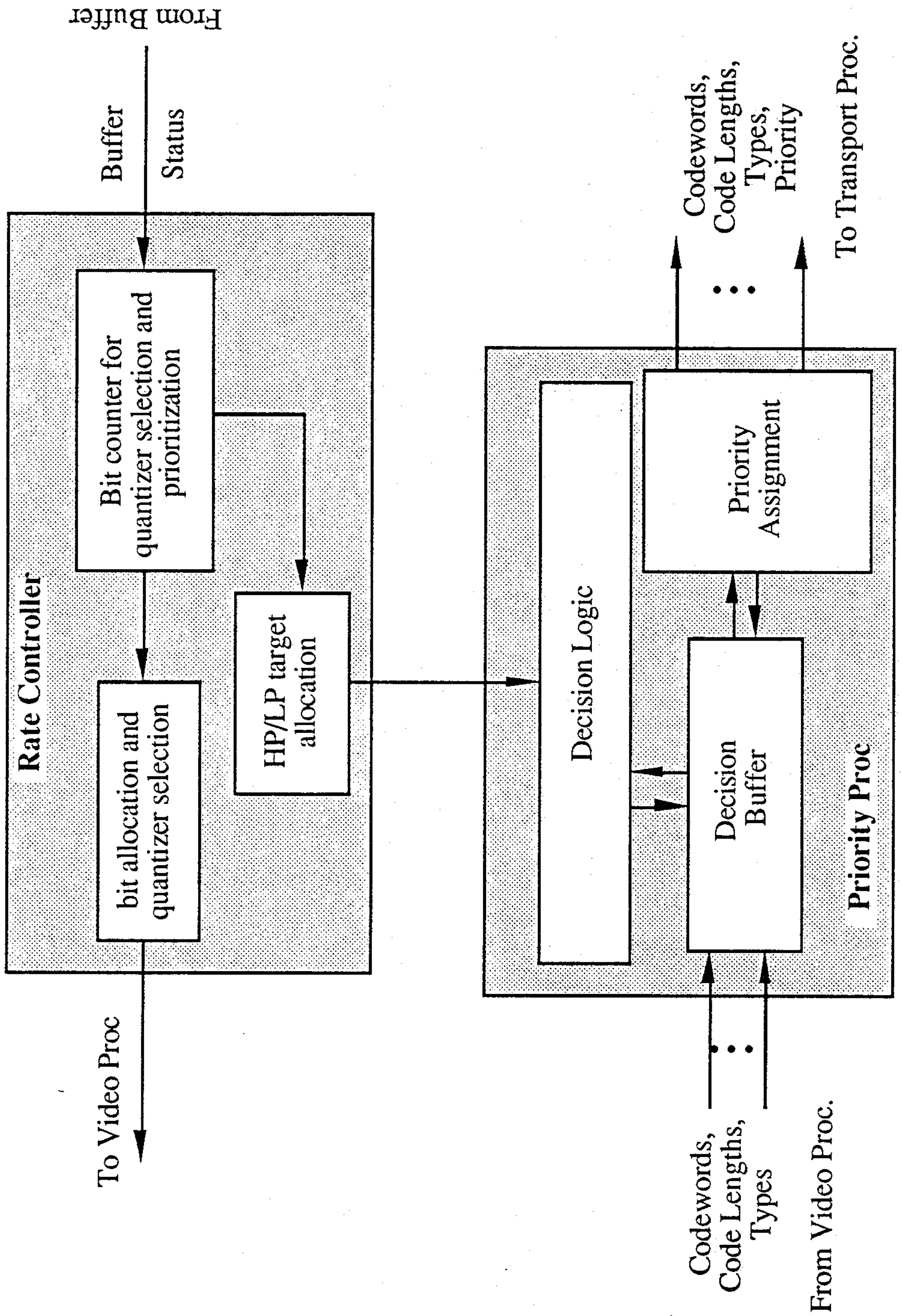


Figure 3.1: Data Prioritization Encoder.

