Optibase MPEG Primer

First there was Analog

Several methods exist for the transmission of video signals. The earliest of these was analog. In an analog video signal, each frame of the video is represented by a fluctuating voltage signal. This is known as an analog wave form. One of the earliest of these analog video formats was composite video.

Composite analog video has all the video components (brightness, color, sync, 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.

Composite video quickly gave way to component video, which takes the different components of the video and breaks them into separate signals. Improvements to component video have led to numerous video formats, such as S-Video, RGB, Y, Pb, Pr. All of these are still analog formats however, susceptible to quality loss from one generation to another. Generational loss with video is similar to photocopying, in which a copy of a copy is never as crisp and sharp as the original.

Defining Digital Video

These limitations led to the birth of digital video. Think of digital video as a digital representation of the analog video signal. In the professional video world, there are quite a variety of digital video formats; D1, D2, Digital BetaCam, etc. Unlike analog video that degrades in quality from one generation to the next, digital video does not. Each generation of digital video is virtually identical to the parent.

Even though the video data is digital in nature, virtually all the digital formats are still stored on sequential tape. As a result, most video professionals are more accustomed to working with tape media. Although tape holds considerably more data than a computer hard drive, there are two significant advantages in using computers for digital video. The ability to random access the storage of video and to also compress the video you store.

There are also a number of issues related to the migration of video from traditional video equipment to the computer desktop. These are discussed in detail later in this paper. Considering these issues, digital video for computers requires a different definition and understanding than the digital video formats we previously mention. 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 or sound sample) are represented as a series of binary digits, or bits.

Previous attempts were made to find the best procedure for capturing, storing, and playing back video from the computer desktop. Unfortunately, these early attempts were of a proprietary nature and resulted in various formats and incompatibilities. Subsequently, the International Standards Organization (ISO) worked to define the internationally accepted formats for digital video capture, storage, and playback. These formats will be reviewed in detail later in this paper.

Four Factors of Digital Video

With digital video, we should keep in mind four major factors. These are Frame Rate, Spatial Resolution, Color Resolution and Image Quality.

- **Frame Rate**: The standard for displaying any sort of non-film video is 30 frames per second (film is 24 frames per second). This simply 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 is called "fields". Here again, there is a major difference between the way computers and television handle video. When a television set displays its analog video signal, it displays the odd lines (the odd field) first. Then it displays the even lines (the even field). Each pair forms a frame and there are 60 of these fields displayed every second (or 30 frames every second). This is referred to as "interlaced" video. A computer monitor, however, uses a process called "progressive scan" to update the screen. With this method, the screen is not broken into fields. Instead, the computer displays each line in sequence, from top to bottom. This entire frame is displayed 30 times every second. This is often referred to as "non-interlaced" video.

- **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 (number of maximum colors). Typical RGB color resolutions are 8 bits/pixel (256 colors), 16 bits/pixel (65,535 colors) and 24 bits/pixel (16.7 million colors). Typical YUV color resolutions are 7 bit, 4:1:1 or 4:2:2 (approximately 2 million colors), and 8 bit, 4:4:4 (approximately 16 million colors).

- **Spatial Resolution:** The third factor is spatial resolution--or in other words, "How big is the picture?" Since PC and Macintosh monitors generally have resolutions of 640 x 480 pixels, most people assume that this resolution is the video standard. It isn't. As with RGB and YUV, there is no direct correlation between analog video resolutions and computer display resolutions. A standard analog 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 American and Japanese television uses a 768 x 484 display. The Phase Alternative system (PAL) standard for European television is slightly larger at 768 x 576. Most countries endorse one or the other, but not both. Since the resolution between analog video and computers is different, conversion of analog video to digital video at times must take this into account. This can often result in the downsizing of the video and the loss of some of your resolution.

- **Image Quality**: The last, and ultimately most important factor is video quality. The final objective is video that looks acceptable for your application. For some, this may be a 1/4 screen, 15 frame per second video, at 8 bits per pixel. Others require full screen (768 x 484), full frame rate video (24 or 30 frames per second), at 24 bits per pixel (16.7 million colors).

The Need for Compression

Determining your compression needs is not difficult but does require an understanding of how the four factors mentioned previously (Frame Rate, Color Resolution, Spatial Resolution, and Image Quality) affect your selection. As with most things in this world, there is a price to pay for quality. With more colors, higher resolution, faster frame ates, and better quality, the more horsepower you will need and the more storage space your video will require. By adjusting these factors, you can dramatically change your digital video compression requirements.

Doing some simple math shows that 24-bit color video, with a 640 x 480 resolution, at 30 frames per second, requires an astonishing 26 megabytes of data per second! Not only does this surpass the capabilities of the standard PC-AT's data bus, but it quickly overburdens existing storage systems!

For some users, the way to reduce this amount of data down to a manageable level is to compromise on one of the four factors described above. Certain applications like games, low-end training systems, low-end kiosks, and business presentations may not need a full 30 frame per second frame rate. Instead, the application may be displaying its video in a window, and does not require that the entire frame be captured and stored digitally. As a quick exercise, lets reduce these factors and do our equation again.

As you can see, reducing these parameters, significantly reduces the data requirements. However, the standard PC ISA bus is only capable of sustained data transfer rates of approximately 600 Kilobytes per second. Even if you seriously compromise the video by reducing the size and frame rate, you still have 6 times too much data! In addition, 3.3 megabytes is the amount needed for just one second. For a two hour movie you would still require 23.73 Gigabytes of disk storage!

