

Chapter 5

Fundamental Concepts in Video

- ✓ Video is a series of images. When this series of images are displayed on screen at fast speed (e.g 30 images per second), we see a perceived motion. It projects single images at a fast rate producing the illusion of continuous motion. These single images are called frames. The rate at which the frames are projected is generally between 24 and 30 frames per second (fps). The rate at which these images are presented is referred to as the Frame Rate
- ✓ This is fundamental to the way video is modeled in computers.
- ✓ A single image is called frame and video is a series of frames.
- ✓ An image just like conventional images is modeled as a matrix of pixels.
- ✓ To model smooth motion psychophysical studies have shown that a rate of 30 frames a second is good enough to simulate smooth motion.
 - Old Charlie Chaplin movies were taken at 12 frames a second and are visibly jerky in nature.

Each screen-full of video is made up of thousands of pixels. A pixel is the smallest unit of an image. A pixel can display only one color at a time. Your television has 720 vertical lines of pixels (from left to right) and 486 rows of pixels (top to bottom). A total of 349,920 pixels (720 x 486) for a single frame.

There are two types of video:

- ✓ Analog Video
- ✓ Digital Video

Analog Video

Analog technology requires information representing images and sound to be in a real-time continuous-scale electric signal between sources and receivers. It is used throughout the television industry. For television, images and sound are converted into electric signals by transducers. Distortion of images and noise are common problems for analog video.

In an analogue video signal, each frame is represented by a fluctuating voltage signal. This is known as an analogue waveform. One of the earliest formats for this was composite video.

Analog formats are susceptible to loss due to transmission noise effects. Quality loss is also possible from one generation to another.

This type of loss is like photocopying, in which a copy of a copy is never as good as the original.

Digital Video

Digital technology is based on images represented in the form of bits. A digital video signal is actually a pattern of 1's and 0's that represent the video image. With a digital video signal, there is no variation in the original signal once it is captured on to computer disc. Therefore, the image does not lose any of its original sharpness and clarity. The image is an exact copy of the original. A computer is the most common form of digital technology.

The limitations of analog video led to the birth of digital video. Digital video is just a digital representation of the analogue video signal.

Unlike analogue video that degrades in quality from one generation to the next, digital video does not degrade. Each generation of digital video is identical to the parent.

Even though the data is digital, virtually all digital formats are still stored on sequential tapes. There are two significant advantages for using computers for digital video :

- ✓ the ability to randomly access the storage of video and
- ✓ compress the video stored.

Computer-based digital video is defined as a series of individual images and associated audio. These elements are stored in a format in which both elements (pixel and sound sample) are represented as a series of binary digits (bits).

Advantages:

- Direct random access -> good for nonlinear video editing
 - No problem for repeated recording
 - No need for blanking and sync pulse
- ✓ Almost all digital video uses component video

Analog vs. Digital Video

An analog video can be very similar to the original video copied, but it is not identical. Digital copies will always be identical and will not lose their sharpness and clarity over time. However, digital video has the limitation of the amount of RAM available, whereas this is not a factor with analog video. Digital technology allows for easy editing and enhancing of videos. Storage of the analog video tapes is much more cumbersome than digital video CDs. Clearly, with new technology continuously emerging, this debate will always be changing.

Displaying Video

There are two ways of displaying video on screen:

- ✓ Progressive scan
- ✓ Interlaced scan

Interlaced Scanning

Interlaced scanning writes every second line of the picture during a scan, and writes the other half during the next sweep. Doing that we only need 25/30 pictures per second. This idea of splitting up the image into two parts became known as interlacing and the splitted up pictures as fields. Graphically seen a field is basically a picture with every 2nd line black/white. Here is an image that shows interlacing so that you can better imagine what happens.