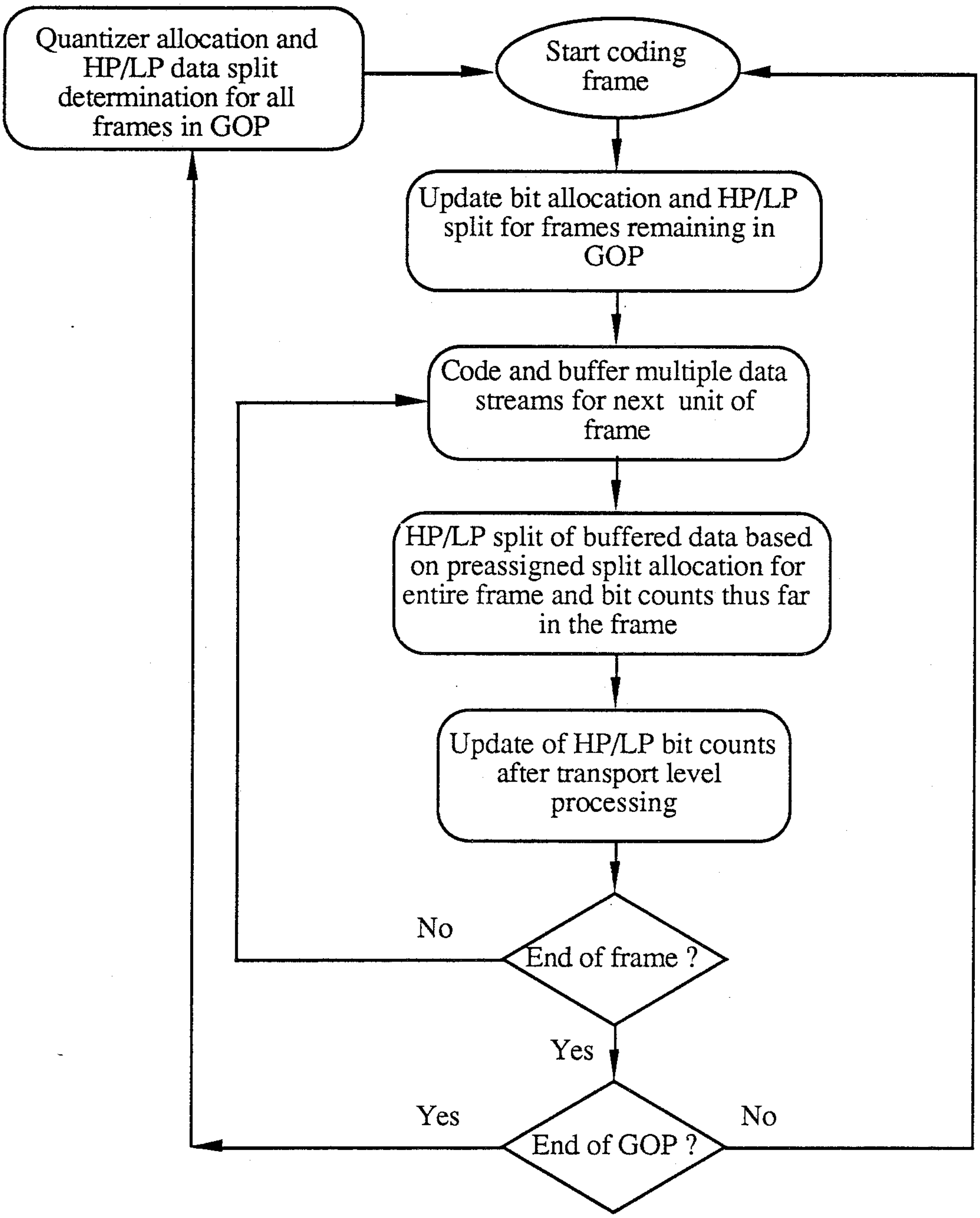


Figure 3.2: Flow Diagram of Priority-Split Logic.

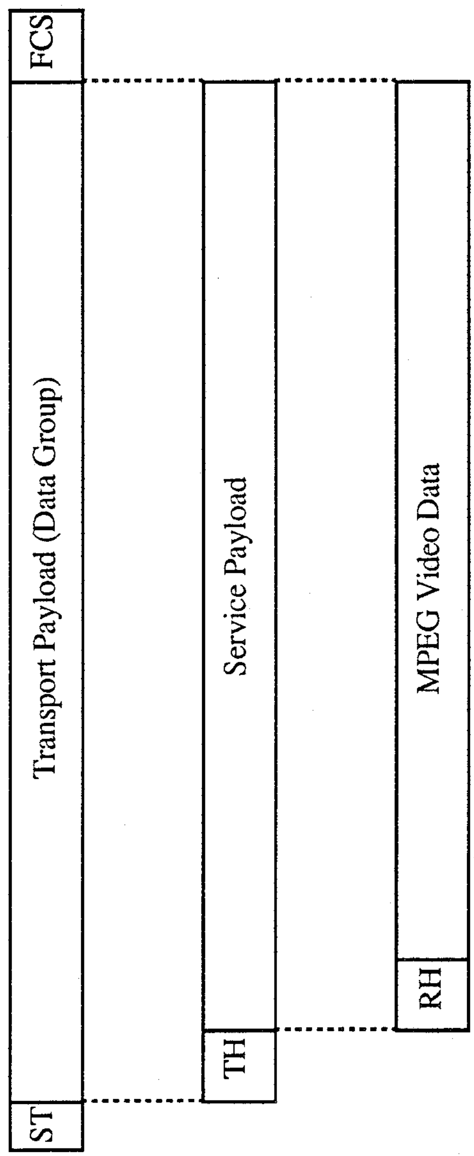


Figure 4.1: Transport Cell Structure.

