**Course Outcome to be mentioned which is mapped with this lecture topic.**

Co5 :- To perform different operations on videos.

**Reading Material of this lecture topic.**

**Video Compression**

Video compression is the process of encoding a video file in such a way that it consumes less space than the original file and is easier to transmit over the network/Internet.

It is a type of compression technique that reduces the size of video file formats by eliminating redundant and non-functional data from the original video file.

Video compression is performed through a video codec that works on one or more compression algorithms. Usually, video compression is done by removing repetitive images, sounds and/or scenes from a video. For example, a video may have the same background, image or sound played several times or the data displayed/attached with video file is not that important. Video compression will remove all such data to reduce the video file size.

Once a video is compressed, its original format is changed into a different format (depending on the codec used). The video player must support that video format or be integrated with the compressing codec to play the video file.

There are two approaches to achieving video compression, viz. intra-frame and inter-frame. Intra-frame compression uses the current video frame for compression: essentially image compression. Inter-frame compression uses one or more preceding and/or succeeding frames in a sequence, to compress the contents of the current frame. An example of intra-frame compression is the Motion JPEG (M-JPEG) standard. The MPEG-1 (CD, VCD), MPEG-2 (DVD), MPEG-4, and H.264 standards are examples of inter-frame compression.

Digital video communication is a rapidly developing field, especially with the progress made in video coding techniques. This progress has led to a high number of video applications, such as High-Definition Television (HDTV), videoconferencing and real-time video transmission over multimedia. Due to the advent of multimedia computing, the demand for these video has increased, their storage and manipulation in their raw form is very expensive and it significantly increases the transmission time and makes storage costly (Khalifa and Dlay, 1998). When an ordinary analog video sequence is digitized, it can consume up to 165 Mbps (Jeremiah, 2004; Sullivan and Wiegand, 2005; White Paper, 2008). With most surveillance applications infrequently having to share the network with other data intensive applications and data transfer of uncompressed video over digital networks requires very high bandwidth (Khalifa, 2003). To circumvent this problem, a series of techniques called video compression techniques have been derived to reduce the number of bits required to represent a digital Video data while maintaining an acceptable fidelity or Video quality. Their ability to perform this task is quantified by the compression ratio. The higher the compression ratio is the smaller the bandwidth consumption is.

Data compression is possible because images are extremely data intensive and contain a large amount of redundancy which can be removed by accomplishing some kind of transform, with a reversible linear phase to de-correlate the image data pixels (Khalifa and Dlay, 1998).

To understand the video formats, the characteristics of the video and how these characteristics are used in defining the format need to be understood. Video is a sequence of images which are displayed in order. Each of these images is called a frame. Since, we cannot notice small changes in the frames like a slight difference of colour, video compression standards do not encode all the details in the video; some of the details are actually lost (Abomhara et al., 2010). This is called lossy compression. It is possible to get very high compression ratios when lossy compression is used. Whereas there are some compressions techniques are reversible or non destructive compression (Haseeb and Khalifa, 2006). It is guaranteed that the decompression image is identical to the original image. This is an important requirement for some applications where’ high quality is demanded. This called lossless compression (Khalifa and Dlay, 1998, 1999). Typically, 30 frames are displayed on the screen every second. There will be lots of information repeated in the consecutive frames. If a tree is displayed for one second then 30 frames are used for that tree. This information can be used in the compression and the frames can be defined based upon previous frames. Frames can be compressed using only the information in that frame (intraframe) or using information in other frames as well (intraframe). Intraframe coding allows random access operations like fast forwarding and provides fault tolerance. If a part of a frame is lost, the next intraframe and the frames after that can be displayed because they only depend on the intraframe. Every color can be represented as a combination of red, green and blue. Images can also be represented using this colour space. However, this colour space called RGB is not suitable for compression since it does not consider the perception of humans.

However, the human eye is more sensitive to changes is Y which is part of the YUV colour space where only Y gives the greyscale image. Thus this is used in compression. The Compression ratio is the ratio of the size of the original video to the size of the compressed video. To get better compression ratios pixels are predicted based on other pixels. In spatial prediction, a pixel can be obtained from pixels of the same image while in temporal prediction; the prediction of a pixel is obtained from a previously transmitted image. Hybrid coding is applied if a prediction in the temporal dimension with a suitable decorrelation technique in the spatial domain is used. Motion compensation establishes a correspondence between elements of nearby images in the video sequence. The main application of motion compensation is providing a useful prediction for a given image from a reference image.

