

**Video**

**Training**

**Document**

Contents

[I. Setup Github account and push data: 4](#_Toc132192084)

[II. Introduction of Video 7](#_Toc132192085)

[III. Display Resolution 9](#_Toc132192086)

[ Overscan and underscan- 9](#_Toc132192087)

[ Interlaced and progressive scan- 9](#_Toc132192088)

[ What is a Pixel ? 10](#_Toc132192089)

[ What is the Aspect Ratio? 11](#_Toc132192090)

[ What is the meaning of P and I in Resolution? 11](#_Toc132192091)

[ Different Types of Video Resolutions 12](#_Toc132192092)

[ Video Resolution Chart 12](#_Toc132192093)

[ Standard Definition 13](#_Toc132192094)

[ High Definition 13](#_Toc132192095)

[ Full HD 13](#_Toc132192096)

[ 2K Resolution 14](#_Toc132192097)

[ Ultra High Definition 14](#_Toc132192098)

[ 8K Resolution 14](#_Toc132192099)

[IV. Frame Rate- 14](#_Toc132192100)

[ STANDARD FRAME RATES: 14](#_Toc132192101)

[ 24fps vs 30fps vs 60fps vs 120fps 15](#_Toc132192102)

[ Frame types 15](#_Toc132192103)

[ I Frame (intra, keyframe) 15](#_Toc132192104)

[ P Frame (predicted) 15](#_Toc132192105)

[ B Frame (bi-predictive) 16](#_Toc132192106)

[ Summary 16](#_Toc132192107)

[V. Bit-Depth- 17](#_Toc132192108)

[VI. Video Containers – 18](#_Toc132192109)

[ Formats of video files- 19](#_Toc132192110)

[ Video Codecs vs. Containers: 21](#_Toc132192111)

[ Best Video Codecs for Streaming- 21](#_Toc132192112)

[1. H.264/AVC 21](#_Toc132192113)

[2. H.265/HEVC 21](#_Toc132192114)

[3. AV1 21](#_Toc132192115)

[4. VP9 21](#_Toc132192116)

[5. H.266/VVC 21](#_Toc132192117)

[ H.264/AVC 21](#_Toc132192118)

[ H.265/HEVC 21](#_Toc132192119)

[ VP9 22](#_Toc132192120)

[ H.266/VVC 22](#_Toc132192121)

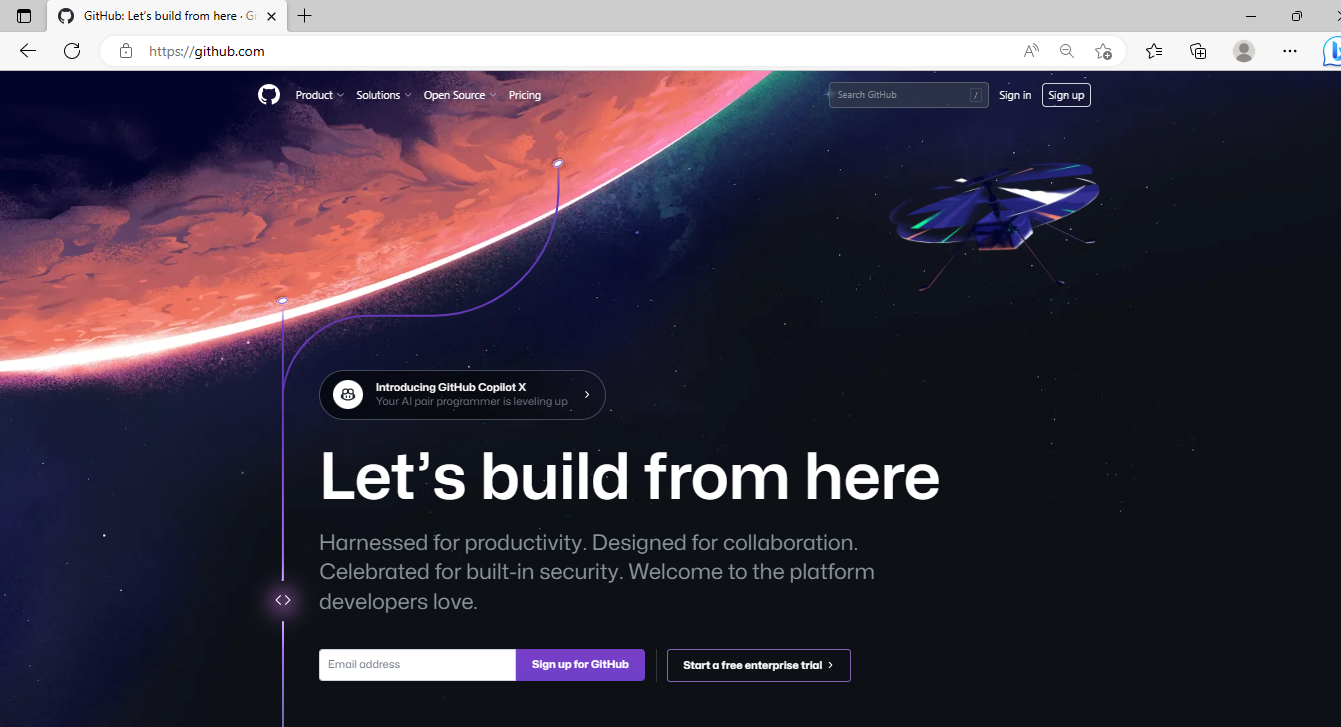
[ Image resizing- 25](#_Toc132192122)

[ Why is Encoding Important? 25](#_Toc132192123)

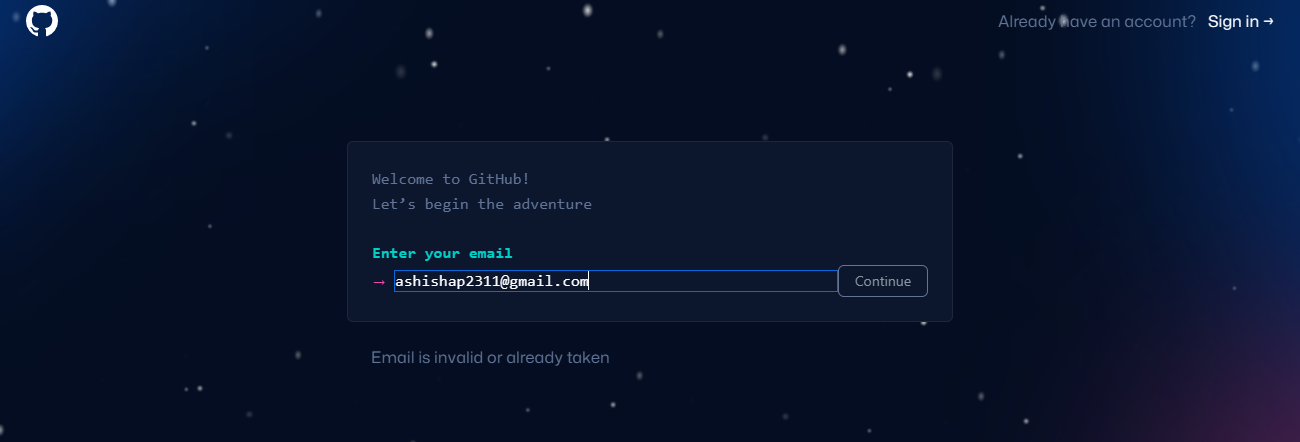
# Setup Github account and push data:

1. Create github account or sign in with your existing github account.

(website: <https://github.com/> )

