A Guide to Encoding and crack points of Low Latency Streaming:

From Video Encoding Basics to Optimizing Streaming Workflows

Video streaming over the internet is gaining importance in many industries, including broadcast, enterprise, and government. It has become popular for a number of reasons, particularly because live video is a great way to contribute content and engage with consumers, employees, and the community. For broadcast engineers video streaming over the internet is a cost-effective and flexible alternative to satellite services. For AV professionals, video streaming, if correctly implemented, can be an efficient and flexible means of communicating across an organization. Flexibility is not only important for content creation and keeping up with demand, but also for scalability and business continuity.

Video streaming begins with video encoding. For content creators, video encoding can be the most important part of a workflow, which is why it is so important to have a solid grasp of the basics of video encoding before embarking with video streaming.

This guide will explore the principle concepts of video encoding and streaming, including compression, codecs, latency, and network transport considerations when streaming from encoders.

The best way to identify your encoding priorities is to first understand the use case and end goal for your live video stream – what is it needed for, for who, and what will be the measure of its success? By establishing these priorities, one can review the four factors that make up a successful live video streaming experience:

- Quality
- Bandwidth
- Security and Reliability
- Latency

BLANKOM

AN INTRODUCTION TO VIDEO ENCODING

What is Video Encoding?

Video encoding is the process of compressing raw video for transport over IP networks such as office LANs and the internet. As IP networks have limited bandwidth, the encoder needs to be able to compress the content accordingly. There are two types of video encoding: file-based and live, and it's important to make the distinction between them.

When working with video files, encoders are used to compress and reduce the size of video content so that it can take up less storage space and be easier to transfer from one part of a video production workflow to another. Since the video files are not live, latency is usually not a key concern.

Live video encoding is the process of compressing real-time video and audio content prior to streaming. Compression significantly reduces the bandwidth required, making it possible for real-time video to be transmitted across constrained networks while maintaining picture quality at levels suitable for viewing. However, depending on the type of encoder used, compressing live video can also add latency which if too great can negatively impact the overall quality of experience.

Decoding and Transcoding: A Brief Overview

Video decoding is essentially the opposite of encoding. It is the process of decoding or uncompressing encoded video. A decoder can output uncompressed video through SDI for further video processing or over HDMI for displaying directly on a screen.

Decoders can also extract embedded audio tracks for sound production. Embedded metadata can be passed on by the decoder to other production components for information on video formatting, time codes, subtitles, and closed captioning.

Synchronizing Feeds Some decoders, support multiple incoming streams and can resync them based on timecode prior to decoding to SDI. This is especially useful for live broadcasts with multiple Camera angles that share an audio source.

For live video, it is imperative that video decoders add as little latency as possible in order to minimize the impact on production and provide a broadcast quality experience.

Video transcoding is the process of converting an already encoded stream from one format to another, or from one size to another. Most transcoders use a two-step process of decoding and re-encoding. Video transcoding is commonly used for enabling OTT (over the top) internet streaming services with a high quality source or mezzanine video transcoded into a cascade of different bitrates and resolutions. These multiple video transcodes or profiles are needed for ABR (adaptive bitrate) streaming which adapts picture quality in real-time based on available bandwidth. This enables a single video source to be delivered to different viewing devices including connected televisions, computers, and smartphones.

Video Encoders

There are two types of video encoders - software and hardware-based.

Software encoders can be installed on standard off the shelf hardware or as virtual machines (VM) in data centers and cloud platforms. Although software can be a great option for encoding file-based video content, depending on the computer hardware they run on, they don't always offer ultra-low latency levels like dedicated hardware encoders and therefore are not always suitable for live broadcast contribution applications.

Hardware encoders are turnkey devices with dedicated processing power for low latency encoding of video streams. Whereas software encoders have to share CPU and other resources, hardware encoders can use purpose-designed micro-processing chips and can therefore encode and stream live video with very little latency.
Hardware video encoders are used by a wide range of organizations for delivering pristine quality, low latency video for many different applications including:

- Broadcast – for backhaul, bi-directional interviews, return feeds, and remote production (REMI)
- Enterprise – for internet streaming of all-hands meetings, product training, and employee briefs as IPTV, and digital signage
- Defense – for mission critical Intelligence, Surveillance and Reconnaissance (ISR) applications

VIDEO CODECS AND COMPRESSION

What is a Video Codec?