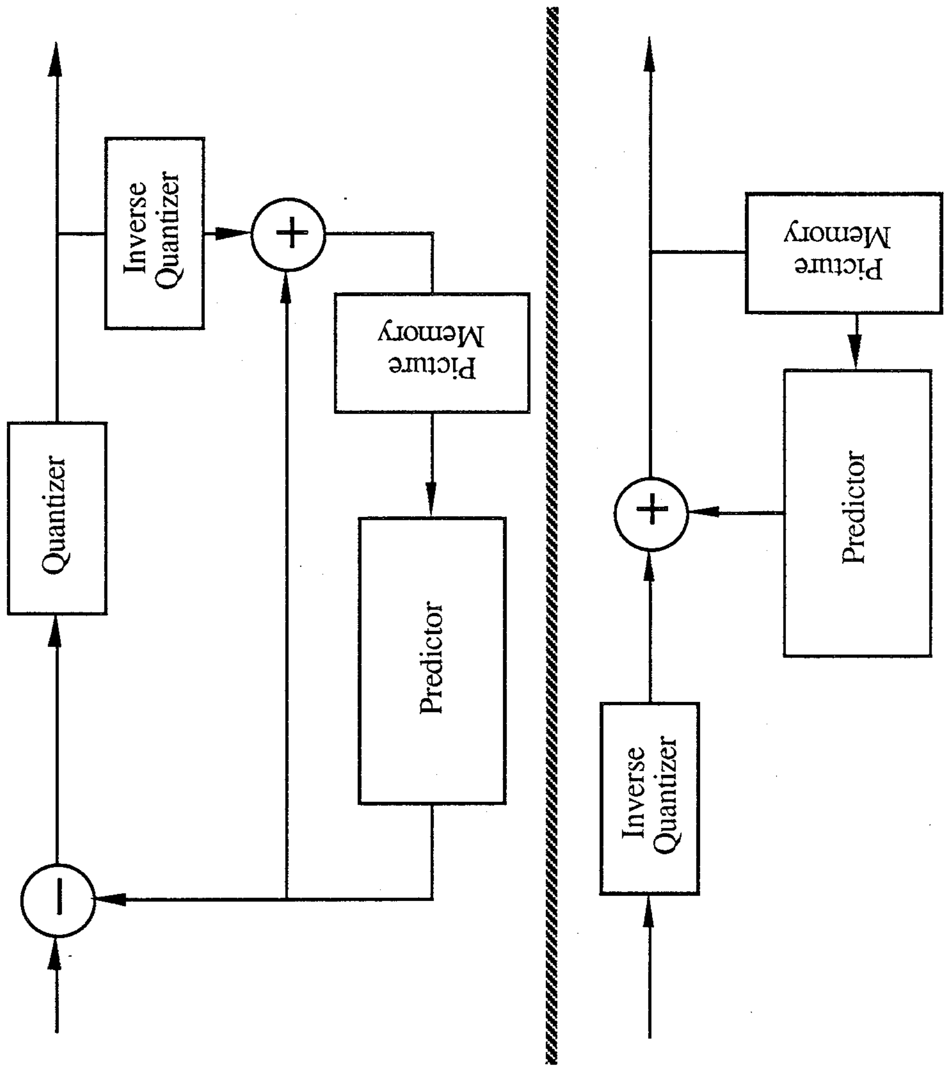


Figure A.1: A Generic Predictive Video Compression Codec.