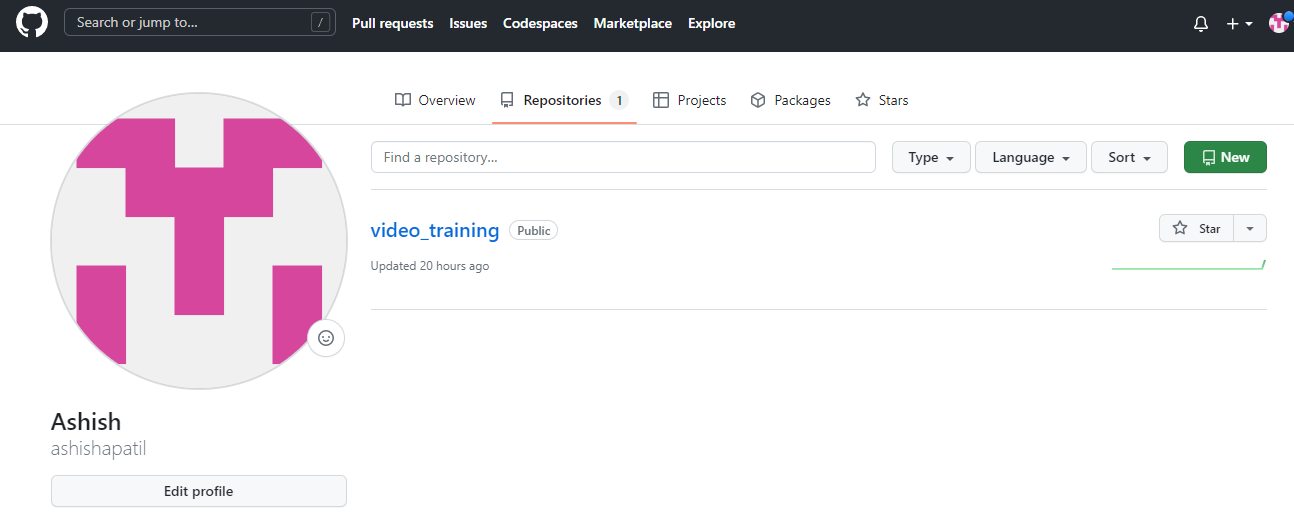


1. By using your mail-id you can sign in or sign up on github.

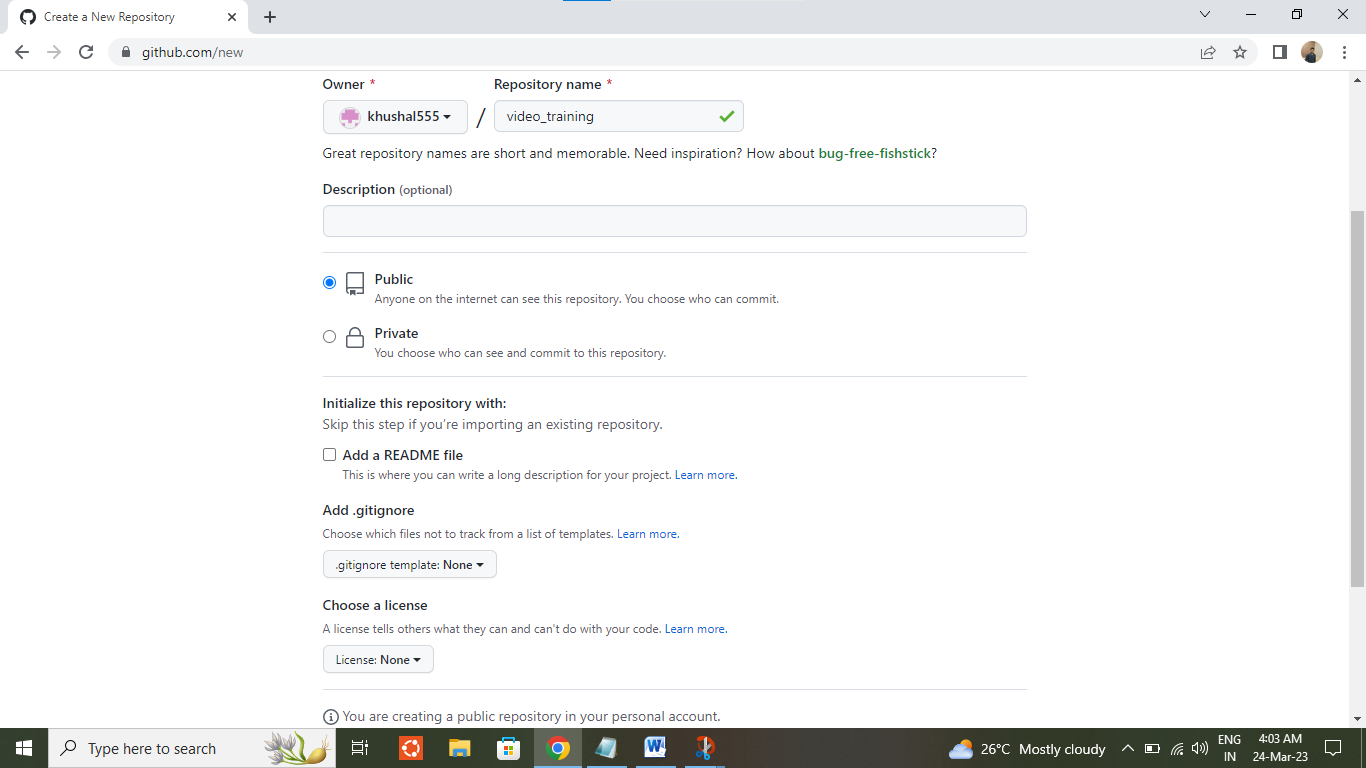


3. Fill the details and create password and solve the puzzle.

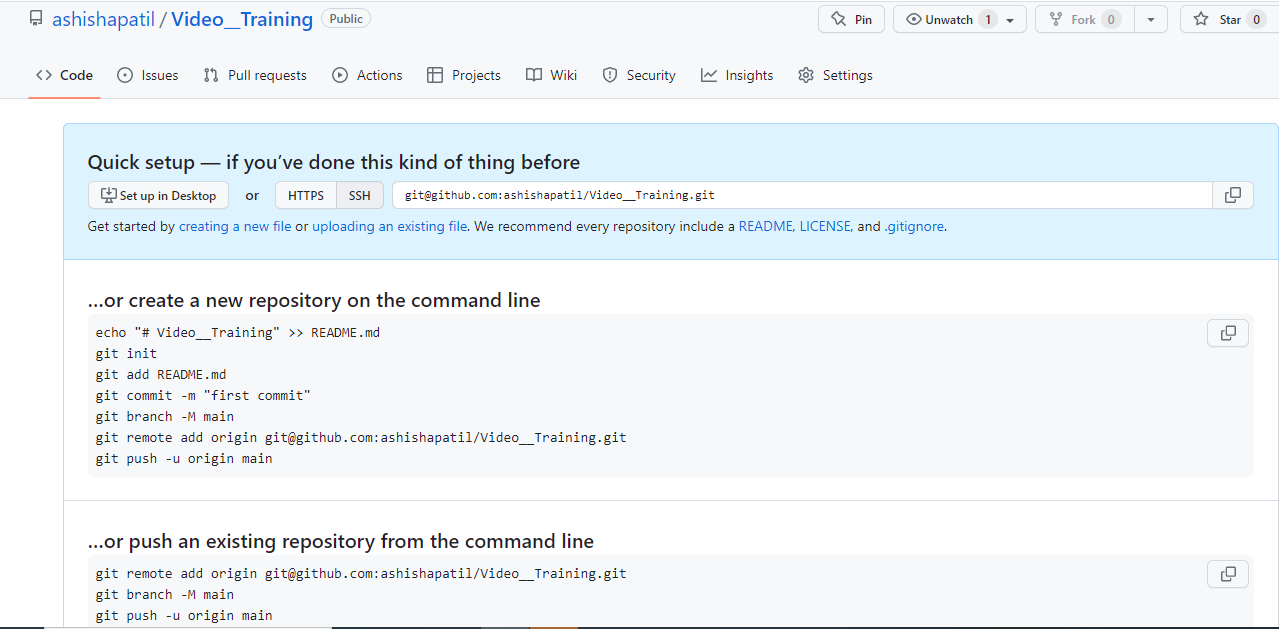
1. At the end, click on create account and your github account will be created.
2. After login to your account, create one repository as shown.



1. **And create new repo**
2. **Open repo**
3. Enter your repo name and select privacy as public or private as per requirement. And select create repository.



1. By using SSH url will be used for push or pull data from github account.



1. Configuration and setup-

You have to add a remote repo url to your local git config by using command-

**git remote add <remote\_name > <remote\_repo\_url>**

This command will map remote repository at to a ref in your local repo under.

After mapping you can push local branches to it.

**git push -u <remote\_name> <local\_branch\_name>**

1. While push any data username and password, have to enter. This password will be a token number and can be generate by following link-

<https://docs.github.com/en/authentication/keeping-your-account-and-data-secure/creating-a-personal-access-token>

# Introduction of Video

A sequence of images processed electronically into an analog or digital format and displayed on a screen with sufficient rapidity as to create the illusion of motion and continuity.

A video is the art of integrating audio and visual elements into one form of multimedia, providing the audience with an image that is recorded electronically by a video camera and displayed on a television or device screen. Videos can be used for education, entertainment, or other purposes, which use pictures, graphics, or text.

Technology that involves the recording and playing back of moving pictures and sound: The next generation of interactive video technology will blend the viewer's image directly into the action on screen.

Light waves of object passes through lens, that light converted into electrical signals using Charge Coupled Device (CCD). In cameras three CCD use for three different hue (R, G, and B).

Video follow three standards:

1. Luminance: Brightness of pixel.
2. Chrominance: Color information
3. Synchronization: Series of electronic pulses that control by time of each frame of video.

* Video is categorized in two types-

1. Analog Video
2. Digital Video

The analog process encodes video and audio in complete frames (modulation), with the receiving device interpreting and translating the signal into video and audio on a monitor (de-modulation). This process can introduce a progressive loss of data leading to a general loss of video quality. NTSC can only deliver 720 pixels wide video or stills from video.