DCT (Discrete Cosine Transform) is used in almost all of the standardized video coding algorithms. The DCT is typically done on each 8x8 block (Xiang-Wei et al., 2008, 2009). When DCT is performed, the top left corner has the highest coefficients and the bottom right has the lowest thus making compression easier (Ali, 1999). The coefficients are numbered in a zigzag order from the top left to the bottom right so that there will be many small coefficients at the end. The DCT coefficients are then divided by the integer quantization value to reduce precision. After this division it is possible to lose the lower coefficients if they are much smaller than the quantization.

VIDEO COMPRESSION/DECOMPRESSION TECHNIQUES

When used to convey multimedia transmissions, video streams contain a huge amount of data that requires a large bandwidth and subsequent storage space. As a result of the huge bandwidth and storage requirements, digital video is compressed in order to reduce its storage or transmitting capacity. This technology (video compression) reduces redundancies in spatial and temporal directions. Spatial reduction physically reduces the size of the video data by selectively discarding up to a fourth or more of unneeded parts of the original data in a frame. Temporal reduction, Inter-frame delta compression or motion compression, significantly reduces the amount of data needed to store a video frame by encoding only the pixels that change between consecutive frames in a sequence. Several important standards like Moving Picture Experts Group (MPEG) standard, H.261, 263 and 264 standards are the most commonly used techniques for video compression.

H.261: It was developed in 1990 by the International Telecommunication Union (ITU) developed the H.261 standard for data rates that are multiples of 64 Kbps. H.261 standard uses motion compensated temporal prediction. It supports two resolutions, namely, Common Interface Format (CIF) with a frame size of 352x288 and quarter CIF (QCIF) with a frame size of 172x144 (Girod et al., 1995; Roden, 1996; Choi et al., 1998). The coding algorithm is a hybrid of the following:

Inter-picture prediction: It removes temporal redundancy transform coding, removes spatial redundancy motion compensation and uses motion vectors to compensate.

A macro block, the basic unit of temporal coding, is used to represent a 16x16 pixel region. Each macro block is encoded using intra (I-coding) or predictive) P-coding. Motion prediction uses only the previous picture to minimize delay (Marcel et al., 1997). H.261 is intended for carrying video over ISDN in teleconferencing applications such as videoconferencing and videophone conversations. H.261 is not suitable for usage in general digital video coding.

H.263: It was developed by the International Telecommunication Union (ITU) in 1996. It uses an encoding algorithm called test model (TMN), which is similar to that used by H.261 but with improved performance and error recovery leading to higher efficiency. It is optimized for coding at low bit rates (Nilsson and Naylor, 2003; Raja and Mirza, 2004). H.263 provides the same quality as H.261 but with half the number of bits. A block motion-compensated structure is used for encoding each picture into macroblocks (Ashraf and Chong, 1997). The functionality of H.263 is enhanced by features like: bi-directionally encoded B-frames, overlapped-block motion compensation on 8x8 blocks instead of 16x16 macroblocks, unrestricted motion vector range outside the picture boundary, arithmetic encoding and fractional-pixel motion-vector accuracy (Rijkse, 1996). H.263 supports three other resolutions in addition to QCIF and CIF:

• SQCIF: Approximately half the resolution of QCIF

• 4CIF and 16CIF: 4 and 16 times the resolution of CIF

H.263 is like H.261, is not suitable for usage in general digital video coding. However, H.261 and 263 are a bit contradictory since they both lack some of the more advanced techniques to really provide efficient bandwidth use (Girod et al., 1995; Ashraf and Chong, 1997).

