**Video**

**The horizontal blanking interval**: is the separation between consecutive lines.

**The vertical blanking interval:** is the separation between consecutive frames.

Horizontal blanking actually occurs slightly before the beginning of each line of video information. Vertical blanking occurs after each frame.

**Persistence of Vision:**

The threshold of retention is 1/ 30 to // 32 of a second. If the images change on the retina at a rate slower than that, the eye sees the light and then the dark that follows. If the images change at a faster rate, the eye sees the images as continuous motion and not as individual images.

**Fields:**

The 30-frames-per-second frame rate of video could potentially allow the flicker of the changing frames to be noticeable. Therefore, the frame rate needed to be increased. The simplest way to do this would be to double the frame rate to 60 frames per second. This was a technological challenge in the early days of tube electronics. The engineers of the time solved the problem by breaking the 30 frames they were able to capture into 60 half frames called fields, each field containing alternating scan lines of the frame.

**Interlace Scanning:**

The process of this field-by-field scanning is known as interlace scanning because the lines in each field interlace with the alternate lines of the other field. There are two fields for each frame. Because the images are appearing at the rate of 1/ 60 of a second, the eye does not see the interval between the two fields. Therefore, the eye perceives continuous motion.

A television image is created through Interlace scanning. Interlace scanning is the process of scanning every other line from top to bottom. The beam first scans the odd lines top to bottom, and then it scans the even lines top to bottom. Each scan from top to bottom is a field. It is the combination of the two successive fields that make up an entire frame of a video image.

**Color Video:**

Color, like sound, is based on frequencies. Each color has its own specific frequency in the visible spectrum and its own specific wavelength. Each color is defined scientifically by its frequency or its wavelength. Wavelength is related to frequency in that the higher the frequency, the shorter the physical length of the wave. Color frequencies are extremely high in the spectrum, and therefore are easier to notate by wavelength rather than frequency (cycles per second or hertz). Wavelengths of light are measured in nanometers, or billionths of a meter.

**Primary and Secondary Colors:**

The definition of a primary color is that it cannot be created through a combination of any of the other primary colors. For example, in a red, green, and blue (RGB) color system. When defining a color system, any set of colors can be used as the primaries. A primary color system can include more than three primary colors. Again, the only rule that must be adhered to is that the combination of any two of the primary colors cannot create one of the other primaries. The basic color system in television is a three-primary color additive system. The primary colors in television are red, green, and blue (RGB). Combining two primary colors creates a secondary color. Combining secondary colors creates a tertiary color. In the color television system, there are three secondary colors— yellow, cyan, and magenta— which are each combination of two primary colors. Yellow is the combination of red and green, cyan is the combination of green and blue, and magenta is a combination of red and blue

**Chrominance and Luminance:**

Color can be defined by three measurements: (1) chrominance, the amount of color information; (2) luminance, the amount of white light that is mixed with the color information; and (3) hue, the particular color pigment. The combination of color and light, or chrominance and luminance, produces saturation. Saturation is the ratio between how much color and how much light there is in the signal (Figure 6.4, Plate 5). If more white light is added to a color, it becomes desaturated. If the white light is removed, the color becomes more saturated. Consider the difference between red and pink. Pink is the same hue as red, except it has more white light in it, which desaturates it. The method used to add more color in an additive system is to add proportionate amounts of the other two primaries, as the combination of all three is white.

**Color Difference Signals:**

The color information minus the luminance information is known as the color difference signal. Mathematically, this would be shown as Red-Y (Red minus luminance), Green-Y, and Blue-Y, or R-Y, G-Y, and B-Y. Based on the Pythagorean theorem, only two signals are needed to calculate the color information. R-Y and B-Y were chosen because they contain the least amount of luminance information and therefore conserve bandwidth. The designations R-Y and B-Y were later changed to Pr and Pb for analog signals and Cr and Cb for digital signals. Thus, Y, R-Y, B-Y would be designated either as Y, Pb, Pr or as Y, Cb, Cr. As analog video is replaced by digital, the Pb and Pr designations are no longer in common use and are being replaced by the Cb, Cr designations.

**Video Encoding:**

Video information is constrained by available transmission bandwidth and the limitations of recording media. To overcome this limitation, the video information must first be condensed or compressed so it will fit within the available space. This is done through a process called encoding, which involves taking all the parts of the video and audio information and combining or eliminating the redundant material mathematically.