Digital video, or DV, on the other hand, remains digital (such as '0's and '1's ) with the data describing the colors and brightness of each video frame. On the receiving end of this data transmission, there is no translation or interpretation, just the delivery of pure data. The consistency of delivery is the crucial advantage that digital video has over analog video when it comes to working with images on a PC. As opposed to NTSC, there is no limit to resolution so images or movies as wide as 4000 pixels are easily obtainable with the digital cameras we sell for microscopy.

|  |  |  |
| --- | --- | --- |
| **Factor** | **Digital** | **Analog** |
| Signal Type | Digital Systems use discrete signals as on/off representing binary format. Off is 0, On is 1. | Analog Systems use continuous signals with varying magnitude. |
| Wave Type | Digital Systems use square waves. | Analog systems use sine waves. |
| Technology | Digital systems first transform the analog waves to limited set of numbers and then record them as digital square waves. | Analog systems records the physical waveforms as they are originally generated. |
| Transmission | Digital transmission is easy and can be made noise proof with no loss at all. | Analog systems are affected badly by noise during transmission. |
| Flexibility | Digital system hardware can be easily modulated as per the requirements. | Analog system's hardware are not flexible. |
| Bandwidth | Digital transmission needs more bandwidth to carry same information. | Analog transmission requires less bandwidth. |
| Memory | Digital data is stored in form of bits. | Analog data is stored in form of waveform signals. |
| Power requirement | Digital system needs low power as compare to its analog counterpart. | Analog systems consume more power than digital systems. |
| Best suited for | Digital systems are good for computing and digital electronics. | Analog systems are good for audio/video recordings. |
| Cost | Digital systems are costly. | Analog systems are cheap. |
| Example | Examples of digital systems include Computers, CD, and DVD. | Examples of analog systems include analog electronics, voice radio using AM frequency. |

# Display Resolution

Display resolution or screen resolution expresses how many pixels are present in a display or entire screen. A given display will have a maximum resolution that depends on its physical ability to focus light. For example, a 15-inch 640x480 monitor will display approximately 50 dots per inch.

### Overscan and underscan-

Most television display manufacturers "overscan" the pictures on their displays (CRTs and PDPs, LCDs etc.), so that the effective on-screen picture may be reduced from 720 × 576 (480) to 680 × 550 (450), for example. The size of the invisible area somewhat depends on the display device. Some HD televisions do this as well, to a similar extent.

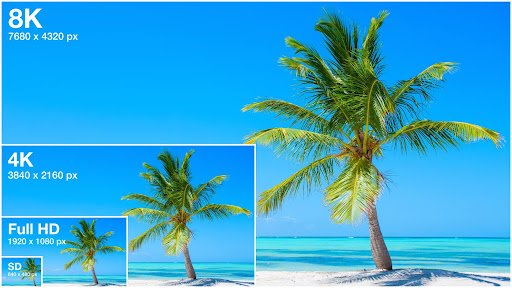
Computer displays including projectors generally do not overscan although many models (particularly CRT displays) allow it. CRT displays tend to be underscanned in stock configurations, to compensate for the increasing distortions at the corners.

### Interlaced and progressive scan-

[Interlaced video](https://en.wikipedia.org/wiki/Interlaced_video) (also known as **interlaced scan**) is a technique for doubling the perceived [frame rate](https://en.wikipedia.org/wiki/Frame_rate) of a video display without consuming extra [bandwidth](https://en.wikipedia.org/wiki/Bandwidth_(signal_processing)). The interlaced signal contains two [fields](https://en.wikipedia.org/wiki/Field_(video)) of a video frame captured consecutively. This enhances motion perception to the viewer, and reduces [flicker](https://en.wikipedia.org/wiki/Flicker_(screen)) by taking advantage of the [phi phenomenon](https://en.wikipedia.org/wiki/Phi_phenomenon).

The [European Broadcasting Union](https://en.wikipedia.org/wiki/European_Broadcasting_Union) has argued against interlaced video in production and broadcasting. The main argument is that no matter how complex the deinterlacing algorithm may be, the artifacts in the interlaced signal cannot be completely eliminated because some information is lost between frames. Despite arguments against it, television standards organizations continue to support interlacing. It is still included in digital video transmission formats such as [DV](https://en.wikipedia.org/wiki/DV), [DVB](https://en.wikipedia.org/wiki/Digital_Video_Broadcasting), and [ATSC](https://en.wikipedia.org/wiki/ATSC). New video compression standards like [High Efficiency Video Coding](https://en.wikipedia.org/wiki/High_Efficiency_Video_Coding) are optimized for [progressive scan](https://en.wikipedia.org/wiki/Progressive_scan) video, but sometimes do support interlaced video.

[Progressive scanning](https://en.wikipedia.org/wiki/Progressive_scanning) (alternatively referred to as **noninterlaced scanning**) is a format of displaying, storing, or transmitting [moving images](https://en.wikipedia.org/wiki/Moving_image) in which all the lines of each [frame](https://en.wikipedia.org/wiki/Film_frame) are drawn in sequence. This is in contrast to [interlaced video](https://en.wikipedia.org/wiki/Interlaced_video) used in traditional [analog television](https://en.wikipedia.org/wiki/Analog_television) systems where only the odd lines, then the even lines of each frame (each image called a [video field](https://en.wikipedia.org/wiki/Video_field)) are drawn alternately, so that only half the number of actual image frames are used to produce video.



when a video is said to have a resolution of 3840×2160 pixels, it essentially means the video has a width of 3840 pixels and a height of 2160 pixels.

So right now, it must be obvious that the higher the resolution, the sharper and crisper the video image will be. Theoretically, it has more clarity because, since the number of pixels is high, there are more blocks for colors to fill in, thus giving a better image quality.

Currently, there are various types of video resolutions. There exist multiple types of video resolution rather than “One Ultimate Resolution”. When devices are manufactured, may it be mobile phones, monitors, televisions, or digital billboards, the screen resolution is predetermined. So in order to choose a video resolution, you have to understand your target devices.

For example, you might have watched movies either on television or mobile devices with the video dimension of 640 x 360, 640 x 480, or 720 x 480. Sometimes users prefer to watch clear & sharper images with brighter colors. Movies like Avatar, Avengers, Interstellar can be better viewed in high definition with the value of 1280 x 720 (720p) or 1920 x 1080.

## What is a Pixel ?

The word “pixel” refers to a picture element in an image. These pixels are the smallest unit of an image. The images comprise hundreds of thousands (and often millions) of pixels. When every pixel gets combined a resolution is formed. It will be in 3 components, known as RGB (Red, Green, Blue), or 4 components, known as CMYK (Cyan, Magenta, Yellow, blacK).

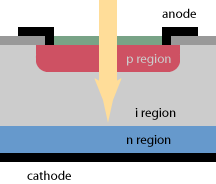
### What is an image sensor?

A device that is used to detect information usable for creating images. It uses photodiodes to capture light and transform it into an electric signal.

#### What is a photodiode?

A photodiode works quite similar to a normal diode, which is an electrical component that ideally has an infinitely small resistance in one direction and an infinitely big resistance in the other direction (therefore directing a flow of current). It uses a [P-N junction](https://en.wikipedia.org/wiki/P%E2%80%93n_junction) of two semiconductors (mostly silicon), meaning one side is over-saturated with electrons and the other is lacking electrons. The photodiode has been depleted of mobile charge carriers (by applying [reverse bias](https://en.wikipedia.org/wiki/P%E2%80%93n_junction#Reverse_bias)), which creates an electric field, which when hit by a photon (e-), hole(h+) pair attracts the photon to the positive charged N-layer and the positive hole charge to the P-layer. The energy of the photon has to be big enough, to attract an electron across the depletion region of the semiconductor, which in the case of silicon fits every light-wave with a wave-length shorter or equal to near-infrared light. That current flow is then amplified and later converted into a digital signal.

This is just a quick & dirty summary, if you want to go deeper into this topic start with: [[4]](https://youtu.be/4Deyx3RighA).

 [[www.rp-photonics.com]](https://www.rp-photonics.com/photodiodes.html)

#### What are common types of image sensors?

**CCD (Charge Coupled Device)**

* Each cell of the CCD is an analog device, which generates an electric charge that is shifted towards the single amplifier for the column.
* the signal has to be converted to a digital signal with an ADC (Analog to Digital Converter).
* Used for high-quality video cameras as they produce less noise than a CMOS

**CMOS (Complementary Metal Oxid Semiconductor)**

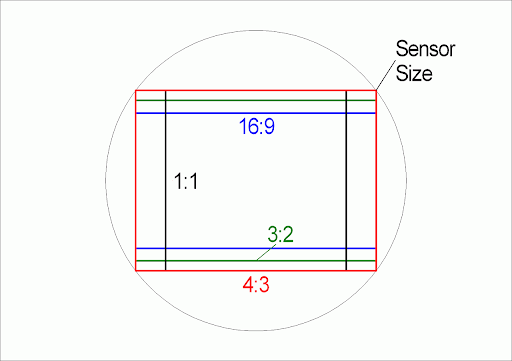
* Uses an amplifier for each sensor cell together with micro lenses in order to not lose information of photons hitting the amplifier instead of the cell.
* Cheaper and more commonly used, with a lower power consumption than a CCD.