H.263+: It is an extension of H.263 with higher efficiency, improved error resilience and reduced delay. It allows negotiable additional modes, spatial and temporal scalability (Berna et al., 1998; Raja and Mirza, 2004). H.263+ has enhanced features like:

• Reference picture re-sampling motion compensation and picture prediction

• Reduced resolution update mode that permits a high frame rate during rapid motion

• Independent segment decoding mode that prevents the propagation of errors from corrupt frames

• Modified quantization mode improves bit rate control by controlling step size to detect errors and reduce decoding complexity

MPEG-1: The first public standard for the Moving Picture Experts Group (MPEG) committee was the MPEG-1. MPEG-1 was approved in November 1991 and its first parts were released in 1993 (Morris, 1995). It has no direct provision for interlaced video applications (Sikora, 1999) (Roden, 1996). MPEG frames are encoded in three different ways (White Paper, 2008):

• Intra-coded (I-frames): Encoded as discrete frames (still frames), independent of adjacent frames

• Predictive-coded (P-frames): Encoded by prediction from a past I-frame or P-frame, resulting in a better compression ratio (smaller frame)

• Bi-directional-predictive-coded (B-frame): Encoded by prediction using a previous and a future frame of either I-frames or P-frames; offer the highest degree of compression

MPEG-1 decoding can be done in real time using a 350 MHz Pentium processor. It is also suitable for playback from CD-ROM (Ali, 1999).

MPEG-2: The MPEG-2 project was approved in November 1994, focused on extending the compression technique of MPEG-1 to cover larger pictures and higher quality at the expense of higher bandwidth usage. MPEG-2 is designed for digital television broadcasting applications that require a bit rate typically between 4 and 15 Mbps (up to 100 Mbps), such as Digital high definition TV (HDTV), Interactive Storage Media (ISM) and cable TV (CATV) (Sikora, 1997; Ali, 1999). Profiles and levels were introduced in MPEG-2 (Morris, 1995). The profile defines the bit-stream scalability and the color space resolution. With scalability, it is possible to extract a lower bit stream to get a lower resolution or frame rate. The level defines the image resolution, the Y (Luminance) samples/sec, the number of video and audio layers for scalable profiles and the maximum bit-rate per profile. The MPEG compatibilities include upward (decode from lower resolution), downward (decode from higher resolution), forward (decode from previous generation encoding) and backward (decode from new generation encoding). The MPEG-2 input data is interlaced making it compatible with the television scanning pattern that is interlaced.

The MPEG-2 is suitable for TV broadcast applications and high-quality archiving applications. It is not however designed for the internet, as it requires too much bandwidth (Puri et al., 2004).

MPEG-4: It was approved in October 1998 and it enables multimedia in low bit-rate networks and allows the user to interact with the objects (Puri and Eleftheriadis, 1998; (ISO/IEC JTC1/SC29/WG11 N4668, 2002). The objects represent aural, visual or audiovisual content that can be synthetic like interactive graphics applications or natural like in digital television. These objects can then be combined to form compound objects and multiplexed and synchronized to provide QoS during transmission. Media objects can be in any place in the coordinate system. Streamed data can be applied to media objects to change their attributes (Nemcic et al., 2007).

The MPEG-4 compression methods are used for texture mapping of 2-D and 3-D meshes, compression of time-varying streams and algorithms for spatial, temporal and quality scalability, images and video. Scalability is required for video transmission over heterogeneous networks so that the receiver obtains a full resolution display. The MPEG-4 provides a high coding efficiency for storage and transmission of audio-visual data at very low bit-rates (Ali, 1999). About 5-64 Kbps is used for mobile or PSTN video applications and up to 2 Mbps for TV/film applications (Puri et al., 2004).

MPEG-7: It was approved in July 2001 (Chang et al., 2001) to standardize a language to specify description schemes. The MPEG-7 is a different kind of standard as it is a multimedia content description standard and does not deal with the actual encoding of moving pictures and audio. With MPEG- 7, the content of the video is described and associated with the content itself, for example to allow fast and efficient searching in the material.

The MPEG-7 uses XML to store metadata and it can be attached to a timecode in order to tag particular events in a stream. Although, MPEG-7 is independent of the actual encoding technique of the multimedia, the representation that is defined within MPEG-4, i.e., the representation of audio-visual data in terms of objects, is very well suited to the MPEG-7 standard. The MPEG-7 is relevant for video surveillance since it could be used for example to tag the contents and events of video streams for more intelligent processing in video management software or video analytics applications (Avaro and Salembier, 2001; Martinez, 2002).