The term codec is a portmanteau of the words ENCoding (coding) and DECoding. It describes a process for compressing and decompressing data as files or real-time streams. For engineers, a codec usually refers to the compression format used by a video encoder, decoder, or transcoder.
Codecs for live video, mainly H.264/AVC or H.265/HEVC, can reduce raw content data by as much as 1,000 times, saving much-needed bandwidth and enabling real-time video streams or files to be easily transmitted across constrained networks and to end-user devices. For example, a typical uncompressed HD stream is about 1.5 Gbps and is compressed to around 5 Mbps for live broadcast television.

Bitrates A bitrate is the number of video data bits that are processed within a unit of time and is commonly measured in bits per second, often abbreviated as bps. Generally, higher bitrates result in higher image quality, but newer and more complex codecs, like HEVC, often deliver better video quality at lower bitrates than older codecs.

VIDEO COMPRESSION TECHNIQUES

The compression techniques used by different codecs can fall into two main categories: lossy and lossless.

Most codecs use “lossy” compression methods which, at a high level, means that when a video is compressed, some redundant spatial and temporal information is lost. “Lossless” compression is used when the goal is to reduce file and stream sizes by only a slight amount in order to keep picture quality identical to the original source.

Lossless compression formats such as JPEG-XS can reduce bit-rates up to 10 times with no loss of picture information and are well-suited for supporting SMPTE 2110 workflows within production studios.

The reduction needed for compression methods can be divided into two main categories: intra-frame and inter-frame. Intra-frame compression or spatial reduction, reduces the size of each individual frame within a video file or stream. Examples of intra-frame formats are JPEG-2000, used for high bitrate broadcast contribution and video archiving, and JPEG-XS used for SMPTE 2110 workflows.

Inter-frame compression or temporal reduction works by grouping multiple frames within a group of pictures or GOP, and encoding only the pixels that change between consecutive frames, based on an initial reference or key frame, known as the I-frame. AVC / H.264 and HEVC / H.265 are the most common inter-frame compression formats as they offer significant bitrate reductions suitable for streaming SD, HD and 4K video content.

Framing Options

Choosing the right combination and number of I, P, and B-frames is key to optimizing video quality. The main reference comes from the I-frames, which contain the most amount of data. P-frames only contain the differences between it and the previous I-frame, and B-frames contain both forward and backward changes resulting in even more efficient compression.

Here are some examples of framing options available with high end encoders:

I: I-frames only (highest quality, least amount of bandwidth efficiency) IP: I and P-frames only (high picture quality, efficient compression) IBP: I, B and P-frames

IBBP: I, BB (two B-frames in sequence) and P-frames

IBBBP: I, BBB (three B-frames in sequence) and P-frames

IBBBBP: I, BBBBB (four B-frames in sequence) and P-frames (highest latency; highest bitrate efficiency)

Some more basics & details to perfect balancing your encodings:

There are many critical elements to producing a successful high-quality video production, such as selecting different camera lenses based on the available focal lengths, exposure (*see below*), frame rate setting (*see below*), and color -> *See below details*. In addition to these elements, which we have introduced before, codec, container format, and bitrate are the last few factors determining your video output quality. And in this article, we are going to explain the importance of codec, container format, and video file format.

What is a Codec?

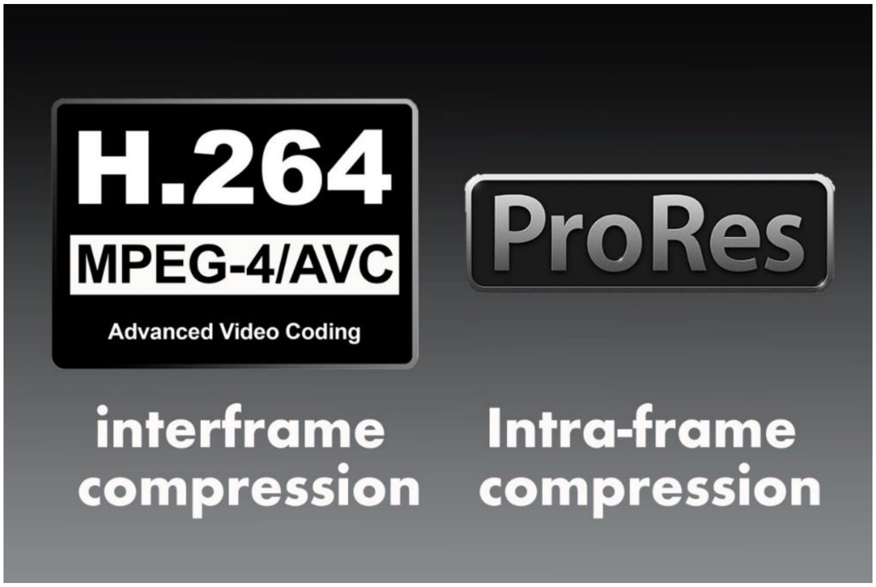
A codec, an essential element of video compression, is a device or computer program that encodes or decodes a data stream or signal. Codec is a portmanteau of coder/decoder. A coder or encoder encodes a data stream or a signal for transmission or storage, possibly in encrypted form, and the decoder function reverses the encoding for playback or editing. Hence, codec will be used in video recording, post-editing, and file export in video production.