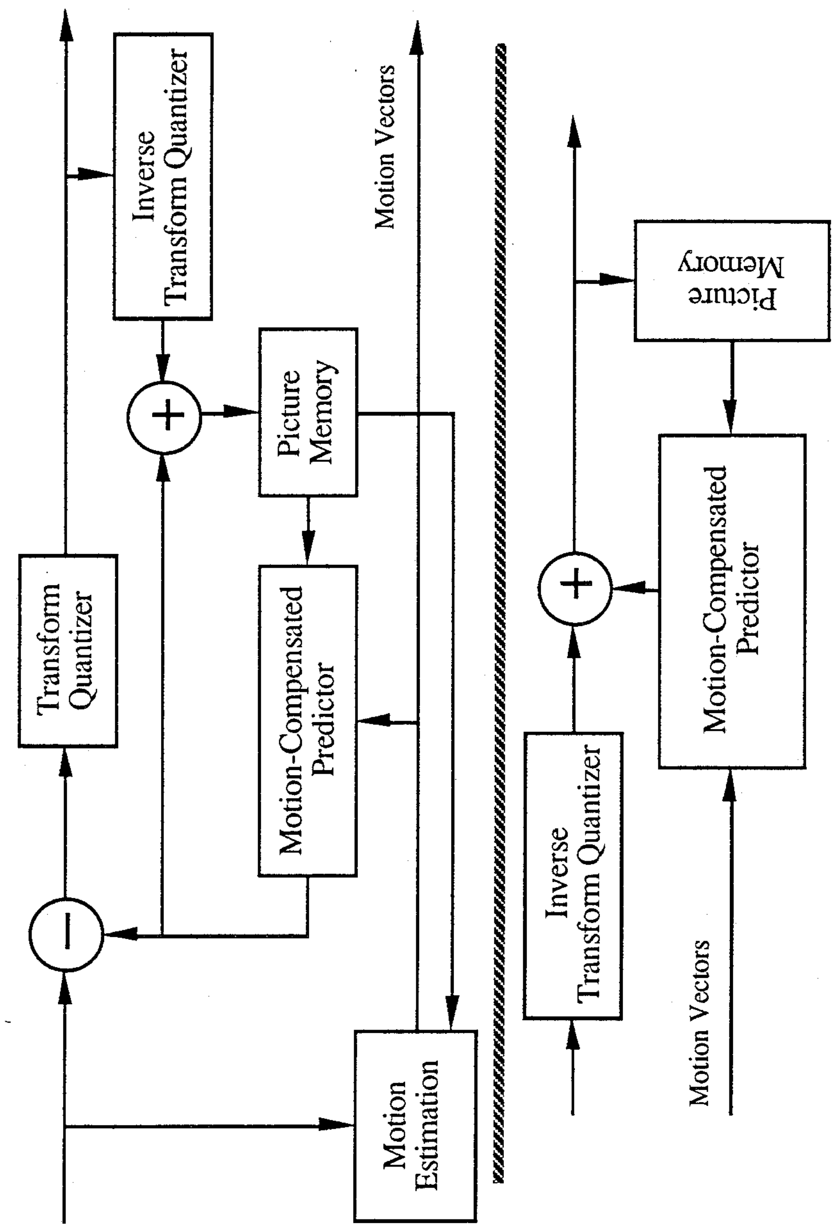


Figure A.2: A Generic Motion-Compensated Predictive Video Codec.

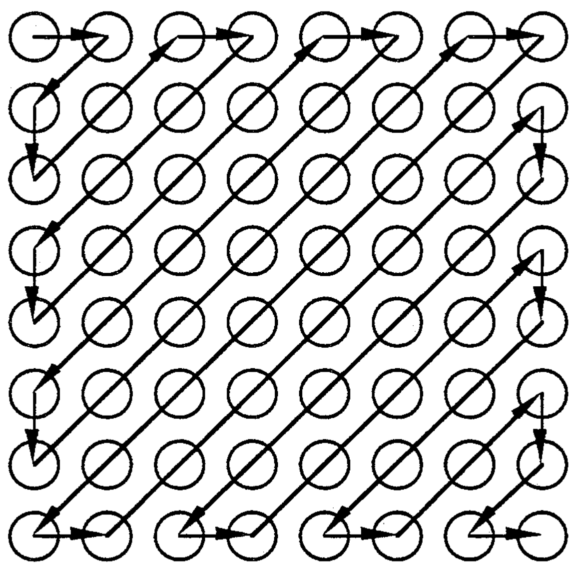


Figure A.3: Zigzag Scanning for an 8 by 8 Block of DCT Coefficients.