## What is the Aspect Ratio?

If you are to understand the different types of video resolutions, you need to understand what the [video aspect ratio](https://www.vdocipher.com/blog/video-aspect-ratio/) is and why it is important.

Aspect ratios, in simple words, are the ratio of horizontal pixels (width) to vertical pixels (height).

For example, if a device is known to have a predefined resolution of 1920×1080 pixels, it has an aspect ratio of 16:9.

[](https://www.vdocipher.com/blog/wp-content/uploads/2021/09/video-resolution-aspect-ratio.png)

Most videos on the internet have an aspect ratio of 16:9, as it is said to support the majority of devices, even television screens.

The reason aspect ratios are an important factor to consider is that it affects the final display and makes sure the cropping is not limited where it should not be.

If you are producing a video of aspect ratio 1:1, it can be concluded that wherever it is displayed, it is displayed as a proportional square. So knowing about the [video aspect ratio](https://www.vdocipher.com/blog/video-aspect-ratio/), in the beginning, helps you understand how the final video image will be produced and how it will be displayed on different screens.

## What is the meaning of P and I in Resolution?

‘P’ refers to Progressive video where it makes use of horizontal lines to make up the image. examples to consider 480p, 720p or 1080p. Displays both the scan lines (even & odd) simultaneously. This type of videos results in better, fast sequences providing sharpness to the video. The progressive display is used in all monitor screens.

‘I’ refer to interlaced video. This technology was introduced in the early days of television to fit a picture into the desired bandwidth. An example to consider 480i or 1080i. It displays even and odd scan lines in separate fields. Two of these even and odd scan line fields form one video frame.

## Different Types of Video Resolutions

Every new type of video resolution is clearer and crisper than its predecessor. Let’s go through the different types of video resolutions.

Sometimes, rather than saying a dimension, we refer to different resolutions by their names given based on how they appear on screens. The different types of video resolutions are:

* Standard Definition (SD)
* High Definition (HD)
* Full HD
* 2K resolution
* Ultra High Definition (UHD) or 4K
* 8K resolution

Each type has its own number of pixels and aspect ratios.

### Video Resolution Chart

|  |  |  |  |
| --- | --- | --- | --- |
| Resolution | Name | Aspect Ratio | Pixel Size |
| SD (Standard Definition) | 480p | 4:3 | 640 x 480 |
| HD (High Definition) | 720p | 16:9 | 1280 x 720 |
| Full HD (FHD) | 1080p | 16:9 | 1920 x 1080 |
| 2K video (Quad HD) | 2k or 1440 | 16:9 | 2560 x 1440 |
| 4K video or Ultra HD (UHD) | 4K or 2160p | 1:1.9 | 3840 x 2160 |
| 8K video or Full Ultra HD | 8K or 4320p | 16∶9 | 7680 x 4320 |

### Standard Definition

Standard Definition, as the name would suggest, does not have a lot of pixels, resulting in lower resolution. Two popular SD resolutions are 640×360 and 640×480. These resolutions are best for small mobile devices. The aspect ratio for standard video is 4:3

In Standard Definition, videos are of low quality, with lower bit rates and smaller file sizes. A good option to consider if you are on slow internet connection.  With the lower bit rate videos can be streamed more smoothly with less chance of buffering compared to higher quality videos.

### High Definition

High Definition or HD is a step ahead of standard definition. When it comes to HD resolutions, the video or screen is being specific about only two resolutions – 1920×1080 pixels or 1280×720 pixels. 1920×1080 resolution (1080p) refers 1920 pixels horizontally and 1080 pixels vertically. HD videos come with a fixed aspect ratio of 16:9. HD videos are best for small online content. You can read more about [SD vs HD](https://www.vdocipher.com/blog/sd-vs-hd) in this blog.

### Full HD

Full HD is the same as HD, but the resolution is strictly 1920×1080. In other words, full HD videos are strictly 1080p.

‘p’ here refers to progressive scan. It is the type of display and signal used to refresh the pixels periodically. This resolution is widely used in most smartphones currently.

### 2K Resolution

2K Resolution is also known as Quad High Definition (QHD). Quad High Definition is so much more advanced than HD or Full HD. Most commonly used in advanced smartphones and gaming monitors. It has an aspect ratio of 1:1.77 and the dimensions are 2048×1152 pixels.

### Ultra High Definition

Ultra High Definition is commonly known as 4K resolution. Generally speaking, there isn’t much difference between 2K and 4K. The only difference is, since the size is comparatively bigger, it adds more space to zoom in. There won’t be any noticeable difference unless the color grading deep inside is very different.

4K resolutions are mostly used for the theatrical releases of movies and work well on big screens.

The dimensions of UHD are 3840×2160 and the aspect ratio is 16:9. It is commonly referred to as 2160p.

### 8K Resolution

8K Resolution is very advanced, and not widely used. But the image quality is very intense and it has a lot of space for editing and zooming. If 8K resolution is chosen, it is primarily to focus on reframing and enhanced visual effects. The dimensions are 7680×4320 pixels.

# Frame Rate-

Frame rate is the number of individual video frames that your camera captures, per second. In video production, a video’s frame rate is expressed as frames per second (fps).

## STANDARD FRAME RATES:

1. [24fps: Cinematic Standard](https://wistia.com/learn/production/what-is-frame-rate#24fps-cinematic-standard)
2. [30fps: Video Standard](https://wistia.com/learn/production/what-is-frame-rate#30fps-video-standard)
3. [60fps: Slow-Motion Standard](https://wistia.com/learn/production/what-is-frame-rate#60fps-slow-motion-standard)

### [24fps vs 30fps vs 60fps vs 120fps](https://wistia.com/learn/production/what-is-frame-rate#24fps-vs-30fps-vs-60fps-vs-120fps)

The difference between different frame rates has to do with how the image looks. 24fps, 30fps, and 60fps all have different looks, with the main difference between each being the number of frames captured per second.

When choosing your frame rate, you’ll want to keep in mind that the higher the frame rate, the slower the slow motion will be. For example, videos recorded in 60fps or 120fps will then be slowed down to a 24ps frame rate, which creates that smooth slow-motion effect. If you ever play back a 60fps or 120fps how it is recorded, then you’ll definitely notice a strange-looking effect.

## Frame types

Now we can move on and try to eliminate the **redundancy in time** but before that let's establish some basic terminology. Suppose we have a movie with 30fps, here are its first 4 frames.

[](https://github.com/leandromoreira/digital_video_introduction/blob/master/i/smw_background_ball_1.png) [](https://github.com/leandromoreira/digital_video_introduction/blob/master/i/smw_background_ball_2.png) [](https://github.com/leandromoreira/digital_video_introduction/blob/master/i/smw_background_ball_3.png) [](https://github.com/leandromoreira/digital_video_introduction/blob/master/i/smw_background_ball_4.png)

We can see **lots of repetitions** within frames like **the blue background**, it doesn't change from frame 0 to frame 3. To tackle this problem, we can **abstractly categorize** them as three types of frames.

### I Frame (intra, keyframe)

An I-frame (reference, keyframe, intra) is a **self-contained frame**. It doesn't rely on anything to be rendered, an I-frame looks similar to a static photo. The first frame is usually an I-frame but we'll see I-frames inserted regularly among other types of frames.

[](https://github.com/leandromoreira/digital_video_introduction/blob/master/i/smw_background_ball_1.png)

### P Frame (predicted)

A P-frame takes advantage of the fact that almost always the current picture can be **rendered using the previous frame.** For instance, in the second frame, the only change was the ball that moved forward. We can **rebuild frame 1, only using the difference and referencing to the previous frame**.

[](https://github.com/leandromoreira/digital_video_introduction/blob/master/i/smw_background_ball_1.png) <- [](https://github.com/leandromoreira/digital_video_introduction/blob/master/i/smw_background_ball_2_diff.png)

#### Hands-on: A video with a single I-frame

Since a P-frame uses less data why can't we encode an entire [video with a single I-frame and all the rest being P-frames?](https://github.com/leandromoreira/digital_video_introduction/blob/master/encoding_pratical_examples.md#1-i-frame-and-the-rest-p-frames)

After you encoded this video, start to watch it and do a **seek for an advanced** part of the video, you'll notice **it takes some time** to really move to that part. That's because a **P-frame needs a reference frame** (I-frame for instance) to be rendered.

Another quick test you can do is to encode a video using a single I-Frame and then [encode it inserting an I-frame each 2s](https://github.com/leandromoreira/digital_video_introduction/blob/master/encoding_pratical_examples.md#1-i-frames-per-second-vs-05-i-frames-per-second) and **check the size of each rendition**.

### B Frame (bi-predictive)

What about referencing the past and future frames to provide even a better compression?! That's basically what a B-frame is.