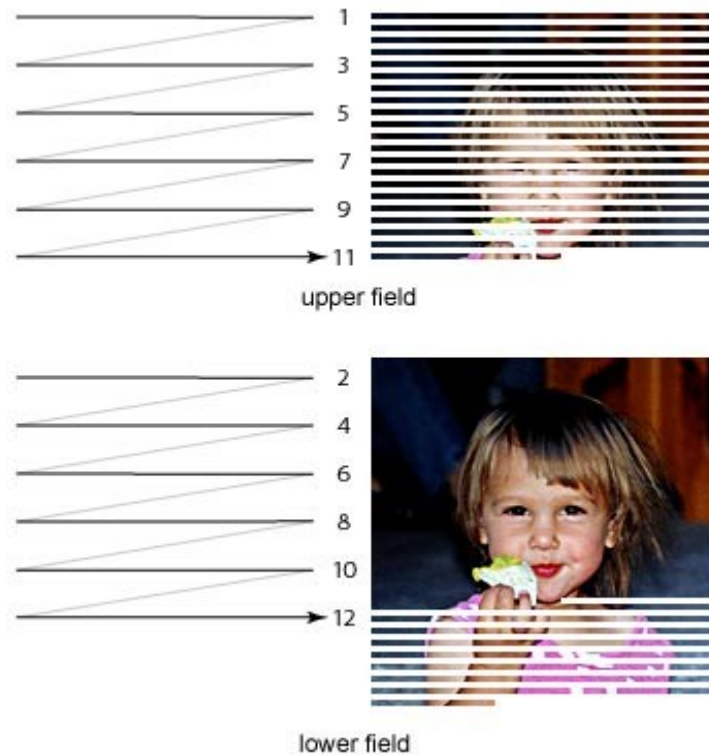
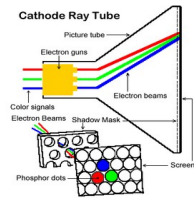


Fig
scanning

During the first scan the upper field is written on screen. The first, 3rd, 5th, etc. line is written and after writing each line the electron beam moves to the left again before writing the next line.

Currently the picture exhibits a "combing" effect, it looks like you're watching it through a comb. When people refer to interlacing artifacts or say that their picture is interlaced this is what they commonly refer to.

Once all the odd lines have been written the electron beam travels back to the upper left of the screen and starts writing the even lines. As it takes a while before the phosphor stops emitting light and as the human brain is too slow instead of seeing two fields what we see is a combination of both fields - in other words the original picture.



Progressive Scanning

PC CRT displays are fundamentally different from TV screens. Monitor writes a whole picture per scan. Progressive scan updates all the lines on the screen at the same time, 60 times every second. This is known as progressive scanning. Today all PC screens write a picture like this.

Fig progressive scanning

Here is a comparison of computer and television display.

Computer

- ✓ Scans 480 horizontal lines from top to bottom
- ✓ Scan each line progressively interlacing system
- ✓ Scan full frame at a rate of typically 66.67 HZ or higher
- ✓ Use RGB color model and restricted

Television

Scans 625, 525 horizontal

Scan line using

Scan 25-30 HZ for full

Uses limited color palette

luminance (lightness or

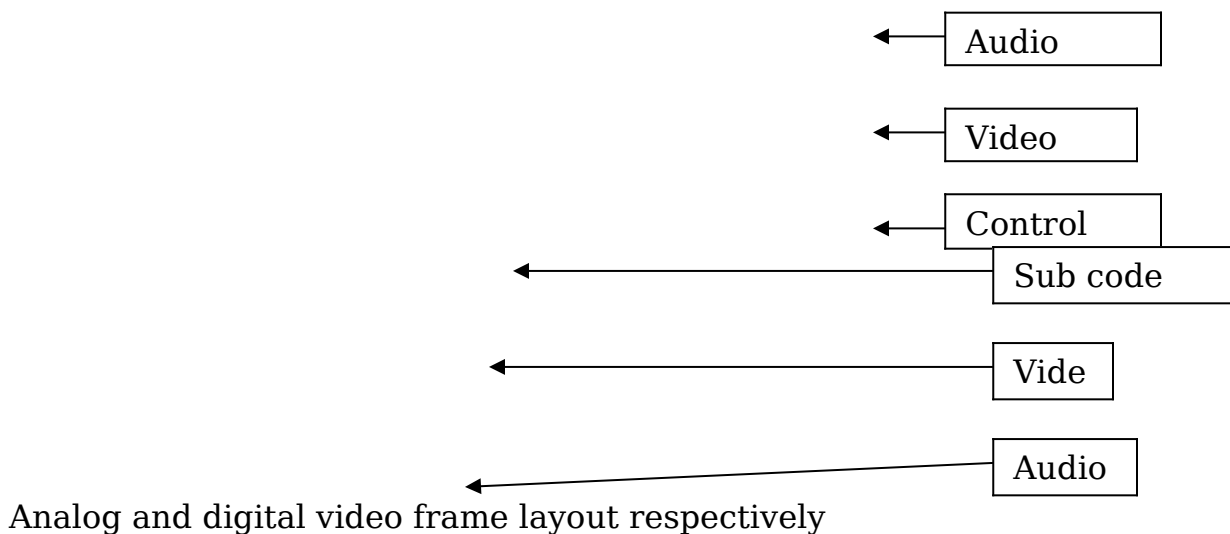
darkness)