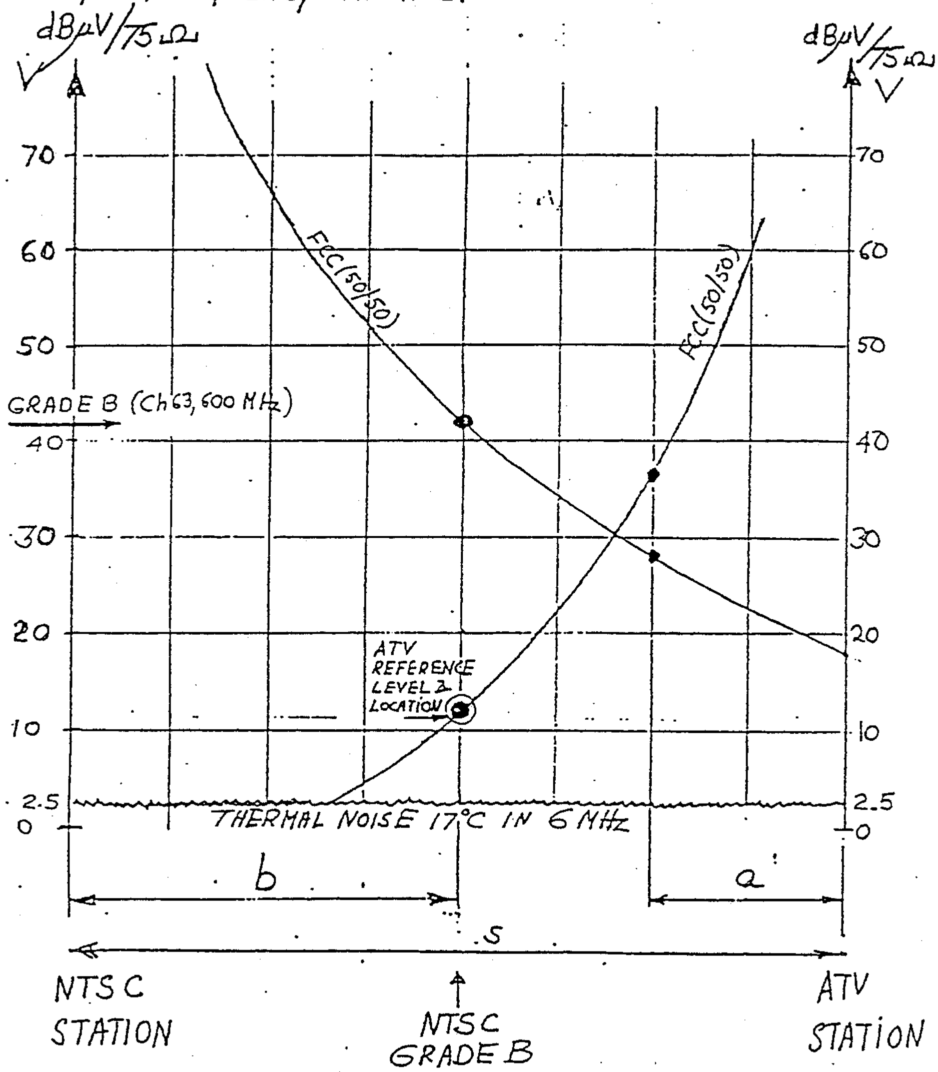
$V = E - D \quad dB\mu V/75\Omega$

V = VOLTAGE OVER 75 OHM LOAD DELIVERED BY MATCHED DIPOLE IN $dB\mu V/75\Omega$

E = FIELD STRENGHT IN $dB\mu V/meter$

$D = 20 \log(f_{MHz}/48.5)$ = DIPOLE FACTOR dB.

f = FREQUENCY IN MHz.



SIGNAL DELIVERED TO 75Ω BY DIPOLE RECEIVING FCC(50/50) FIELD AT UHF. GRADE B IS AT E=47, 56, AND 64 $dB\mu V/m$ FOR LO-VHF, HI-VHF AND UHF RESPECTIVELY.

Figure B.1: Chart for Qualitative Analysis of ATV Coverage.

APPENDIX A: PREDICTIVE VIDEO COMPRESSION

A.1 PREDICTIVE CODING

In classical video compression systems based on predictive coding, the encoder generates a prediction of the current picture. Figure A.1 shows a generic predictive coder configuration. The difference between the prediction and the actual current pictures is called the prediction error, or as used in this document, the prediction residue. The prediction residue values tend to have a peaked distribution. Such a distribution lends itself to more efficient coding using statistical coding techniques such as Huffman coding. However, the number of possible levels for the prediction residue is twice that of the original picture, i.e., the residue has twice as much dynamic range as the video signal. In most video communication systems, the prediction is quantized to fewer levels than the original pictures before further processing by a statistical coder. The coded data for the prediction residue, often of variable rate due to the variable-length codes of the statistical coder, is transmitted to the receiver. Additional side information may also be transmitted, as will be discussed shortly.

The reconstruction of the picture from the coded residue information involves the inverse operations of an appropriate statistical decoder (e.g., a Huffman decoder using the correct code table,) and the generation of the identical prediction as that used by the encoder in the generation of the residue. To allow the receiver to generate the identical prediction, the encoder tracks the operations of the decoder in such a way that at all times, the encoder has access to information (pictures, side information, etc.,) that the decoder has. The encoder then generates the prediction using *only* information that the decoder has, or will have. Any additional information needed to generate the prediction must be transmitted to the receiver as side information. A common example of side information is the motion vector information that is used in motion-compensated prediction. Other side information includes quantization parameters to allow the inverse operations at the receiver. For the remainder of this appendix, only temporal predictive coders are considered, i.e., the prediction is made from pictures corresponding to frames other than the current frame. Temporal predictive coders are sometimes called inter-frame predictive coders.