It is important to compress the video captured by a camera to save storage space. Let's review the video file size at 1080p30 as an example. Given that the resolution is 1920 x 1080 so each frame contains about 2,073,600 pixels. In an RGB color space, each pixel is represented by 8 bits so the file size of each frame is (2,073,600 pixels x 8 bits) / (8 x 1024) = 2025 kilobytes. Because there are 30 frames in one second, the file size of a one-second uncompressed video is (2025x30) / 1024 = 59.33 megabytes. For a 10-second uncompressed video, the file size can be as large as 593.3 MB. As a result, long uncompressed videos can consume substantial storage on the PC or the video streaming platform and are thus impractical.

Many manufacturers have developed their codec algorithms for playing a specific video file. For example, H.264 and Apple ProRes are standard codecs. There are two types of image compression algorithms: inter-frame compression and intra-frame compression. The so-called inter-frame compression means that when encoder only records the different parts and the information in the changing process of the two frames. Inter-frame compression does not have to record the same segments of the two frames repeatedly. The inter-frame compression can significantly reduce the size of the video; however, it consumes more processing power of the computer when playing video. H.264 codec adopts an inter-frame compression algorithm, which is also widely applied in network streaming media, such as Vimeo, YouTube, and iTunes Store.

The intra-frame compression adopts a lossy compression algorithm to compress each frame without reference to any others. Each frame is encoded as a separate image. The intra-frame compression reduces the loading of computer processing, but the file requires a significant storage space. Apple ProRes is an intra-frame compression algorithm.

Suppose the video content is a static topic, such as an interview program or a talk show. In that case, the inter-frame compression can compress the video file without the expense of the image quality. For the action theme content like sports games, intra-frame compression is the algorithm to retain better image quality and reduce the computer processing load.



What is a Container Format?

It's essential to distinguish codecs from container formats, though sometimes they share the same name. Briefly, container formats, or wrappers, are file formats that can contain specific types of data, including audio, video, closed captioning text, and associated metadata. A container also refers to a file extension since the term is often noted at the end of file names. Popular video containers include .mp4, .mov, or .avi which are frequently used on various major live streaming platforms such as YouTube and Facebook. MP4 is the most common type of video file format as it works across multiple platforms, operating systems, and devices.

Container Format	Name	Developer	Feature
avi	Audio Video Interleave	Microsoft	An old version format. It cannot contain videos using modern compression algorithm.
wmv	Windows Media Video	Microsoft	
mov	Movie Digital Video Technology	Apple	High compatibility. It can be played by QuickTime on Windows and MAC systems.
mp4	MPEG-4 Part 14	International Organization for Standardization	The most widely accepted container, MP4 can work across multiple platforms, operating systems and devices.

CURRENT AND UPCOMING VIDEO CODECS

MPEG-2 MPEG-2 has been around since the 1990s and introduced the world to digital television and DVDs. It is slowly being phased out, but used in many legacy applications and terrestrial broadcast systems.	**JPEG2000** JPEG2000 is a codec used for digital cinema, medical imaging, geospatial data, and document archiving.
• High quality Not efficient	• Intra-frame encoding for high image quality Requires a lot of bandwidth and storage space
H.264 / AVC (Advanced Video Coding) H.264 makes up the majority of multimedia traffic; it is used for high quality streaming and HD television. However, as the demand for 4K continues, it is being replaced with HEVC.	**H.265 / HEVC (High Efficiency Video Coding)** The successor to H.264, HEVC or H.265, is fast becoming ubiquitous thanks to the proliferation of 4K content.
• Fast encoding speed • Efficient for HD video Not efficient enough for 4K UHD	Highly efficient – can deliver same quality as H.264 at half the bitrate • Excellent quality – ideal for 4K UHD video Requires significant processing power – not always suitable for software encoders
VP9 VP9 is a royalty-free and open source codec developed by Google and used by YouTube, but with few features, is rarely used in commercial or broadcast applications.	**JPEG-XS** JPEG-XS is a lightweight “lossless” compression standard which is useful for 4K and 8K contribution workflows over existing 3G SDI and 12G SDI or hybrid SDI/IP networks including SMPTE-2110.
• Open source Few benefits compared to codecs like HEVC Limited availability of real-time VP9 encoders	Ideal for sending 4K video over a 10GB network to devices that aren’t equipped for 12GB Limited configuration settings
AV1 AV1 is a newer open and royalty-free standard. However, it is still very much a work in progress.	**VVC (Versatile Video Coding)** VVC is a next-generation compression standard, in its early development phase. First hardware implementations are slated for later in 2021.
• Open source and evolving Requires significantly more computing power than HEVC A lack of details has caused hesitation on its adoption	More efficient that HEVC – it is focused on achieving 30% better compression efficiency Still in development and will not be tested in real-world implementation for a while