Recording Video

CCDs (Charge Coupled Devices) a chip containing a series of tiny, light-sensitive photosites. It forms the heart of all electronic and digital cameras. CCDs can be thought of as film for electronic

cameras. CCDs consist of thousands or even millions of cells, each of which is light-sensitive and capable of producing varying amounts of charge in response to the amount of light they receive.

Digital camera uses lens which focuses the image onto a Charge Coupled Device (CCD), which then converts the image into electrical pulses. These pulses are then saved into memory. In short, Just as the film in a conventional camera records an image when light hits it, the CCD records the image electronically. The photosites convert light into electrons. The electrons pass through an analog-to-digital converter, which produces a file of encoded digital information in which bits represent the color and tonal values of a subject. The performance of a CCD is often measured by its output resolution, which in turn is a function of the number of photosites on the CCD's surface.



Types of Color Video Signals

Component video - each primary is sent as a separate video signal. The primaries can either be RGB or a luminance-chrominance transformation of them (e.g., YIQ, YUV).

- Best color reproduction
- Requires more bandwidth and good synchronization of the three components

Component video takes the different components of the video and breaks them into separate signals. Improvements to component video have led to many video formats, including S-Video, RGB etc.

Composite video – color (chrominance) and luminance signals are mixed into a single carrier wave. Some interference between the two signals is inevitable.

Composite analog video has all its components (brightness, color, synchronization information, etc.) combined into one signal. Due to the *compositing* (or combining) of the video components, the quality of composite video is marginal at best. The results are color bleeding, low clarity and high generational loss.

S-Video (Separated video) – a compromise between component analog video and the composite video. It uses two lines, one for luminance and another for composite chrominance signal.

Video Broadcasting Standards/ TV standards

There are three different video broadcasting standards: *PAL*, *NTSC*, and *SECAM*

PAL (Phase Alternate Line)

PAL uses 625 horizontal lines at a field rate of 50 fields per second (or 25 frames per second). Only 576 of these lines are used for picture information with the remaining 49 lines used for sync or holding additional information such as closed captioning. It is used in Australia, New Zealand, United Kingdom, and Europe.

- ✓ Scans 625 lines per frame, 25 frames per second (40 msec/frame)
- ✓ Interlaced, each frame is divided into 2 fields, 312.5 lines/field
- ✓ For color representation, PAL uses YUV (YCbCr) color model
- ✓ In PAL,
 - 5.5 MHz is allocated to Y,
 - 1.8 MHz each to U and V

SECAM (Sequential Color with Memory)

SECAM uses the same bandwidth as PAL but transmits the color information sequentially. It is used in France, East Europe, etc. SECAM (Système Electronique Pour Couleur Avec Memoire) is very similar to PAL. It specifies the same number of scan lines and frames per second. It is the broadcast standard for France, Russia, and parts of Africa and Eastern Europe.