**Digital Component Signal:**

This is the most common method of video signal encoding. The majority of cameras and image creation devices process in digital component. Digital component video takes the elements that comprise a video signal (YCbCr) and keeps them separate in recording and transmission. This is similar to analog component in the way these elements are treated. The inputs and outputs from digital cameras and recording devices can either be RGB or the YCbCr components. Equipment that is capable of creating or reproducing these digital elements will have separate paths for each of these elements.

**Image Resolution:**

Image Resolution Image resolution is the detail or quality of the video image as it is created or displayed in a camera, video monitor, or other display source. The amount of detail is controlled by the number of pixels (picture elements) in a horizontal scan line multiplied by the Television Standards number of scan lines in a frame. The combined pixel and line count in an image represents what is known as the spatial density resolution, or how many total pixels make up one frame. For standard definition ATSC digital video, the image resolution is 720 pixels per line with 480 active scan lines per frame (720 × 480). Increasing the number of pixels in an image increases the amount of detail. This corresponds to an increase in the resolution of the image. If the displayed image is kept at the same size, an increase in resolution would increase the detail in the image. Alternatively, a higher-resolution image could be displayed much larger while keeping the same degree of detail as the lower-resolution image.

**Aspect Ratios:**

Video images are generally displayed in a rectangular shape. To describe the particular shape of an image, the width of the image is

compared to its height to come up with its aspect ratio. This ratio describes the shape of the image independent of its size or resolution. Two common aspect ratios in video are 16 × 9 and 4 × 3. An image with a 16 × 9 aspect ratio would be, for example, 16 units across and 9 units tall. The actual ratio does not depend on any particular unit of measure. If a 16 × 9 image or display were 16 feet wide, it would be 9 feet tall. If it were 16 inches wide, it would be 9 inches tall. A display that is 9 yards tall would be 16 yards wide. Think of a stadium scoreboard at that size. Older standard definition displays were a bit closer to square. A display that is 4 feet wide by 3 feet tall would be referred to as a 4 × 3 ratio. Another common way to note aspect ratio is to divide the length of the image by the height. In the 16 × 9 example, when the width amount of 16 is divided by the height of 9, the result would be 1.777777 units long. This is often rounded to 1.78 and expressed as 1.78 to 1, or shown as the ratio 1.78: 1. For each unit of height we need 1.78 units of width. The older 4 × 3 standard definition image would be 1.33: 1. When four is divided by three the result is 1.33333333. The film world uses a number of different image ratios, but they are almost always expressed as comparisons to 1 unit of height. For example, a theatrical release movie, showing in Cinemascope, has an aspect ratio of 21 × 9, but it is more commonly referred to as 2.35: 1. Another common film aspect ratio is 1.85: 1, which is closer to the 16 × 9 aspect ratio of HD (Figure 12.2).

**Pixel Aspect Ratio:**

The pixel aspect ratio is the size and shape of the pixel itself. In computer displays, pixels are square with an aspect ratio of 1 to 1. NTSC pixels have an aspect ratio of 0.91: 1, which makes them tall and thin. When setting the standard for digitizing the NTSC video image, the intent was to digitize the image at the highest practical resolution. While the number of pixels in a horizontal scan line could be set to any amount, the number of scan lines could not be arbitrarily increased since they are part of the NTSC standard. Therefore, the pixels were changed to a narrow, vertical rectangle shape, allowing an increase in the number of pixels per line and added image resolution. In the digital world, the aspect ratio of the pixels is sometimes changed to take fewer samples on each line.

**Interlace and Progressive Scan Modes:**

Interlace scanning is the process of splitting a frame of video into two fields. One field of video contains the odd lines of information while the other contains the even lines of the scanned image. When played back, the two fields are interlaced together to form one complete frame of video.

When listing the criteria for a particular standard, the indication of whether the scanning mode is interlaced or progressive appears as an “i” or “p” following the line count, such as 480p, 720p, 1080i, and so on.

**Frame Rate**:

The frame rate, regardless of the pixel or line count, is the number of full frames scanned per second. This rate represents what is known as the temporal resolution of the image, or how fast or slow the image is scanned. Frame rates vary depending on the rate at which the image was captured or created, the needs or capability of the system using the video image, and the capability of the system reproducing the image. For example, an image may have been captured at 24 frames per second (fps), edited at 29.97 fps, and reproduced at 30 fps.