A.2 MOTION-COMPENSATED PREDICTION

The MPEG algorithm uses the same underlying structure as shown in the generic predictive coder of Figure A.1. The predictor uses a motion-compensated prediction in which reconstructed pictures, displaced by an amount specified by a motion vector computed elsewhere, constitute the prediction. Quantization is carried out in the Discrete Cosine Transform (DCT) coefficient domain.

For all intent and purposes, the DCT is a linear operation, and the coefficients can be considered as the matrix transformation outputs of a block of input values. The input values come from an 8 by 8 block of pixels or prediction residue values. When the values to the DCT correspond to blocks of *video pixels*, the processing is equivalent to spatial or intra-frame coding.

To obtain the motion vectors required for a motion-compensated predictive coder, a process known as motion estimation is carried out. In practice, this is done separately from the motion compensation process. Figure A.2 shows an example of a motion-compensated predictive coder configuration. The quantizer in Figure A.2 is labelled Transform Quantizer, reflecting the MPEG approach.

A.3 MOTION ESTIMATION

Within the context of ADTV, motion estimation refers to the operation of computing two-dimensional vectors (i.e., a vector with two elements) that indicate some measured motion. These vectors are called motion vectors. For temporal prediction algorithms, motion vectors are used in motion compensation operations prior to the generation of a prediction. Two types of motion estimations are used in ADTV. When a motion estimation of a particular picture, say the current frame, is made based on past pictures (i.e., pictures that are displayed before the current picture,) the estimation is said to be in a *forward* direction. *Backward* motion estimation refers to a situation when the motion estimation of the current picture is based on future pictures projecting *backward* in time to the current picture.

For both types of motion estimation, the operations involve a search of a block of pixels over a predefined area called a *search region.* The size of the search region determines the maximum bit length for the motion vectors. In general, a search region of size M by N will result in motion vectors with length no greater than $\log_2(MxN)$ bits. Within the search region, a total of MxN locations need to be compared with the current picture block. The location in the search region that produces the minimum *distortion* from the current picture block is the "best match" location for the current picture block. The spatial vector describing that minimum distortion location with respect to the current picture block spatial location is the motion vector for that motion estimation process. Common distortion measures are mean square block error or mean absolute block error. The mean absolute error criterion is often preferred because of its simpler hardware realization.

In ADTV, half-pixel motion estimation is carried out in two steps. An initial full search motion estimation is performed to obtain the integer-pixel motion vector. The second stage of the motion estimation process is a half-pixel motion estimation centered at the integer-pixel motion vector from the first stage search.

A.4 DISCRETE COSINE TRANSFORM (DCT) QUANTIZATION

As mentioned, the Discrete Cosine Transform is a linear transformation that, in simple terms, can be thought of as a matrix transformation. Typically, a block of values, say of size 8 by 8 and denoted as $f(x,y)$, is transformed by the processor (or DCT chip) to a block (of size 8 by 8) of DCT coefficients $F(u,v)$.

The equation for the forward transformation of an 8 by 8 block is

$$F(u,v) = \frac{1}{4} C(u)\, C(v) \sum_{x=0}^{7} \sum_{y=0}^{7} f(x,y) \cos\left[\frac{(2x+1)u\pi}{16}\right] \cos\left[\frac{(2y+1)v\pi}{16}\right]$$

and the inverse DCT (IDCT) is defined by the equation

$$f(x,y) = \frac{1}{4} \sum_{u=0}^{7} \sum_{v=0}^{7} C(u)\, C(v)\, F(u,v) \cos\left[\frac{(2x+1)u\pi}{16}\right] \cos\left[\frac{(2y+1)v\pi}{16}\right]$$

where

$f(x,y)$ = spatial domain values;

x,y = spatial coordinates in the pixel domain;

$F(u,v)$ = frequency/transform domain values;

u,v = frequency coordinates in the transform domain; and

$$C(\Omega) = \begin{cases} 1/\sqrt{2} & \text{for } \Omega = 0; \\ 1 & \text{otherwise.} \end{cases}$$

The inverse Discrete Cosine Transformation (IDCT) takes a block of coefficients $F(u,v)$ and produces the original 8 by 8 block of values $f(x,y)$. In the context of DCT quantization, an additional stage of coefficient quantization is applied after the forward DCT and before the inverse DCT operations.