NTSC (National Television Standards Committee)

NTSC is a black-and-white and color compatible 525-line system that scans a nominal 30 interlaced television picture frames per second. Used in USA, Canada, and Japan.

- 525 scan lines per frame, 30 frames per second (or be exact, 29.97 fps, 33.37 sec/frame)
- Interlaced, each frame is divided into 2 fields, 262.5 lines/field
- 20 lines reserved for control information at the beginning of each field (Fig. 38)
- So a maximum of 485 lines of visible data

NTSC Video

- ✓ 525 scan lines per frame, 30 frames per second
- ✓ Interlaced, each frame is divided into 2 fields i.e. 262.5 lines/field
- ✓ 20 lines reserved for control information at the beginning of each field
 - So a maximum of 485 lines of visible data

NTSC Video Scan Line

- ✓ Each line takes 63.5 microseconds to scan. Horizontal retrace takes 10 microseconds (with 5 microseconds horizontal synch pulse embedded), so the active line time is 53.5 microseconds.

Figure Digital video raster

NTSC Video Color Representation/Compression

- ✓ For color representation, NTSC uses YIQ color model.
- ✓ Basic Compression Idea
 - Eye is most sensitive to Y, next to I, next to Q.*
 - This is still Analog Compression:
- ✓ In NTSC,
 - 4 MHz is allocated to Y,
 - 1.5 MHz to I,
 - 0.6 MHz to Q.

Fig Television standards used in different countries

HDTV (High Definition Television)

High-Definition television (HDTV) means broadcast of television signals with a higher resolution than traditional formats (NTSC, SECAM, PAL) allow. Except for early analog formats in Europe and

Japan, HDTV is broadcasted digitally, and therefore its introduction sometimes coincides with the introduction of digital television (DTV).

- ✓ Modern plasma television uses this
- ✓ It consists of 720-1080 lines and higher number of pixels (as many as 1920 pixels).
- ✓ Having a choice in between progressive and interlaced is one advantage of HDTV. Many people have their preferences

HDTV vs Existing Signals (NTSC, PAL, or SECAM)

The HDTV signal is digital resulting in crystal clear, noise-free pictures and CD quality sound. It has many viewer benefits like choosing between interlaced or progressive scanning.

File Formats

File formats in the PC platform are indicated by the 3 letter filename extension.

.mov= QuickTime Movie Format

.avi= Windows movie format

.mpg =MPEG file format

Four Factors of Digital Video

With digital video, four factors have to be kept in mind. These are :

- Frame rate
- Spatial Resolution
- Color Resolution
- Image Quality

Frame Rate

The standard for displaying any type of non-film video is 30 frames per second (film is 24 frames per second). This means that the video is made up of 30 (or 24) pictures or *frames* for every second of video. Additionally these frames are split in half (odd lines and even lines), to form what are called *fields*.

Color Resolution

This second factor is a bit more complex. Color resolution refers to the number of colors displayed on the screen at one time. Computers deal with color in an *RGB* (red-green-blue) format, while video uses a variety of formats. One of the most common video formats is called *YUV*. Although there is no direct correlation between RGB and YUV,

they are similar in that they both have varying levels of color depth (maximum number of colours).

Spatial Resolution

The third factor is spatial resolution - or in other words, "*How big is the picture?*". Since PC and Macintosh computers generally have resolutions in excess of 640 by 480, most people assume that this resolution is the video standard.

A standard analogue video signal displays a full, over scanned image without the borders common to computer screens. The National Television Standards Committee (NTSC) standard used in North America and Japanese Television uses a 768 by 484 display. The Phase Alternative system (PAL) standard for European television is slightly larger at 768 by 576. Most countries endorse one or the other, but never both.

Since the resolution between analogue video and computers is different, conversion of analogue video to digital video at times must take this into account. This can often the result in the *down-sizing* of the video and the loss of some resolution.

4.4 Image Quality

The last, and most important factor is video quality. The final objective is video that looks acceptable for your application. For some this may be 1/4 screen, 15 frames per second (fps), at 8 bits per pixel. Other require a full screen (768 by 484), full frame rate video, at 24 bits per pixel (16.7 million colours).