**HD Image Resolution:**

With the 16 × 9 aspect ratio, there is a larger image area and therefore more room for additional pixels. For example, one high def format has an image resolution of 1920 × 1080. In this format, there are 1920 pixels across one line and 1080 lines in one frame of the image. The 1920 pixel count is the horizontal resolution, and the 1080 line count is the vertical resolution.

**4K Image Resolution:**

4K As of this writing, the highest resolution equipment available is called 4K, as the image has approximately 4000 pixels in each scan line. The actual number for a true 4K image is 4096 pixels per line, with 2160 lines in each frame. This size is used for Digital Cinema photography and projection. The aspect ratio for this format is 1.9: 1, just a little wider than the common 1.85: 1 used by many feature films.

**Progressively Segmented Frames:**

HDTV standards can use either of the two scanning modes, interlace or progressive. When a high def standard uses an interlace mode, the odd fields are transmitted first, followed by the even fields, just as they are in CDTV or SDTV. In progressive scanning, the entire image is scanned as one complete frame. This data may then be transmitted as a complete frame and received as a complete frame. If there is insufficient bandwidth to transmit the complete frame, the data may be segmented and transmitted in two parts. Because progressively scanned images are complete frames of video, they require more bandwidth to transmit than may be available. To transmit these images within an interlace environment using narrower bandwidths, a process of segmenting progressively scanned frames was developed. To do this, the image is divided into two fields by scanning every other line as in interlace scanning. The difference is that, in the interlace scanning process, the two fields are from different instances in time. When a progressively scanned image is segmented, the two fields are from the same instance in time. When the separate fields are recombined, the result is a complete progressively scanned frame.

**Lossy Compression:**

Compression used for active video is usually in the lossy category. With lossy compression, the restored image will be an approximation of the original. When a lossy image is reproduced or uncompressed, not all the data left out during compression will be restored exactly as it was. To minimize the visible loss of data, lossy compression techniques generally compress the data that comprise those parts of the image the human eye is less sensitive to, or that contain less critical image data. The human eye is more sensitive to changes in light levels or luminance than it is to changes in color, both hue and saturation. Within the color gamut, the human eye is more sensitive to the yellow-green-blue range. The human eye is also more sensitive to objects in motion than to still objects. In lossy compression, the data compressed is the data that does not fall within the human sensitivity range or data that contains a great deal of motion. Two commonly used lossy compression techniques are JPEG and MPEG.

**Data Reduction**:

Video files generally contain redundancy that can be used to reduce the amount of data to be stored. This can be done by registering the differences within a frame (intraframe) and between frames (interframe) rather than storing all the information from each frame.

**Intraframe Compression**

Intraframe compression utilizes spatial redundancy. This compression is achieved through a technique called sub-sampling. In sub-sampling, the number of bits needed to describe an image is reduced by storing only some of the pixels that make up the image. For instance, every second pixel in a row and the entirety of every second row could be ignored. Those pixels that are retained would then be increased in size to compensate for the data that has been left out

Another sub-sampling strategy uses the average values for a group of pixels. This average is substituted for the original values for those pixels. Sub-sampling effectively reduces the number of pixels in an image. Alternatively, rather than reduce the number of pixels in an image, the amount of information about each pixel can be reduced. This, however, also reduces the number of gradations of color and grays in the image.

Intraframe Compression Quantization is another method for reducing the amount of data. Quantizing sets a limit on the quantity for the range of values to be stored and the increments between values for the data to be coded, thereby compressing the amount of data needed to be stored. Transform coding uses a complex conversion process to turn coded blocks of the image into a compressed image or video file for storage or transmission.

**Interframe Compression**

Interframe compression compares the unchanging portions of successive frames from one frame to the other.

Two known ways are : block-based difference coding, and block-based motion compensation.

**In basic difference coding**, each frame is compared to the preceding one and only those pixels that are changed are stored.

**Interframe Compression Block-based difference coding** works in the same fashion but at a block level rather than a pixel level. The frames are divided into blocks of pixels, and it is these blocks that are compared between frames.