H.264/AVC: In early 1998, the Video Coding Experts Group (VCEG) ITU-T issued a call for proposals on a project called H.26L, with a target of doubling the coding efficiency in comparison to any other existing video coding standards for various applications. The Moving Picture Expert Group (MPEG) and the Video Coding Expert Group (VCEG) have developed a new and outstanding standard that promises to outperform the earlier MPEG-4 and H.263 standard. Even though the first draft design for the new standard was adopted in October 1999, it provides the most current balance between the coding efficiency, cost and implementation complexity. It has been finalized by the Joint Video Team (JVT) as the draft of the new coding standard for formal approval submission referred to as H.264/AVC and was approved by ITU-T in March 2003 (known also as MPEG-4 part 10) (Wiegand et al., 2003; Nukhet and Turhan, 2005; Jian-Wen et al., 2006). The standard is further designed to give lower latency as well as better quality for higher latency. In addition, all these improvements compared to previous standards were to come without increasing the complexity of design so much that it would be impractical or expensive to build applications and systems. An additional goal was to provide enough flexibility to allow the standard to be applied to a wide variety of applications: for both low and high bit rates, for low- and high-resolution video and with high and low demands on latency. The main features that improve coding efficiency are the following (Ostermann et al., 2004):

• Variable block-size motion compensation with the block size as small as 4x4 pixels

• Quarter-sample motion vector accuracy

• Motion vectors over picture boundaries

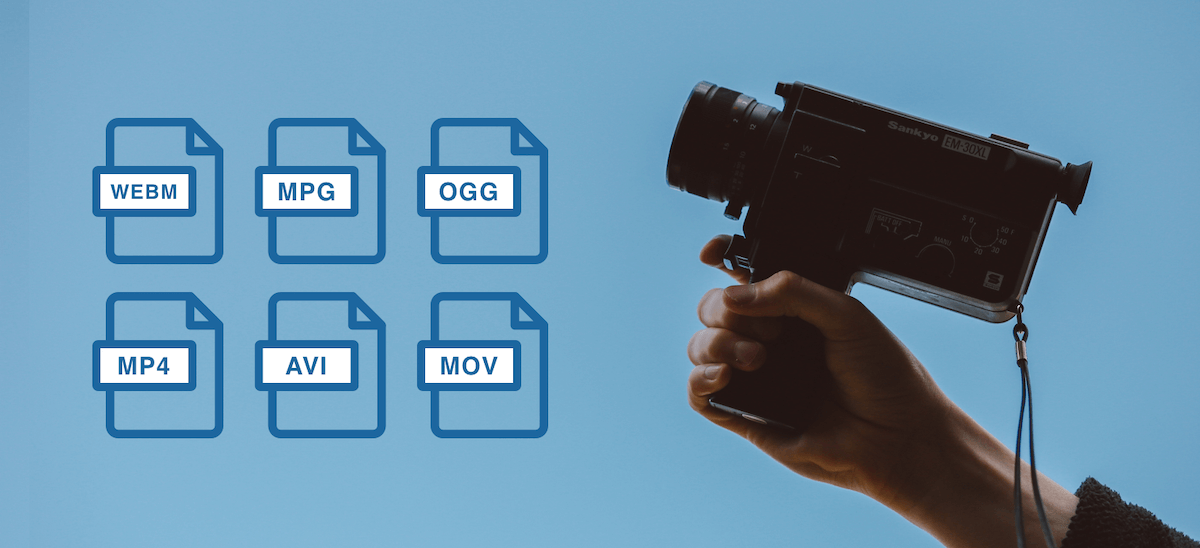
• Multiple reference picture motion compensation

• In-the-loop deblocking filtering

• Small block-size transformation (4x4 block transform)

• Enhanced entropy coding methods (Context- Adaptive Variable-Length Coding (CAVLC) and Context Adaptive Binary Arithmetic Coding (CABAC))

Video File Formats



From .WEBM, .OGG, .MP3, .MP4 and more, there are many video file formats you can expect to come across as a developer handling audio/video content in your application. Like in our last article [The Complete Image File Extension List for Developers](https://blog.filestack.com/api/complete-image-file-extension-list/), and [How to Choose the Best Audio File Format and Codec](https://blog.filestack.com/thoughts-and-knowledge/audio-file-format-codec/), in this article we’ll delve into the different types of audio/video file types, and when to use which type of audio or video format for your application.