THE MUST-HAVE CODECS – HEVC AND H.264

In most use cases, broadcast engineers and AV professionals need to be able to support both H.264 and HEVC depending on the source, video format, and devices they are streaming to. H.264 is often required for SD and HD content streamed to digital broadcast systems, displays, and set top boxes, while HEVC is rapidly being deployed by organizations with a focus on live video workflows and 4K content.

Today's live video encoders need to support either H.264, HEVC, or both, as they are the most widely used and fit for purpose codecs for streaming over IP networks. Encoding in H.264 or HEVC also ensures interoperability with all types of decoders, viewing devices, and web browsers.

HEVC (H.265) vs. AVC (H.264): What's the Difference?

AVC (H.264)?
H.264 (also called AVC, or Advanced Video Coding) is an industry standard for video compression that allows for the recording, compression, and distribution of digital video content.
What is HEVC (H.265)?
H.265 is newer and more advanced than H.264 in several ways. H.265 (also called HEVC, or High Efficiency Video Coding) allows for further reduced file size, and therefore reduced required bandwidth, of your live video streams.

H.265 compared to H.264

The H.265 codec compresses information more efficiently than H.264, resulting in files of comparable video quality that are about half the size. The benefits of this are twofold: H.265 video files don't take up as much storage space, and they require less bandwidth to stream. This is a big advantage especially when it comes to storing and streaming 4K video and other high-resolution video content.
What accounts for this difference is how each video compression standard processes frames. H.264 uses what are called macroblocks, processing units that span 4×4 to 16×16 pixels. H.265 uses a newer block structure called coding tree units (CTUs), which can process sizes of up to 64×64 pixels.
There are other technological enhancements at work, such as superior motion compensation and spatial prediction. But the shift from macroblocks to CTUs is the most significant contributor to H.265's greater efficiency.

Required bandwidth for broadcasting in 4K:

mbps: Megabits per second

Because H.265 compresses your data so much more efficiently, using it as your video compression tool will drop your bandwidth and storage requirements by roughly 35-50%. The table below compares the recommended bandwidth for H.264 vs. H.265 encoding.

Recommended bandwidth for video encoding:

Resolution	Minimum Upload Speed*	
	H.264	H.265
480p	1.5 mbps	0.75 mbps
720p	3 mbps	1.5 mbps
1080p	6 mbps	3 mbps
4K	32 mbps	15 mbps

*These values are rough estimates based on stable network environments, calculating upload requirements is very subjective and depends on a number of factors.
mbps: Megabits per second

H.264 vs. H.265: Which is better?

The whole point of H.265 is to succeed H.264. So why isn't everyone using it?
What H.265 offers in efficiency it demands in processing power. Advanced hardware is needed not only to create H.265 video files but also to decode them for playback. This limits who can benefit from H.265's superior efficiency to those with the right gear. It's why H.264 is still the go-to codec for many.
That said, there are plenty of current-day video applications where H.265 encoding is an asset, particularly those involving 4K. Plus, the number of content consumers with H.265-capable hardware continues to grow. It's only a matter of time before H.265 fulfills its purpose and succeeds H.264.

About the Bitrate

Mar 28 2022

Once you are finished editing your video, the next phase in video production is to encode your video. Video encoding plays a critical role in ensuring streaming quality. In network streaming, H.264 is the most widely used codec, with the generated video packaged into MP4 format. A critical factor determining image quality is the "bitrate" which we would like to talk about in this article.

What is a bitrate?