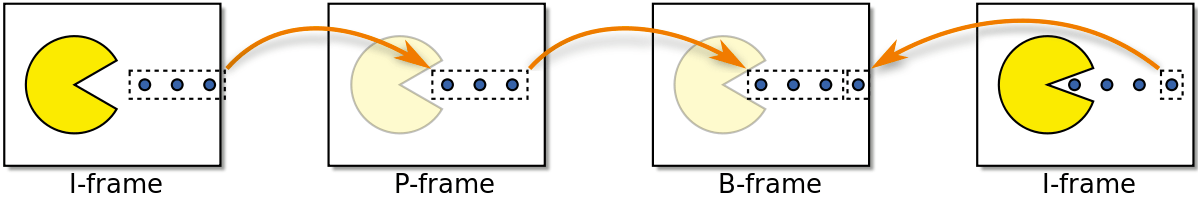
[](https://github.com/leandromoreira/digital_video_introduction/blob/master/i/smw_background_ball_1.png) <- [](https://github.com/leandromoreira/digital_video_introduction/blob/master/i/smw_background_ball_2_diff.png) -> [](https://github.com/leandromoreira/digital_video_introduction/blob/master/i/smw_background_ball_3.png)

#### Hands-on: Compare videos with B-frame

You can generate two renditions, first with B-frames and other with [no B-frames at all](https://github.com/leandromoreira/digital_video_introduction/blob/master/encoding_pratical_examples.md#no-b-frames-at-all) and check the size of the file as well as the quality.

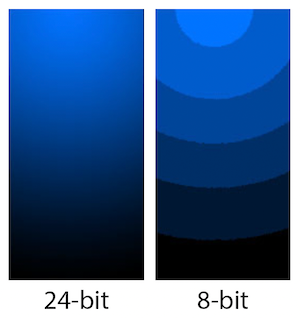
### Summary

These frames types are used to **provide better compression**. We'll look how this happens in the next section, but for now we can think of **I-frame as expensive while P-frame is cheaper but the cheapest is the B-frame.**

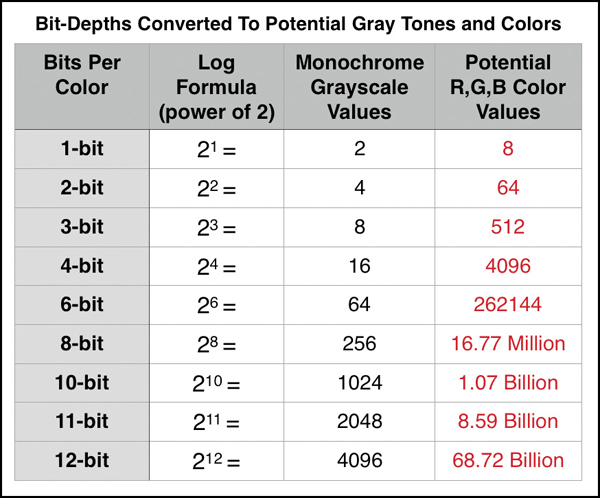
[](https://github.com/leandromoreira/digital_video_introduction/blob/master/i/frame_types.png)

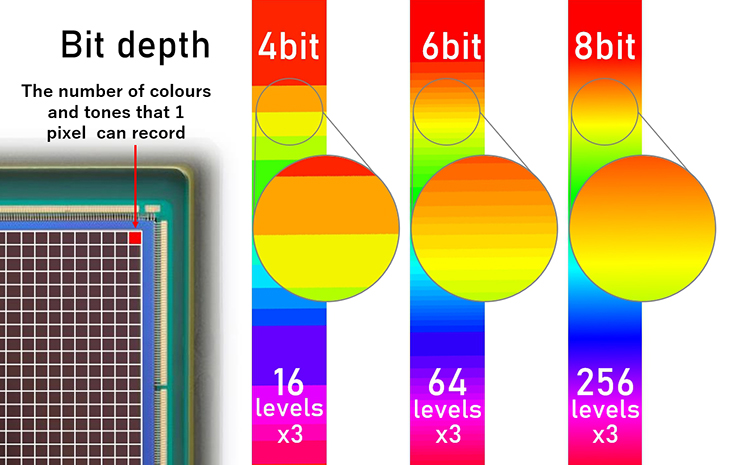
# Bit-Depth-

Bit depth describes the level of detail for the color information stored in an image. The higher the bit depth an image has, the more colors it can store. The more colors that can be stored, the smoother and more accurate an image looks. The lower the bit depth is for an image, the more likely you are to see sharp delineations between color changes in the images. Sharp changes that don't look like an image in real life are referred to as banding.



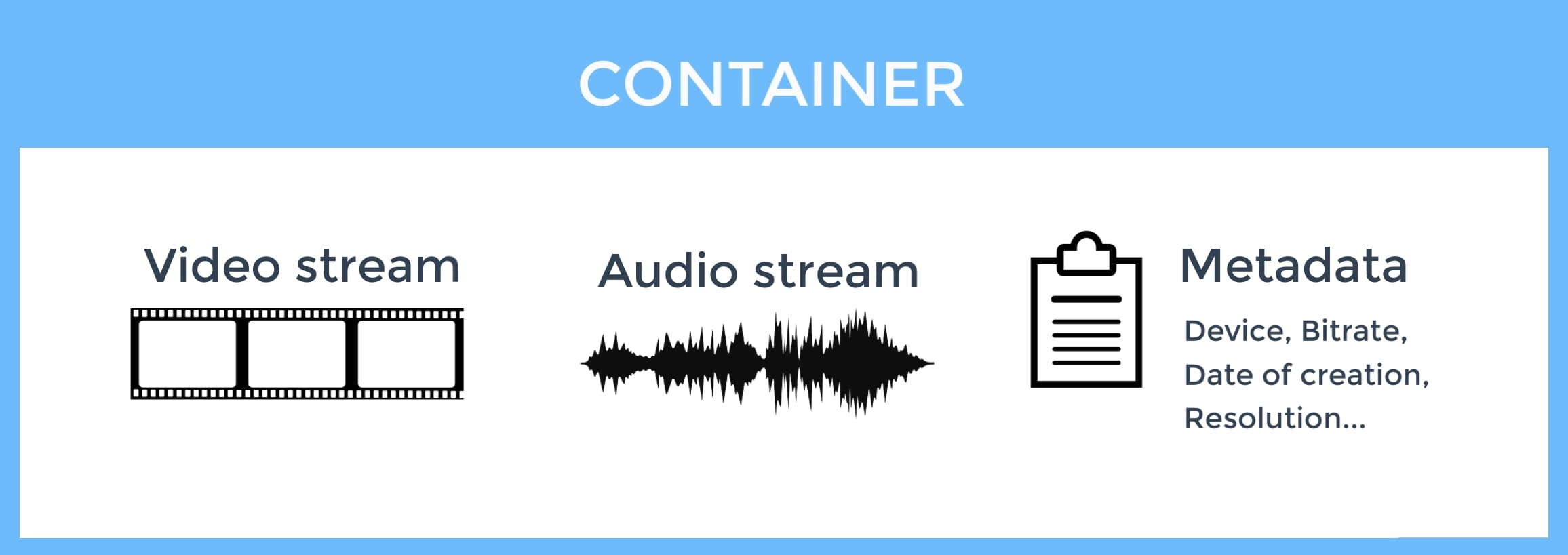
In the image presented above you can see how increased bit depth (the image on the left) yields a smoother image, while less bit depth results in what is known as banding - where you can see separate bands of color instead of a smooth blend between the colors. With techniques like [dithering](https://api.video/what-is/dithering), you can conceal banding with varying degrees of success and avoid going up to a higher and more data expensive quality level. It depends on the image, the technique you use, and how high quality you need the final output to be.

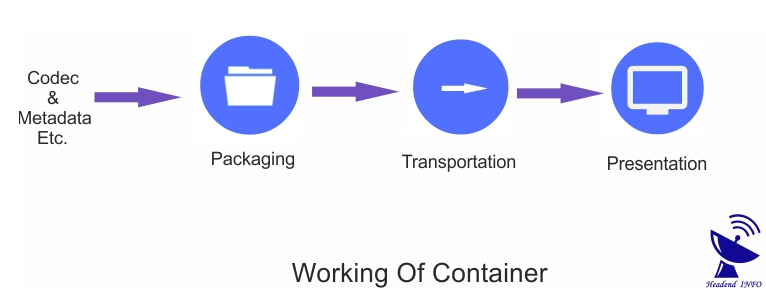




# Video Containers –

A container for video files packages the various parts of the file. This includes the visual images as well as the audio tracks and graphics. Containers are also referred to as a the format of that file. The common video containers are .mov .mp4.





## Formats of video files-

There are numerous video file formats available that can be used for videos, which provide different-different video quality. A list is given below that has some common kinds of video files.