First, let’s get a couple of common questions out of the way:

**Which is the best video format?**

Unfortunately, there is no single “best” video format. The best video format for you depends on how you would like to balance the quality and size of the video file. Some formats are extremely small and are great for web [video streaming](https://www.appypie.com/live-tv-video-streaming), but are low quality. Other formats are high quality and the right choice for commercial videography but are very large in size.

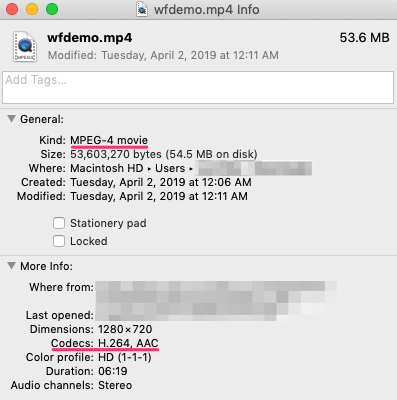
**What is the difference between a video codec, video container, and a video file format?**

Video files are made of 2 parts: a codec and a container.

A video codec is a protocol for encoding and decoding video (the word codec comes from “enCOde / DECode”). Common codecs include H.264, MPEG-4, and DivX. A well-engineered codec has high efficiency, or ability to preserve quality while reducing file size.

The container format is a definition of how the file metadata and data are structured, excluding how the video is actually encoded (which the codec determines). The container file holds the metadata and compressed video data which is encoded using the codec. The container format is also more generally called “the format”, and is reflected in the file’s extension. Common container formats include .AVI, .MP4, and .MOV. Container formats can be paired with different codecs that influence what devices and programs the file will be compatible with.

**Which file format is my video file?**



On Mac, right-click the video file and click “Get Info”, then under “More Info” you should see both the video and audio codec.

On Windows, right-click the file and click “Properties”. Under the “Details” tab you will see the file format and codecs used.

**Which video format is the smallest?**

As of now, the [HEVC](https://en.wikipedia.org/wiki/High_Efficiency_Video_Coding) or H.265 codec is one of the most efficient available on the market and is commonly used to compress 8K UHD video. However, using the codec requires paying a licensing fee so it is not widely compatible or supported by devices or browsers. On the web, .WEBM and its corresponding VP8/VP9 codec are a widely compatible and popular way of making video files smaller.

However, it’s important to consider factors besides size: where the files will be played and the required quality of the video. Next, we’ll dive into each of the most popular container formats to understand the tradeoffs.