A bitrate is also known as bit rate or compression rate. A video bitrate is the number of bits that are processed in a unit of time. An encoder applies a higher bitrate per unit time; the lower compression ratio used for the video improves image quality. On the contrary, more compressed files have lower bitrates, resulting in an awful viewing experience. Thus, the bitrate is proportional to the image quality, but the video file's size also increases as the bitrate increases.

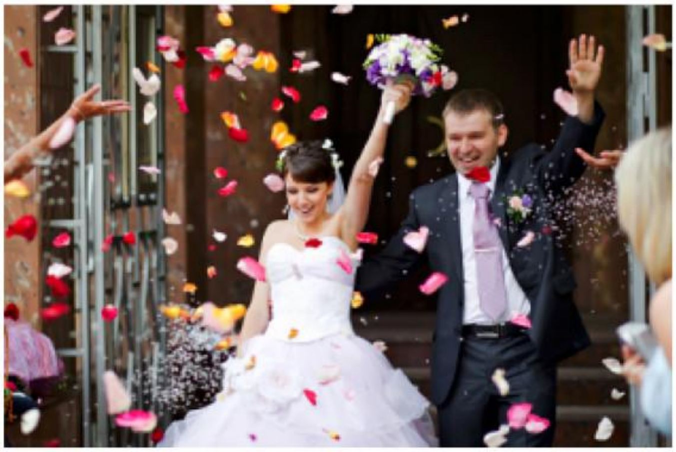
We can calculate the video file's size by the given bitrate. For example, if three-minute footage is recorded at 6000 kbits per second (or 750KBytes per second), the bitrate in KBytes per second (750KBytes per second) is multiplied by the footage duration (180 seconds) to yield 135,000KB which equals 131.8 MB (135000KB/ 1024).

Note: Video bitrate is often measured in kilobits per second (kbps) or Megabits per second (MBps). A byte is an uppercase B, and a bit is a lowercase b. One byte consists of 8 bits.

Video encoding attempts to balance compression rate and image quality to create a smaller file size and higher-quality video. Applying high bitrates is not always the best choice because it is difficult for the human eye to detect the slight difference when the extremely high resolution is beyond visual perception.

Since the lower video bitrate requires less network bandwidth, thus a user can lower the video bitrate to ensure the stable and no buffering video playing where the network service is poor.

HD 1920 x 1080 30fps
@6000 kbps
131.8MB



HD 1920 x 1080 30fps
@300 kbps
6.6MB File Size

Constant Bitrate (CBR) and Variable Bitrate (VBR)

The Common bitrate settings include CBR (Constant Bitrate) and VBR (Variable Bitrate). The variable bitrate means that the encoder provides more bitrate to compress the more complex segments of frames while less space is allocated to less complicated details. The VBR is suitable for most productions, especially if there are a lot of randomly moving objects in the video, such as raindrops or heavy snow. The CBR applies the constant bitrate to compress the video by the given file size. CBR is appropriate for video production with little or no variability in the background, such as the conference and interview videos. CBR ensures a constant flow of data so that the video will not freeze or buzz while streaming video to the platform.

Bitrate and Network Speed

Most live-streaming platforms, such as YouTube, Vimeo, Facebook, etc., compress the uploaded video. The platform initially compresses video at a low bitrate so that the video is available to stream on a wide variety of devices.

Take uploading a 1080p video on the YouTube platform as an example. YouTube compresses the video bitrate to 6000Kbs. The video can be played smoothly as long as the network speed exceeds 6Mb. On the contrary, if the video is not compressed, assuming that the original bitrate is 20Mb, the network speed must be above 20Mb to support smoothly playing video.

Recommended bitrate settings provided by YouTube are shown in the below table. Please note that these settings may vary on other platforms.

	Acceptable Bite Rate on YouTube	
Resolution	**30fps**	**60fps**
4K/2160p	13,000 - 34,000 Kbps	20,000 - 51,000 Kbps
1440p	6,000 - 13,000 Kbps	9,000 - 18,000 Kbps
1080p	3,000 - 6,000 Kbps	4,500 - 9,000 Kbps
720p	1,500 - 4,000 Kbps	2,250 - 6,000 Kbps