* **.3GP:**  3GP is a compression file format created by 3GPP (3rd Generation Partnership Project), which is a basic version of MPEG-4 part 14 (MP4) format. It is used to compress audio, video, and other multimedia to save on disk space, bandwidth, and data usage. It often creates files for use with cell phones. There are different video players that can play this file format videos on your computer, smartphone, or tablet. The .3GP or .3G2 are the file extensions of the 3GP file format.
* **.MP4:**  MP4 is a file format that was introduced in 1998, which stands for MPEG-4 Part 14 and, it was agreed upon as a standard by the MPEG. It is an audio and video compression standard, commonly used to store video and audio as well as store subtitles and still images. Like the MPEG-1 and MPEG-2 standards, MPEG-4 reduces the file size as much as possible through codified methods for encoding audio and video. It provides some features, such as DRM (Digital Rights Management) support, various forms of interactivity, including features for 3D rendering with VRML.
* **.AVI:**  An AVI file is a sound and motion picture file, stands for Audio Video Interleaved that is used for video files under Microsoft Windows. The file size, 2 GB, is the maximum file size of an AVI file. AVI files have the extension .avi and need a special player to play the video files that may be require downloading or may be included with your Web browser.
* **.WebM:**  WebM provides open video compression for HTML5 videos that are based on the same video format, which uses the MKV file extension. It is a container format for audio and video data that is sponsored by Google Inc. as a WebM Project. Under a BSD license, it was put together to provide support and release WebM content that led to available in the market to users for free of cost. If any file has the WEBM file extension, it will be a WebM Video file. As the format is used on HTML5 websites for video streaming, therefore, most web browsers support WEBM files.
* **.MOV:**  MOV is a MPEG 4 video container file format introduced by Apple in 1998. It is a format of Apple's Quicktime program. An algorithm, Apple's proprietary compression, is used by MOV files, which can be opened with the help of using a video editor, compatible video program, or QuickTime.
* **.WMV:**  The WMV file extension is a video file format, which stands for Windows Media Video and is developed by Microsoft. It is a compressed video file format that supports different video codecs for streaming video over the Internet. The .WMV is the file extension of the WMV file. It contains video encoded with one of Microsoft's WMV proprietary codec and is the same as an. ASF file. Now WMV files can be played on different player software. In 1999, it was introduced as a competitor of the RealVideo format.
* **.FLV:**  An FLV encoder tool is native to the Adobe Flash player, which is commonly used to change audio and video into the FLV format. The .FLV is the file extension of the FLV file.

1. **Codec –**

Content distributors use a video compression technology called a codec to shrink videos into a streamable size. Codecs allow us to tightly compress bulky streams down for delivery and storage.

Streaming requires the use of both audio and video codecs. H.264, also known as AVC (Advanced Video Coding), is the most common video codec; AAC (Advanced Audio Coding) is the most common audio codec.

Literally ‘coder-decoder’ or ‘compressor-decompressor,’ codecs apply algorithms to the video and create a miniature version of it.

## Video Codecs vs. Containers:

A **codec** acts upon the video, both at the source to compress it and before playback to decompress it. This is done through lossy compression, during which any unnecessary data is discarded.

On the other hand, a **video container format** stores the video codec, audio codec, and metadata such as subtitles or preview images. The container holds all the components together and determines which programs can accept the stream.

## Best Video Codecs for Streaming-

## [H.264/AVC](https://www.wowza.com/blog/video-codecs-encoding#H264)

## [H.265/HEVC](https://www.wowza.com/blog/video-codecs-encoding#H264)

## [AV1](https://www.wowza.com/blog/video-codecs-encoding#AV1)

## [VP9](https://www.wowza.com/blog/video-codecs-encoding#VP9)

## [H.266/VVC](https://www.wowza.com/blog/video-codecs-encoding#H266)

## H.264/AVC

The majority of encoding output today takes the form of [H.264](https://www.wowza.com/blog/h264-codec-advanced-video-coding-avc-explained) files, also referred to as AVC (Advanced Video Coding). This widely supported codec was developed by the International Telecommunications Union and the International Organization for Standardization/International Electrotechnical Commission (ISO/IEC) Moving Picture Experts Group.

H.264 also has significant penetration into markets outside of streaming, such as Blu-ray disks and cable broadcasting. It is often incorporated with the AAC audio codec and can be packaged into .mp4, .mov, .F4v, .3GP, and .ts containers.

## H.265/HEVC

The ISO/IEV Moving Picture Experts Group developed [H.265](https://www.wowza.com/blog/h265-codec-high-efficiency-video-coding-hevc-explained) as the successor to H.264. Also called HEVC (High Efficiency Video Coding), this codec aims to improve compression efficiency and support 8K resolution. It generates smaller files than H.264, thus decreasing the bandwidth required to view these streams. This makes it an ideal codec for high-resolution streaming.

That said, only about 10 percent of encoded files take the form of H.265. [Uncertainties about royalties](https://www.ibc.org/delivery/codec-wars-the-battle-between-hevc-and-av1/2710.article) have stifled adoption. Specifically, content distributors are frustrated by the lacking transparency into what they’ll have to pay when using this codec.

* **AV1**

Developed by Alliance for Open Media, a non-profit organization, AV1 is a non-licensable video codec. Due to this, the codec can be used by anyone without paying royalties to the developers. Founded by Intel, Amazon, Apple, Netflix, and other technology giants, the Alliance of Open Media released AV1 in 2018.

The main goal of the AV1 codec is to reduce file size after compression without affecting quality. This reduction was needed as the display resolution of the content we stream keeps increasing. For example, back in the day, DVDs came with content with a resolution of 480p, and the same could be [compressed using MPEG-2,](https://www.makeuseof.com/tag/how-to-add-mpeg-2-to-your-raspberry-pi-media-centre/) but with [Blue Ray, the compression standard had to improve](https://www.makeuseof.com/physical-vs-digital-movies-blu-rays-downloads/) as the resolution was bumped up to 1080p[.](https://www.makeuseof.com/physical-vs-digital-movies-blu-rays-downloads/)

Evolving video quality demand is what led to the development of h.254 (AVC). With 4K and 8k coming into the picture, new compression standards like AV1 were needed.

The main goal of the AV1 codec is to reduce the video's bitrate while maintaining the quality. Due to this, higher quality videos can be uploaded to the internet without degradation in the quality. In terms of numbers, AV1 offers 30 percent better compression at the same bitrates compared to HEVC. This increased compression efficiency enables smoother playback on lower bandwidths for both high and low resolutions.

## VP9

Google developed [VP9](https://www.wowza.com/blog/vp9-codec-googles-open-source-technology-explained) as a royalty-free, open-source alternative to H.265. The Google-owned YouTube platform and Chrome browser support VP9, as well every Android phone, Mozilla’s Firefox, Apple’s Safari, and all new iOS devices. This codec also makes an appearance in many [WebRTC](https://www.wowza.com/blog/what-is-webrtc) workflows, with [more than 90% of Chrome-encoded WebRTC video using VP9 or its predecessor VP8](https://blog.chromium.org/2020/05/celebrating-10-years-of-webm-and-webrtc.html).

VP9 was released in 2013, which puts it in the middle of the pack as far as age goes. Even so, it’s a better option than most for several reasons. For one, VP9 performs about the same as H.265/HEVC. This makes it well suited for 4K video, especially when publishing to YouTube.

Beyond that, VP9 ranks second only to H.264/AVC in terms of compatibility across browsers and devices. Samsung, Sony, LG, Roku, and many other household names support it. Plus, Google’s implementation of the codec in YouTube and Netflix’s use of it for some content will continue to drive this trend.

## H.266/VVC

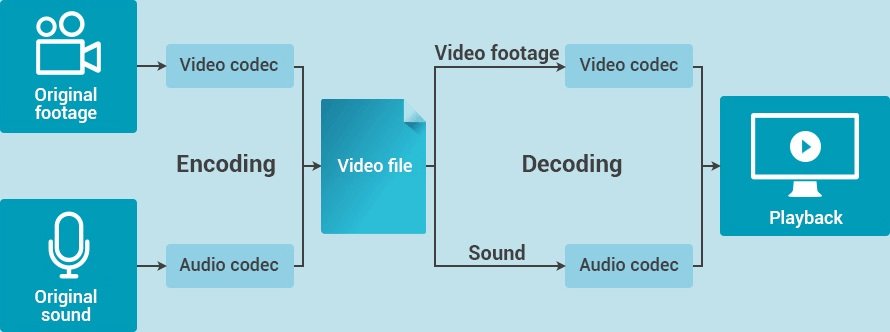
As the newest kid on the video compression block, the [H.266/VVC (Versatile Video Coding) specification](https://www.wowza.com/blog/h266-codec-versatile-video-coding-vvc-explained)was only just finalized in 2020. While intended to usurp H.265 and H.264, it has the same royalty issues as its predecessors.