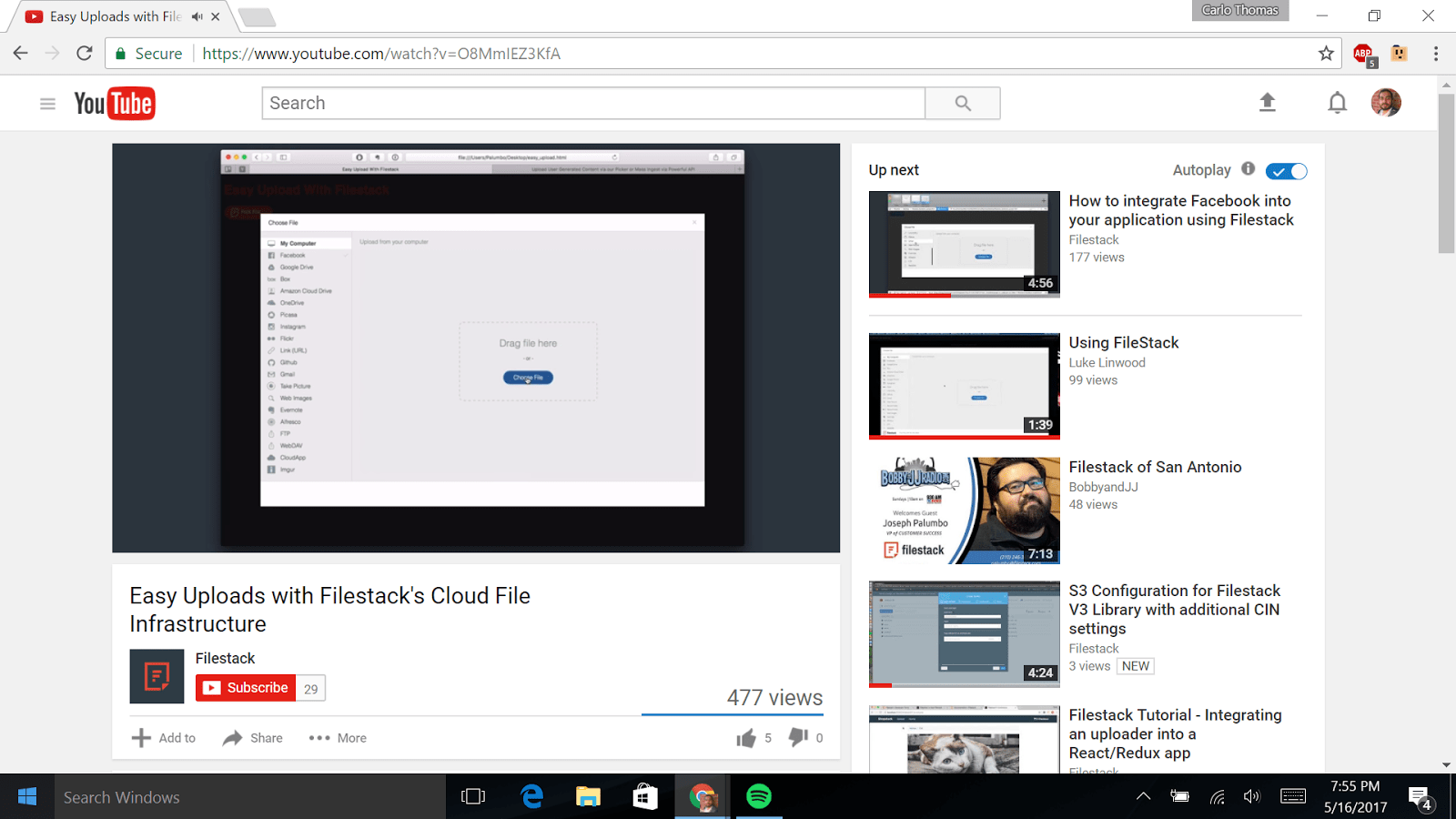
Table of Contents

* [Video File Formats](https://blog.filestack.com/thoughts-and-knowledge/complete-list-audio-video-file-formats/#Video_File_Formats)
  + [.WEBM](https://blog.filestack.com/thoughts-and-knowledge/complete-list-audio-video-file-formats/#WEBM)
  + [.MPG, .MP2, .MPEG, .MPE, .MPV](https://blog.filestack.com/thoughts-and-knowledge/complete-list-audio-video-file-formats/#MPG_MP2_MPEG_MPE_MPV)
  + [.OGG](https://blog.filestack.com/thoughts-and-knowledge/complete-list-audio-video-file-formats/#OGG)
  + [.MP4, .M4P, .M4V](https://blog.filestack.com/thoughts-and-knowledge/complete-list-audio-video-file-formats/#MP4_M4P_M4V)
  + [.AVI](https://blog.filestack.com/thoughts-and-knowledge/complete-list-audio-video-file-formats/#AVI)
  + [.WMV](https://blog.filestack.com/thoughts-and-knowledge/complete-list-audio-video-file-formats/#WMV)
  + [.MOV, .QT](https://blog.filestack.com/thoughts-and-knowledge/complete-list-audio-video-file-formats/#MOV_QT)
  + [.FLV, .SWF](https://blog.filestack.com/thoughts-and-knowledge/complete-list-audio-video-file-formats/#FLV_SWF)
  + [AVCHD](https://blog.filestack.com/thoughts-and-knowledge/complete-list-audio-video-file-formats/#AVCHD)
* [How to choose the best video file format](https://blog.filestack.com/thoughts-and-knowledge/complete-list-audio-video-file-formats/#How_to_choose_the_best_video_file_format)
* [Conclusion](https://blog.filestack.com/thoughts-and-knowledge/complete-list-audio-video-file-formats/#Conclusion)

**Video File Formats**

**.WEBM**

Like the .WEBP image file, .WEBM was created by Google as an efficient means of disseminating media to a large audience. .WEBM video files are relatively small in size, and as such are not as high in terms of quality as some of the other file types on this list. The .WEBM video file format is used for HTML5 video streaming sites, such as YouTube.