Digital video basics

Analog composite signals, such as PAL, NTSC and SECAM, are subject to cumulative distortions and noise that affect the quality of the reproduced picture. Separate distortions of the luminance and chrominance components, as well as intermodulation between them, are likely to occur.

The cumulative analog video signal impairments and their effect on the reproduced picture can be reduced considerably by using a digital representation of the video signal and effecting the distribution, processing and recording in the digital domain. By a proper selection of two parameters, namely the sampling frequency and the quantizing

accuracy, these impairments can be reduced to low, visually imperceptible values. As long as the digitized signals are distributed, processed and recorded in the digital domain, these impairments are limited.

Sampling

The sampling of the video signal is essentially a pulse amplitude modulation process. It consists of checking the signal amplitude at periodic intervals (T). The sampling frequency ($F_s=1/T$) has to meet two requirements:

- ✓ It has to be higher than twice the maximum baseband frequency of the analog video signal (F_B), as stipulated by Nyquist. This is required in order to avoid aliasing. Aliasing is visible as spurious picture elements associated with fine details (high frequencies) in the picture. The only way to avoid aliasing is to use an anti-aliasing filter ahead of the A/D converter. The task of this filter is to reduce the bandwidth of the sampled base band.
- ✓ It has to be coherent with and related to an easily identifiable and constant video frequency.

An early approach, $3F_{SC}$, sampled the composite video signal at three times the color subcarrier frequency. This resulted in $F_s = 3 \times 3.58\text{MHz} = 10.7\text{MHz}$ in NTSC and $F_s = 3 \times 4.43\text{MHz} = 13.29\text{MHz}$ in PAL. A later approach, $4F_{SC}$, sampled the composite video signal at four times the color subcarrier frequency, or 17.7MHz in PAL and 14.3MHz in NTSC.

While sampling at a multiple of F_{SC} works well in PAL and NTSC, it doesn't work at all in SECAM. This is due to the inherent nature of SECAM, which uses two separate line-sequential frequency-modulated color subcarriers carrying, respectively, the D_B and D_R color-difference signals.

It appeared evident in the 1970s that a digital video system in which the luminance and chrominance are individually coded would ease the program interchange between the PAL and SECAM countries. This resulted in the component digital concept, which is at the core of all contemporary digital video systems.

Quantizing

The pulse amplitude modulation results in a sequence of pulses, spaced at $T=1/F_s$ intervals, whose amplitude is proportional to the amplitude of the sampled analog signal at the sampling instant. There are an infinite number of shades of gray — ranging from black (lowest video signal amplitude) to white (highest video signal amplitude) — that the analog video signal can represent.

The instantaneous sampling pulse amplitudes can be represented in the digital domain by only a limited number of binary values, resulting in quantizing errors. The possible number of shades of gray is equal to 2^n , where n is the number of bits per sample.

Experiments have shown that when less than eight bits per sample are used, the quantizing errors appear as contouring. With eight bits per sample or more, the quantizing errors appear, in general, as random noise (quantizing noise) in the picture. In practical applications, in order to avoid clipping, the signal occupies less than 2^n steps, resulting in a specified quantizing range.

Advantages and disadvantages

- ✓ The advantages of digital video are:
- ✓ Single-pass, analog-type impairments are non-cumulative if the signal stays digital.
- ✓ There is a reduced sensitivity to noise and interference.
- ✓ Digital equipment efficiently and economically performs tasks that are difficult or impossible to perform using analog technology.
- ✓ It is amenable to the application of techniques for efficient retention of essential information such as compression.

The disadvantages of digital video are:

- ✓ Analog-type of distortions, as well unique digital distortions related to sampling and quantizing, result in a variety of visible impairments.
- ✓ Wide bandwidth requirements for recording, distribution and transmission necessitate sophisticated bit-rate reduction and compression schemes to achieve manageable bandwidths.
- ✓ Unlike analog signals, the digital signals do not degrade gracefully and are subjected to a cliff effect.