Quite often, a weighting matrix is applied to the block of DCT coefficients prior to a uniform quantization. This allows perceptual weighting to be introduced into the DCT coefficient domain. To facilitate the quantization process, a zigzag scan is used to index the coefficients. The coefficients are accessed, for the purpose of quantization, in the order as shown in Figure A.3. The adoption of such a scan is based on the observation that the scan is in the direction of DCT coefficients corresponding to increasing spatial frequency.

A.5 VARIABLE-LENGTH CODING

The statistical coding mentioned before falls into a category of coding known as *lossless coding*. In general, lossless coding and decoding will introduce *no* coding artifacts to the coded data in the absence of any bit errors. Statistical coding, also known as entropy coding, exploits the statistical redundancy present in a data set. It assigns codewords to each input value according to the computed or estimated likelihood of occurrence of that input value. The more likely values are assigned shorter codewords relative to the less likely values. In this way, the *average* bits needed to code all the values in a data set is reduced. The codewords, although different in length, are uniquely decodable by the receiver if the codewords satisfy a condition known as the *prefix condition.* A well-known prefix code is the Huffman codes.

To be truly useful, statistical coding requires a good statistical characterization of the data set to be coded. In many practical systems, the statistical characterization is by means of multiple sets of statistical distribution of the values to be coded. For each distribution, or histogram, a code table is generated. Such a multiple-table approach is used in the MPEG algorithm.

In many cases in which the data set to be coded has a particular pattern of values frequently occurring, more compression can be obtained by assigning a codeword for that particular pattern. In both the MPEG algorithm and the CCITT H.261 standard for video compression, similar statistical "pattern coding" approaches have been adopted. Another commonly used statistical pattern coding is run-length coding. For the coding of heavily quantized numbers, such as the quantized DCT coefficients, it is highly likely that a series of zero values will be found in the strings of numbers to be coded. Instead of sending the codeword for the value "zero" corresponding to each individual zero value, a count of the number of consecutive zeroes, i.e., the zero-run, is coded. In particular, the zero-runs form a separate data set, along with their own statistical characterization by means of zero-run histogram and their own code table.

Huffman coding, in general, leads to codewords of different lengths so that the coded bit rate is no longer constant.

APPENDIX B: PREDICTIONS OF COVERAGE

Coverage of Advanced Television (ATV) is primarily limited by co-channel interference. The radiated ADTV signal is designed with particular considerations to over-the-air broadcasting in the TV broadcast band. The spectrum of the radiated ADTV signal is shaped to minimize co-channel interference into NTSC receivers and to maximize robustness of ADTV reception in the presence of co-channel NTSC radiation.

While NTSC-ATV co-channel interference is the major constraint on NTSC-compatible ATV coverage and allocations, there are other constraints: adjacent and taboo channel constraints, out-of-band radiation constraints, ATV-ATV co-channel interference as well as practical and economical constraints on the effective radiated power from ATV transmitters.

It must be strongly emphasized that all attempts to estimate ATV coverage and NTSC compatibility involve a large number of system-independent assumptions, particularly assumptions of "planning factors" at the ATV and NTSC receiver sites. For this reason we want to alert the reader that statements of coverage without detailed specifications of planning factors and other assumptions or experimental conditions are meaningless. As of this time no general agreement has been reached about system independent planning factors, assumptions and experimental conditions.

There are two kinds of estimates of coverage.

1. The first kind is the estimates of coverage from the transmitter point of view ignoring co-channel interference and other constraints. This is strictly an estimate of average and peak effective radiated power to provide an ATV service in an area which is about the same as an NTSC signal could cover with an acceptable quality of reception.

2. The second kind is the estimates of coverage from the co-channel interference point of view, starting from what can be considered as an acceptable level of co-channel interference determined primarily by subjective tests.

For an estimate of the required transmitter power (*first kind of coverage estimate*) we shall use as reference an NTSC TV transmitter and a TV receiver with an antenna installation and a tuner with exactly the standard NTSC planning factors for grade B reception receiving an electromagnetic field at the location of the receiving antenna which has exactly the grade B field strength according to FCC (50/50) propagation charts. For example, at UHF it means that the field exceeds 64 dBμV/meter 50% of the time and that this location is a typical one on this grade B contour where 50% of the locations receive a field strength exceeding 64 dBμV/meter 50% of the time. All the random noise in the tuner and the field is represented by an equivalent noise source in the antenna. With FCC planning factors the available equivalent noise power in a 6-MHz band is then 28.5 dB below the available TV-signal power at the peak of the sync pulses. (With the "old" assumption of a 4.2-MHz noise bandwidth CNR = 30 dB rather than 28.5 dB which is the TASO definition assuming a 6-MHz noise bandwidth.) The unweighted video signal-to-noise ratio is in this case 23.2 dB and the CCIR-weighted video signal-to-noise ratio is 29.8 dB. The picture is "snowy" but watchable.