Image From: [Filestack YouTube Channel](https://www.youtube.com/watch?v=O8MmIEZ3KfA)

**.MPG, .MP2, .MPEG, .MPE, .MPV**

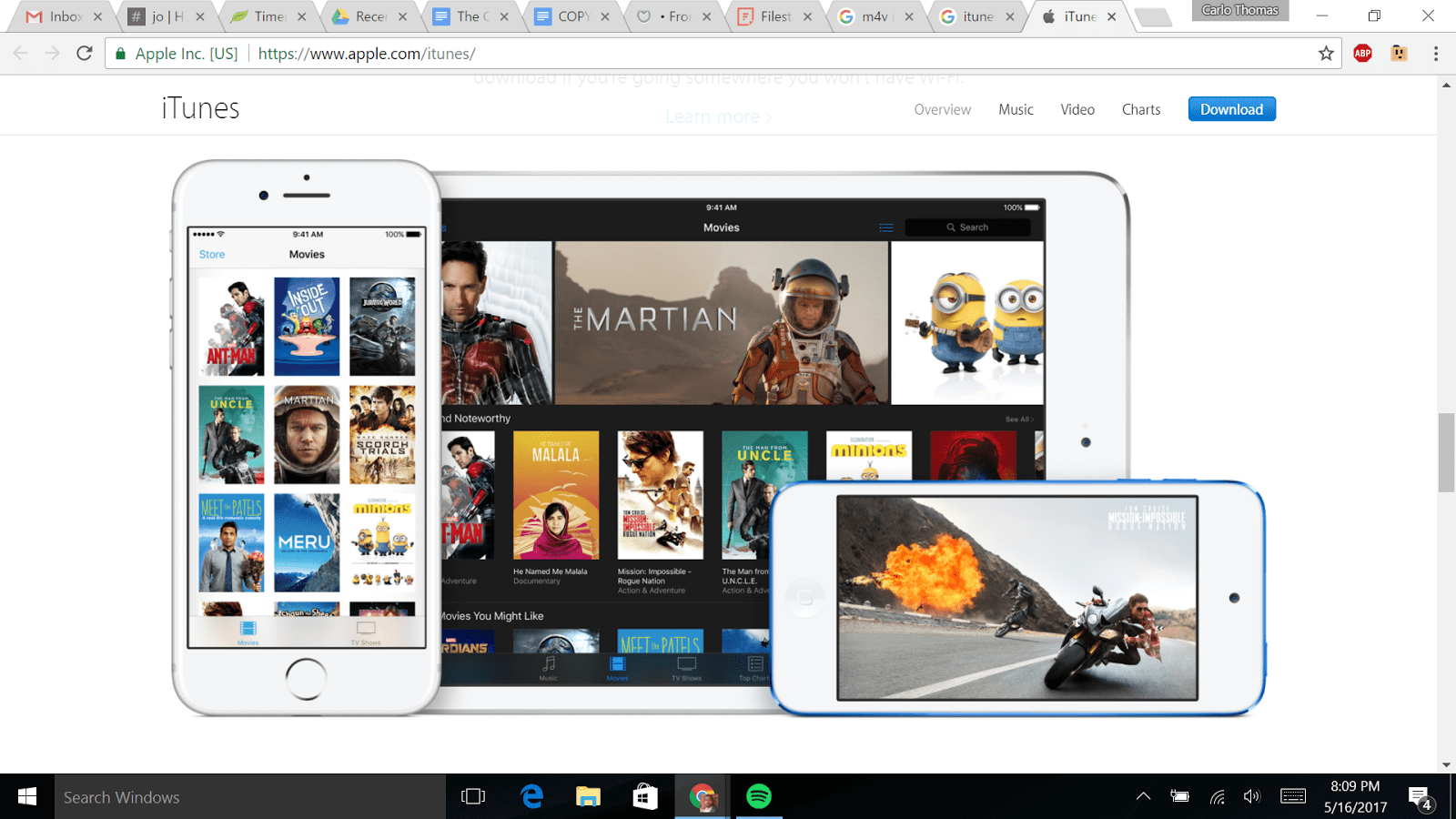
.MPG, .MP2, .MPEG, .MPE, .MPV files can play audio/video media, or simply audio. They are low in file size but also relatively low in quality. They also have lossy compression, meaning their quality will degrade after being edited numerous times. .MPG, .MP2, .MPEG, .MPE, .MPV files are best used when video will be recorded once and never edited.