Reducing the window size even further, sacrificing video quality, and shifting from RGB to YUV 4:1:1 would additionally reduce this data considerably; perhaps to as little as 1.5 megabytes per second. Even this is still too large. This is where digital video compression comes in.

Factors Affecting Compression

The goal of video compression is to massively reduce the amount of data required to store the digital video file, while retaining the quality of the original video. With this in mind, there are several factors that need to be taken into account when discussing desktop digital video compression.

1. Real-Time Versus Non-Real-Time
2. Symmetrical Versus Asymmetrical
3. Compression Ratios
4. Lossless Versus Lossy
5. Interframe Versus Intraframe
6. Bit Rate Control

Real-Time Versus Non-Real-Time

The term "real-time" has been badly abused. In the compression world it means exactly what it says. Some compression systems capture, compress to disk, decompress and play back video (30 frames per second) all in real time; no delays. Other systems are only capable of capturing some of the 30 frames per second and/or are only capable of playing back some of the frames.

Insufficient frame rate is one of the most noticeable video deficiencies. Without a minimum of 24 frames per second, the video will be noticeably "jerky". In addition, the missing frames will contain extremely important lip synchronization data. In other words, if the movement of a person's lips is missing due to dropped frames during capture or playback, it is impossible to match the audio correctly with the video.

Symmetrical versus Asymmetrical

This refers to how video images are compressed and decompressed. Symmetrical compression means that if you can play back a sequence of 640 x 480 video at 30 frames per second, then you can also capture, compress and store it at that rate. Asymmetrical compression means just the opposite. The degree of asymmetry is usually expressed as a ratio. A ratio of 150:1 means it takes approximately 150 minutes to compress one minute of video.

Asymmetrical compression can sometimes be more elaborate and more efficient for quality and speed at playback because it uses so much more time to compress the video. The two big drawbacks to asymmetrical compression are that it takes a lot longer, and often you must send the source material out to a dedicated compression company for encoding (adding time and money to the project).

Compression Ratios

A second ratio is often referred to when working with compressed video. This is the compression ratio and should not be confused with the asymmetry ratio. The compression ratio relates to a numerical representation of the original video in comparison to the compressed video. For example, 200:1 compression ratio means that the original video is represented by the number 200. In comparison, the compressed video is represented by the smaller number, in this case, that is 1. The more the video is compressed, the higher the compression ratio or the numerical difference in the two numbers.

Generally, the higher the compression ratio is, the poorer the video quality will be. With MPEG, compression ratios of 200:1 are common, with good image quality. Motion JPEG provides ratios ranging from 15:1 to 80:1, although 20:1 is about the maximum for maintaining a good quality image. Not only do compression ratios vary from one compression method to another, but hardware and software that perform well on a PC or Mac may be less efficient on a different machine.

Lossless versus Lossy

The "loss" factor determines whether there is a loss of quality between the original image and the image after it has been compressed and played back (decompressed). The more compression, the more likely that quality will be affected. Virtually all compression methods lose some quality when you compress the data. Even if the quality difference is not noticeable, these are considered "lossy" compression methods. At this time, the only "lossless" algorithms are for still image compression. Lossless compression can usually only compress a photo-realistic image by a factor of 2:1.

Interframe versus Intraframe

This is probably the most widely discussed and debated compression issue. The intraframe method compresses and stores each video frame as a discrete picture.Interframe compression, on the other hand, is based on the idea that although action is happening, the backgrounds in most video scenes remain stable--a great deal of the scene is redundant. Compression is started by creating a reference frame. Each subsequent frame of the video is compared to the previous frame and the next frame, and only the difference between the frames is stored. The amount of data saved is substantially reduced.

Bit Rate Control

The final factor to be aware of with video compression is bit-rate control, which is especially important if your system has a limited bandwidth. A good compression system should allow the user to instruct the compression hardware and software which parameters are most important. In some applications, frame rate may be of paramount importance, while frame size is not. In other applications, you may not care if the frame rate drops below 15 frames per second, but the quality of those frames must be of impeccable quality.

The compression hardware and software should allow you to control these parameters to suite your specific application. When evaluating digital video compression systems, look for a system that gives you control. Not all compression systems allow you change every parameter.

Selecting a Compression Type

Compression methods use mathematical algorithms to reduce (or compress) video data by eliminating, grouping and/or averaging similar data found in the video signal. Different algorithms are suited to different purposes. Although there is an alphabet soup of various compression methods, including Motion JPEG (Joint Photographic Experts Group), PLV, Compact Video, Indeo, RTV and AVC, only MPEG-1 and MPEG-2 are internationally recognized standards for the compression of moving pictures.

With so many factors to consider and so many companies touting conflicting solutions, the decision process can be intimidating. First rule of thumb is stick to the standards. Standards don't guarantee that the solution is the best, but they are there for a reason. Years ago two large companies fought over Beta versus VHS video tape formats. Beta was clearly better, but millions of dollars were lost when VHS was adopted as the de facto standard.

The MPEG formats are more than a de facto standard. They are the internationally accepted ISO standard. Some of the most brilliant minds in the video and computer industries have spent years looking at every possible compression solution for full-motion video and are responsible for this specification and standard. This is also not an attempt by one company to push a proprietary format on the computer and video industry. This is an open format available to all.

The second thing to consider is the acceptance of the standard. As many people in the computer and video world are aware, the big joke of standards is that "you have so many to choose from." The unfortunate truth is that many very real standards have been clouded by conflicting pseudo-standards, or in some cases, the real standard just simply wasn't followed, or multiple self-proclaimed "independent standards organizations" developed very different standards and pushed them on the industry.