Assuming these planning factors and this field strength, it is straightforward to determine the transmitter power required to deliver a 16-QAM system to this receiver site with a bit error rate of 10^{-3}. It is well known that for this bit error rate the average signal power to noise power ratio for 16-QAM is 18 dB. Thus, assuming a 6-MHz noise bandwidth, the average effective radiated power from the transmitter can be 10.5 dB below the power at the peak-of-sync of the NTSC signal delivering a grade-B picture at the receiver site. The peak power of the 16-QAM signal is theoretically 2.5 dB higher than the average power, but taking transients under consideration the peak to average ratio is about 4.5 dB. Thus the ATV transmitter peak power can be 6 dB lower than the peak power of the NTSC transmitter.

It should be emphasized that the FCC planning factors are conservative and that better antenna installations and tuners with lower noise figures can be anticipated for ADTV reception. Furthermore, considering co-channel constraints, full-power transmitters with the same coverage as current NTSC transmitters are probably going to be the exception rather than the rule.

More important than this estimate of required transmitter power assuming no co-channel constraints are estimates of coverage taking interference with NTSC co-channel signals under consideration. It is, of course, obvious that additional transmitters radiating in the present TV broadcast bands will cause interference of some level into reception of current NTSC signals. A key issue is how much interference in how many NTSC receivers in the field can be considered tolerable. This basic issue must be answered by subjective tests. It is our experience that the interference of digital signals in NTSC reception looks very much like random noise. Thus, in effect a digital ATV signal increases the noise figure (antenna temperature) of the exposed NTSC receiver. The question is: how many dB can the noise figure be increased by ATV in how many receivers? It is also our experience that spectrum shaping of the digital signal, as proposed in ADTV, can reduce its apparent disturbance relative to the apparent disturbance of random noise with the same power in a 6-MHz channel. Every decibel of such an improvement is of importance for ATV coverage, since the level of tolerable disturbance of ATV into NTSC reception is a starting point for a determination of ATV coverage in the direction of a co-channel NTSC station.

For a qualitative discussion of the issues involved in the prediction of coverage with consideration to co-channel interference the reader is referred to the chart shown in Figure B.1. It shows a line between an NTSC station (to the left) and an ATV station (to the right). It also shows the signal voltage in dBμV delivered to a 75-Ω resistor by a matched dipole 30 feet above ground oriented for maximum reception assuming that the field strength is FCC (50/50). The chart also shows that the thermal noise level is 2.5 dBμV. In particular the chart shows received signals at a distance "a" from the ATV receiver and at a distance "b" from the NTSC receiver, which is on the NTSC grade B contour. At the NTSC grade B contour, the CNR is determined with FCC planning factors (noise figure, field strength statistics, antenna gain, cable losses, etc.) to be 28.5 dB. Planning factors for co-channel interference, however, are not spelled out. They include assumptions of antenna gain in the direction of the ATV signal as well as assumptions about the statistics of ATV field strength fluctuations. Usually it is assumed that co-channel interference must not exceed a certain level more than 10% of the time (FCC (50/10)). With consideration of these additional planning factors for co-channel interference, the most important reference level for the ATV coverage is determined by how much the ATV signal can be allowed to interfere with NTSC reception on the NTSC grade B contour. This reference point is shown in the chart. As we move from the NTSC grade B contour towards the ATV station, ATV reception conditions improve: CNR improves and the desired to undesired signal ratio improves even more rapidly because the desired signal strength increases while the undesired signal strength decreases. At a location a distance "a" from the ATV transmitter ATV reception becomes acceptable. Of key importance for the ATV coverage distance "a" is robustness to NTSC co-channel interference. *In the ADTV system NTSC interference will rarely limit the ADTV coverage.* The exception occurs when the ATV station is located very close to the NTSC grade B contour. Like any ATV system, the signal level of the ADTV signal received by the NTSC receiver must be very weak, only a few decibels above the equivalent input noise level of the receiver/antenna pattern, noise figure and propagation statistics taken into consideration. No ATV signal can be received on the co-channel NTSC grade B contour. The real issue is how close to the co-channel NTSC grade B contour can ATV be received with acceptable error rates. Due to the ADTV robustness towards co-channel interference, ADTV coverage can extend closer to the grade B contour than systems with less tolerance to co-channel interference. The limitation is the ADTV carrier to noise ratio, which depends on antenna pattern, noise figure, and carrier level statistics.