**.OGG**

.OGG files are an open-source alternative to .MPG files, and are used for high-quality videos to be streamed via the internet. Though .OGG files are used for streaming, they are higher in quality than .WEBM files – meaning they will take longer to be delivered to the end-user. Due to .OGG files being open sourced, they can be used in a variety of applications, including [GPS receivers and media players](http://www.online-convert.com/file-format/ogg) (both desktop and portable).

**.MP4, .M4P, .M4V**

.MP4, .M4P, .M4V are similar to .MPG files in that they can contain audio *and* video, or can simply be solely audio files. .MP4, .M4P, and .M4V are used for streaming video via the internet. They are generally higher in quality than .WEBM files, but tend to be larger in file size. .M4V files are proprietary iTunes files that share the same qualities of .MP4 and .M4P files. M4V files are DRM copy-protected.

Image From: iTunes

**.AVI**

.AVI files are one of the oldest and most compatible video file formats. Many different codecs can be used with an .AVI file, which means that this format has more flexibility in choosing a balance between quality and size. However, these files tend to be larger than the previously mentioned formats, which makes it less ideal for the web and more ideal for storing movies on a computer.

**.WMV**

.WMV is a video file format created by Microsoft and stands for Windows Media Video. The codec used by these files results in small file sizes but poor quality. This format is useful if you are sending video to someone with an older Windows computer.

**.MOV, .QT**

.MOV and .QT files were developed by Apple to use with its Quicktime player. These files are of high quality but large in size. And they have poor compatibility with non-Quicktime players. This format is useful if you intend to archive a high-quality video on an Apple computer.

**.FLV, .SWF**

.FLV and .SWF files were designed by Adobe as the video file format for Flash. The use of these file formats has declined rapidly as Flash has become less popular, especially after Flash support ended for iOS devices. The use of these formats is only recommended if you need to support a legacy system that can only accept this type of file.

**AVCHD**

AVCHD or Advanced Video Coding High Definition files are the format generated by many digital camcorders. These files use the H.264/MPEG-4 video codec and are similar to an .MPG file.

**How to choose the best video file format**

Choosing the right video file format depends entirely on what you plan on using the video for. You should choose a format that achieves the quality of video you require, but nothing more. Unnecessarily high-quality video files can be unwieldy to move, share, convert, and manage. In addition, how the video files will be viewed is important. Not all programs, browsers, and devices can open a specific video format. Consider a couple of scenarios:

* **If the video will be viewed on the web, choose a format that is supported by most browsers**. This way, your video will be able to be played without downloading the file and using a separate player. Browser compatible video formats include MP4 and WEBM.
* **If you are archiving a home video, choose a format that is high quality and has a good chance of being playable in the future**. Open source formats are more future-proof than proprietary formats that are controlled by a specific company. Formats that fit this category are MP4 or AVI (using an open codec).
* **If you work at a company that uses older Windows computers, you should choose a format that is highly compressed and compatible with Windows.** In this case, you’d want to use a WMV file.

**Conclusion**

In summary, the most common **video** file types are:

* .WEBM
* .MPG, .MP2, .MPEG, .MPE, .MPV
* .OGG
* .MP4, .M4P, .M4V
* .AVI, .WMV
* .MOV, .QT
* .FLV, .SWF
* AVCHD
* ***Relevant Video links for this lecture topic.***

[MPEG Algorithm - Video & Audio Compression - Data Compression & Encryption - YouTube](https://www.youtube.com/watch?v=Ky4CaKw7tdo)

* ***Any web-link to be attached relevant to the topic.***

[Video Compression Techniques (slideshare.net)](https://www.slideshare.net/cnssources/video-compression-techniques)

**TEXT BOOKS**

1. Tay Vaughan, “Multimedia making it work”, Tata McGraw-Hill, 2008.
2. Rajneesh Aggarwal & B. B Tiwari, “Multimedia Systems”, Excel Publication, New Delhi, 2007.
3. Li & Drew, “Fundamentals of Multimedia”, Pearson Education, 2009.

* **REFERENCE BOOKS**

1. Parekh Ranjan, “Principles of Multimedia”, Tata McGraw-Hill, 2007
2. Anirban Mukhopadhyay and Arup Chattopadhyay, “Introduction to Computer Graphics and Multimedia”, Second Edition, Vikas Publishing House.