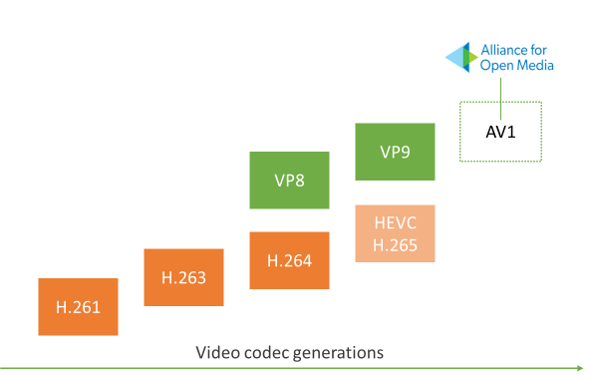
Back in 2008, [Beamr’s chief technology officer Dror Gill explained](https://www.streamingmediaglobal.com/Articles/Editorial/Featured-Articles/Encoding--Transcoding-2018-Part-1-128430.aspx), “It’s okay to pay royalties as long as you know how much you need to pay and when. With H.264, it was very clear how much you need to pay, there was one body collecting all the royalties, and this became the world’s most prominent video codec. The same can happen with VVC, if they get their act together before releasing the standard.”

And yet, the H.266/VVC royalty rate remains a mystery at this time. Unforeseen challenges on the licensing front are up in the air, and we’re also waiting to see how adoption of the codec pans out.

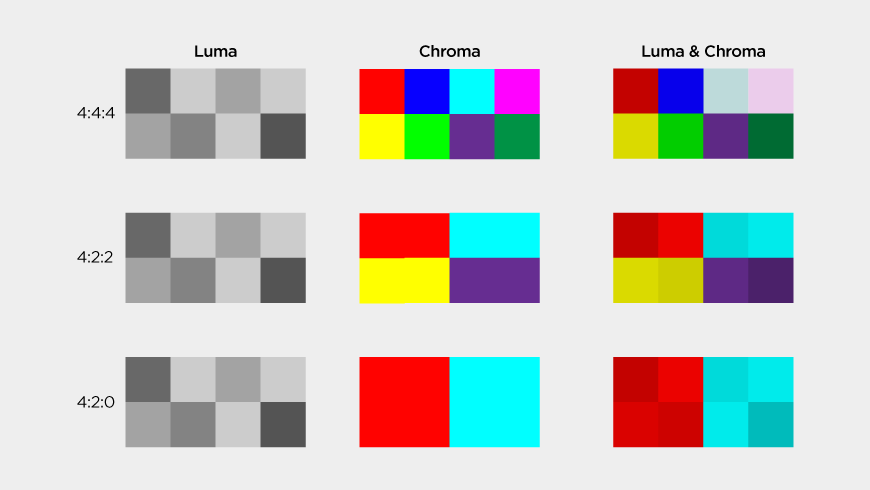
To put it in a nutshell à la codec expert Jan Ozer:

“Overall, the individual H.266/VVC codec developers have made great strides in delivering the promised bandwidth savings, though final performance won’t be known until the royalty policies are set and we know which tools are in which profiles. Beyond that, given the diverse range of other factors in play, it’s impossible to know at this time whether VVC will ever reach critical mass.”





1. **Croma Subsampling -**

****

Chroma subsampling is a type of compression that reduces the color information in a signal in favor of luminance data in order to reduce bandwidth usage without significantly affecting picture quality. Chroma subsampling algorithms are designed to take advantage of the human eye’s higher sensitivity to variations in luminance compared to color. A video signal contains luminance (luma) and chrominance (chroma) components. Luma defines the majority of the picture as contrast is what forms the shapes and details visible, whereas chroma or color information that you see on the screen, although important, has less visual impact. By reducing the amount of color information in a video signal to allow more luminance data instead, chroma subsampling  allows picture clarity to be maintained while effectively reducing the file size by up to 50%.

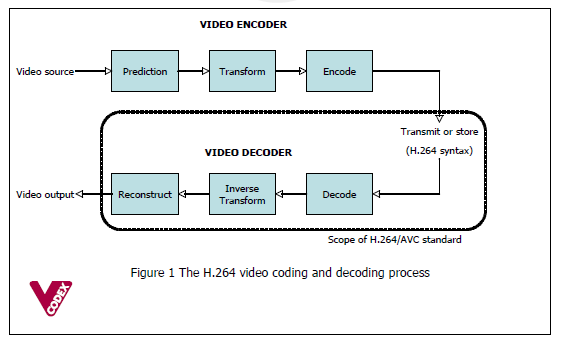
Common chroma subsampling formats include:

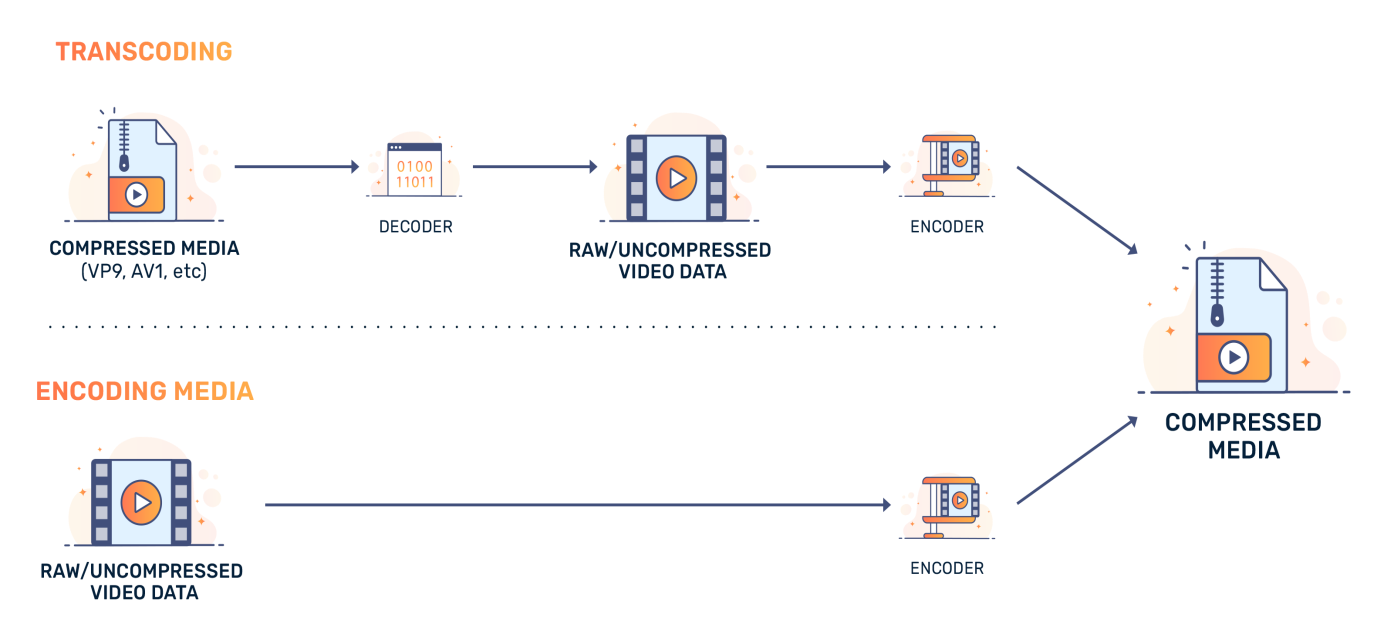
**4:4:4** – uncompressed video with no chroma subsampling, transports both luminance and color data entirely.

**4:2:2**– has half of the chroma of 4:4:4 and reduces the bandwidth of an uncompressed video signal by one-third with little to no visual difference.

**4:2:0** – has one quarter of the chroma of 4:4:4 and reduces bandwidth of an uncompressed video signal by half compared to no chroma subsampling.

1. **Video Encoding-**

****

****

Video encoding is the process of converting digital video files from one standard digital video format into another. The purpose of this is for compatibility and efficiency with a desired set of applications and hardware such as for DVD/Blu-ray, mobile, video streaming or general video editing. The encoding process transforms the video and audio data in the file and then does compression according to the specifications of the encoding standard chosen.

## Image resizing-

A common technique for compression is resizing, or reducing the resolution. This is because the higher the resolution of a video, the more information that is included in each frame. For example, a 1280×720 video has the potential for 921,600 pixels in each frame, assuming it’s an i-frame (more on this in a bit). In contrast, a 640×360 video has the potential for 230,400 pixels per frame.

So one method to reduce the amount of data is to “shrink” the image size and then resample. This will create fewer pixels, reducing the level of detail in the image at the benefit of decreasing the amount of information needed. This concept has become a cornerstone to [adaptive bitrate streaming](https://video.ibm.com/lp/adaptive-streaming-whitepaper?itm_source=blog&itm_medium=onsite&itm_content=what_is_encoding_blog&itm_campaign=ibm_cloud_video). This is the process of having multiple quality levels for a video, and it’s common to note these levels based on the different resolutions that are created.