**Block-based motion compensation** is a further refinement of difference coding. The frame is divided into blocks as before. These blocks are then compared to blocks in the preceding frame to find blocks that are similar. If the similar blocks are in different locations, it is this difference of position that is stored rather than the actual information in the block. A further refinement is bidirectional motion compensation, in which the current frame is compared to both the preceding and following frame and the differences stored rather than the content of the frame

**Bit Rates:**

The complexity of an image is a result of the combination of the amount of movement or change from frame to frame and the quantity of detail contained in the image. To maintain proper image motion in time, the zeros and ones, or digital bits, that comprise the data must be transmitted and received quickly enough to reproduce the image in the proper time frame. Depending on the complexity of the image and the required level of quality, different data rates, or bit rates— that is, the speed at which the data is processed— are used. If the images are less complex in nature, or if the required level of quality is not high, a fixed or constant data rate or bit rate may be used. Where the images are either more complex or the required level of quality is high, variable bit rates may be used in order to maintain a reduced data rate while not compromising the quality. Because of these differences, constant bit rates can be used to compress images in real time, whereas variable bit rates cannot.

**Constant Bit Rates Fixed or constant bit rates (CBR):**

result in varying levels of picture quality because there is no allowance for image complexity. Broadcast media— such as cable, satellite, and terrestrial broadcasting— require constant bit rates for their transmission equipment. Live broadcasts, satellite linkups, and playback of uncompressed video all require immediate real-time compression while being transmitted.

MPEG Compression:

The MPEG Process The MPEG process starts by analyzing a sequence of video frames known as the video stream. Redundant information is encoded and compressed. The compressed video stream is then encoded into a bit stream. The bit stream is then stored or transmitted at the bit rate called for by the playback equipment. The data is decoded and uncompressed when it is to be used and the image restored to its original form. MPEG compression utilizes a combination of two different compression schemes, spatial and temporal. Spatial compression reduces the quantity of data contained in each frame of video by eliminating the redundant data within the image. Temporal compression compares the changes between the images over time and stores the data that represents only the changes. Spatial compression uses the same technique as JPEG compression, to create an intra picture, called an I frame. Unlike the temporal compression frames, the I frames are complete “stand-alone” images that can be decoded and displayed without reference to any surrounding frames. The I frames are interspersed within the video stream and act as references for the temporal compression between frames. The arrangement is somewhat like a picket fence, with the I frames representing the relatively few fence posts while the temporal frames are the many pickets. The temporal compression frames, called B and P frames, contain motion information that describes how the different regions of the I frame have changed between the intervening frames. The B and P frames contain far less data than the I frames. They contain only the data about the changes that have occurred between frames. This accounts for the great efficiency of MPEG encoding. Compression rates of 25: 1 can be achieved with little or no noticeable degradation in the quality of the uncompressed image. I frames and B and P frames are described in more detail below.

**I Frames:**

An I frame (intra picture) is one frame that is a complete image sampled in detail so it can be used as a reference for the frames around it. Each I frame is divided into 8 × 8 pixel blocks which are then placed in groups of 16 × 16 pixel blocks called macroblocks (Figure 15.1 (Plate 20)). These macroblocks are then compressed using a variety of compression techniques. I frames are created as often as needed and particularly when there is a substantial change in the image content. In a typical video stream, this occurs approximately two times per second.

**P Frames:**

The frames before and after the I frame, labeled P and B, contain the data representing the changes that occur between themselves and the I frame. P frames (predicted pictures) contain descriptions of how the pixel blocks in the previous frame have changed to create the current frame. In addition, the P frames are examined to see if data blocks have moved. Subject or camera movement might cause some image blocks to have the same data in each frame, but in a different location. These descriptions of distance and direction of movement are called motion vectors. The decoding process for the current frame looks backward at the previous frame and repositions the pixels based on the P frame motion vectors. The previous frame could be either an I frame or another P frame. If there is a substantial change in the image, new pixel blocks are created for the portion of the image that has changed. These new blocks are derived from the source video and use the same encoding method as the I frame. P frames cannot stand alone or be directly accessed, since they are dependent upon the information in the previous frames from which they are derived. P frames contain much less data than I frames and are therefore simpler to encode.