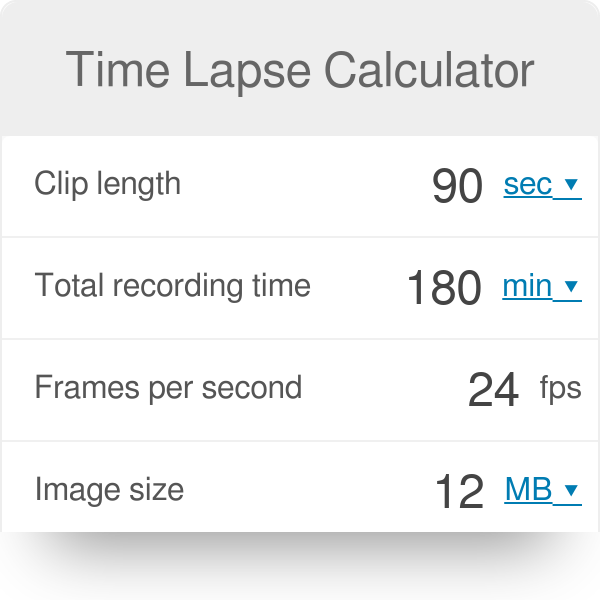
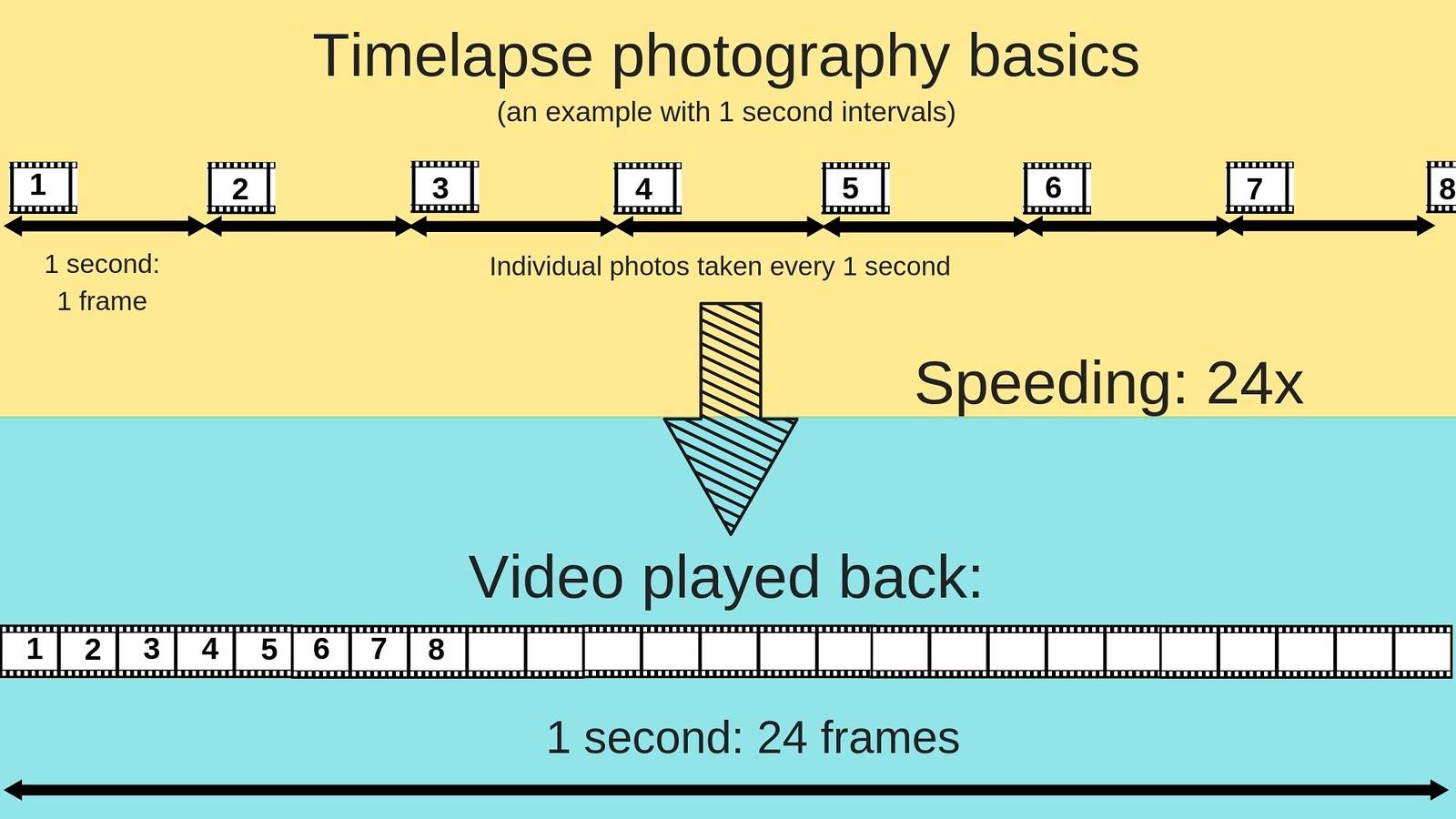
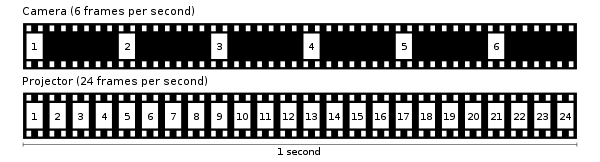
## ****Why is Encoding Important?****

Video encoding is important because it allows us to more easily transmit video content over the internet. In video streaming, encoding is crucial because the compressing of the raw video reduces the bandwidth making it easier to transmit, while still maintaining a good quality of experience for end viewers. If all the video content was not compressed, available bandwidth on the Internet would be inadequate to transmit all of it and prevent us from deploying widespread, distributed video playback services. The fact that we can stream video on multiple devices in our homes, on-the-go using mobile, or even while video chatting with loved ones across the globe, even with low bandwidth, is owed to video encoding.

## ****What is Time-Lapse video?****

A time-lapse video is created by taking photos at regular intervals to record changes that would be too slow for the human eye to see.

A series of still images are captured over a period of hours, days, weeks or even months and then sped up into minutes or seconds in order to show how quickly things change.



## How Time Lapse Video Works?

Time-lapse video is a technique that records the [passage of time](https://filmlifestyle.com/what-is-the-passage-of-time/).

The video captures images of an event at a regular interval and plays them back in rapid succession, usually 24 or 30 frames per second, to create the illusion of fast-moving footage.

It can be used for anything from recording construction projects to documenting the growth cycle of plants.

Ever wondered how time-lapse video works?

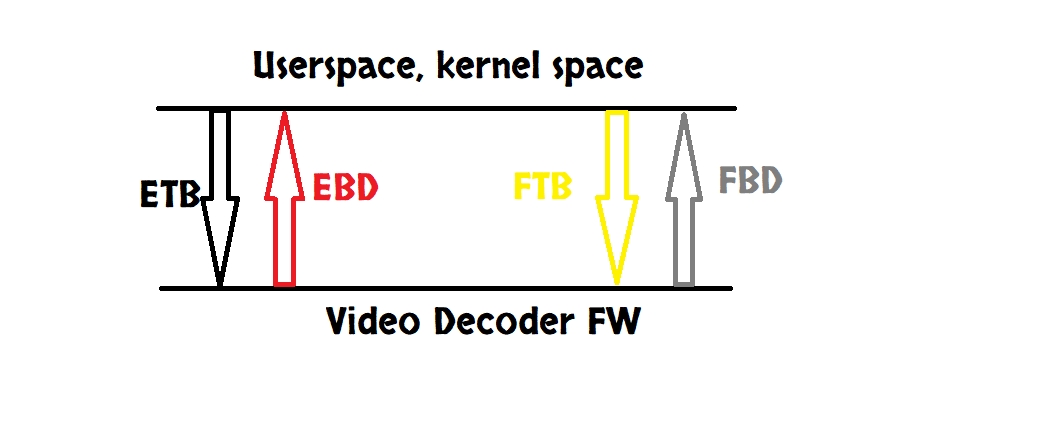
Time-lapse video is a technique that uses the same principle as slow-motion videos.

It captures images at set intervals and then plays them back to create the illusion of time passing quickly.

* **Platform architecture-**



* **Rotating the buffers in circular fashion-**



Input side:

ETB() => Empty This buffer

EBD() => Empty buffer Done

Output side:

FTB() => Fill this Buffer

FBD() => Fill Buffer Done

1. **References-**

Video intro

<https://www.javatpoint.com/what-is-video/>

<https://howvideo.works/>

codec

<https://www.wowza.com/blog/video-codecs-encoding>

<https://www.makeuseof.com/what-is-av1-video-code-what-is-it-for/>

<https://hothardware.com/news/av1-codec-support-and-importance-explained>

sampling-

<https://www.haivision.com/glossary/chroma-subsampling/>

time lapse-

<https://filmlifestyle.com/what-is-time-lapse/>

Device Driver intro:

<https://www.youtube.com/watch?v=fQkObOxVTVo>

Identify V4l2 based devices in ur linux machine:

<https://www.youtube.com/watch?v=_RC1YF2Tv8Y>

V4L2 and basics:

<https://sebastianfricke.me/video-4-linux-the-basics/>