**B Frames:**

B frames (bidirectional pictures) are similar to P frames in that they are made up of motion vectors and picture blocks. The difference is that they look both forward and backward to compare pixel blocks, where the P frames only look backward to the previous frame. When new elements enter the picture, the pixels in a B frame can be compared forward or backward to pixel blocks in either I or P frames. The data representing the difference between the previous and following frames is used to create the B frame. Using B frames causes delays in the transmission of the bit stream. As the encoder must wait to see what is contained in future frames, the data is buffered until that data is available. Transmission of the data cannot occur until the B frames are able to be calculated. Because of this potential delay in playback, B frames are not used in all forms of MPEG.

**The Group of Pictures (GOP):**

I, B, and P frames are grouped together to form a Group of Pictures, or GOP. A GOP must start and end with an I frame to permit its use as a reference for the surrounding B and P frames. A GOP can contain just P frames or both B and P frames in between the I frames. The number of B frames or P frames within a GOP can be increased or decreased depending on image content or the application for which the compressed video is intended. For example, a fast-moving action sequence with complex content (lots of detail) would use shorter groups, hence more I frames. Group lengths typically range from 8 to 24 frames. Figure 15.2 shows the typical GOP structure for 30 frames of IBP-encoded video stream. Group of Pictures (GOP)

**IP Method**

Excellent compression quality can be achieved using just I and P frames, even though the P frames only use backward references in time.

**IBP Method :**

The addition of the optional B frame increases the compression quality but lowers the rate of data transmission. The B frame looks both forward and backward in time so the frame with the most helpful information can be used as a reference

**Color Subsampling:**

Another aspect of the compression process is color subsampling. As mentioned above, 4: 2: 0 and 4: 2: 2 are numbers that represent digital sampling standards. First, the RGB channels are converted into luminance (Y) and two chrominance channels (Cr and Cb). After converting RGB video signals into luminance and chrominance data, the chrominance portion can be compressed with little apparent loss in image quality.

Let’s consider four different digitizing schemes for video: 4: 4: 4, 4: 2: 2, 4: 1: 1 and 4: 2: 0. In each case, for every set of four pixels on a scan line, there are four digital samples of the luminance channel. The schemes differ, however, in the amount of times the chroma in each line is sampled. In the 4: 2: 2 scheme, for example, only two samples are taken for each of the two chrominance channels (Figure 15.5 (Plate 21)). Each chrominance sample is shared by two pixels. As a result, the two chrominance channels are digitized at half the resolution of the luminance channel, reducing the amount of data for those two channels by 50%. This reduces the total data required for each frame of video by 33%. 4: 2: 0 sampling takes the idea of sub-sampling the chrominance channels a step further than 4: 2: 2. If, for every four pixels, there were four samples of luminance but no samples of chrominance, that would be 4: 0: 0 sampling. Of course, with no chrominance data you would only have a black and white picture. However, if every other scan line was digitized at 4: 2: 2 and the lines in-between were digitized at 4: 0: 0, the chrominance data from the 4: 2: 2 scan lines could be shared by the alternating 4: 0: 0 scan lines, further reducing the amount of data required to describe a frame of video.

**Glossary**

**Chroma Subsampling:** The sampling of the color or chroma information in an image. Chroma is sampled less often than the luminance to reduce the amount of data to be stored or transmitted.

**4: 2: 0** -> A chroma subsampling scheme. For every four samples of luminance taken, two are taken for each color difference signal, but only on every other scan line. (See 4: 2: 2.)

**4: 2: 2** -> A chroma subsampling scheme. For every four samples of luminance taken, two are taken for each of the color difference signals on each scan line.

**4: 4: 4** -> A chroma subsampling scheme. For every four samples of luminance taken, four are taken for each of the color difference signals on each scan line.

**Constant or Variable Bit Rate** Two different ways to control the flow of bits in a compressed signal. Constant bit rates can be used to compress images in real time, whereas variable bit rates cannot.

**Combing** When interlaced material is presented on progressive displays, it can lead to a distracting artifact called combing. This leaves fine lines extending from areas of fast motion similar to the teeth of a comb.

**Image Resolution** The amount of detail contained in a video image based on the number of lines in the image and the number of pixels per line.

**Motion Vectors** During the compression process, the descriptions of distance and direction of macroblock movement within I, P or B frames are called motion vectors.

**PsF (Progressive Segmented Frame**) A progressively scanned image that has been divided into two fields, each containing alternate lines from the frame. Unlike a true interlaced frame, both fields are from the same point in time.