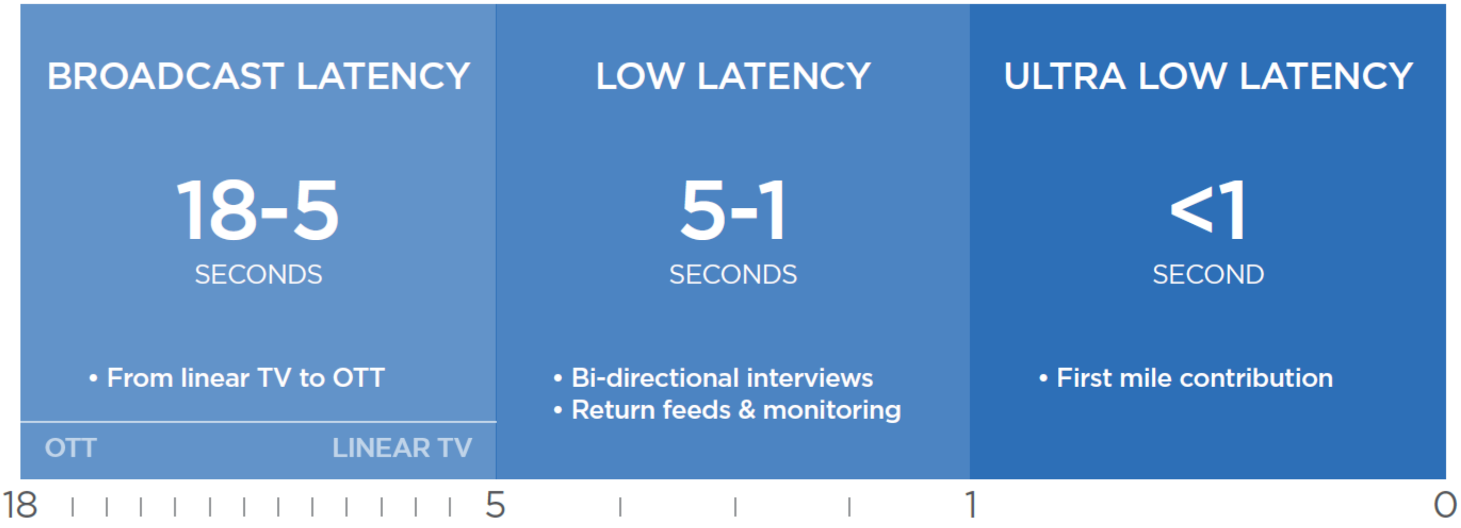
LATENCY

What is Video Latency?

Within the context of video streaming, latency is the amount of time it takes for video to travel from its initial source to its final destination. The complete journey from camera to screen is often referred to as glass-to-glass or end-to-end latency. Encoder latency describes how long it takes for a raw video stream to be processed and encoded prior to streaming, while network latency is the amount of delay on a given IP network. As multiple encoders, transcoders, decoders, and other video processing components may be sharing more than one type of network, these all add up to the total end-to-end latency.

Low end-to-end latency, at under 5 seconds, is critical for broadcasting live events, whereas OTT services are typically delivered with up to 15 seconds of latency. However, for specific tasks such as broadcast contribution, bidirectional interviews, or reconnaissance missions, latency levels need to be much lower or well under 1 second to ensure minimal impact on overall end to end latency.

LIVE VIDEO STREAMING LATENCY



Live video streaming latency

Sources of Latency

There are several factors which contribute to end-to-end latency including the complexity of the content delivery chain and the number of video processing steps involved. While individually these delays might be minimal, cumulatively, they can add a disruptive delay that compromises the viewing experience.

Key contributors to video latency include:

• Individual components in the streaming workflow
Each component in a streaming workflow, including cameras, encoders, decoders and display devices, create delays which contribute to latency in varying degrees. Latency continues to rise with each additional component used to process the video content prior to delivery.

• Network type and speed
The network chosen to transmit a video (ex: public internet, satellite link, or MPLS network) impacts both latency and quality. Network speed is generally defined by throughput or how many megabits or gigabits can it handle in the course of a second and also by the distance traveled.

• Streaming protocols and output formats
The choice of video protocol also impacts video latency. A transport protocol like SRT is low latency while an ABR protocol like HLS is high latency. In addition, the type of error correction used by the selected protocol to counter packet loss and jitter can also add to latency.

Ultra-Low Latency

Broadcast engineers work to keep overall latency as low as possible from the start – keeping the delay from the camera to the production studio at less than 1 second (or ideally under 300 milliseconds) referred to as ultra-low latency.

LIVE LATENCY

For live video, latency should be kept as low as possible. This helps to ensure an optimal viewing experience across a variety of content; low latency prevents news and sports broadcasts from being "spoiled" by social media, and eliminates any awkward pauses that can otherwise arise in bi-directional interviews.

However, a degree of latency can be useful, even for live applications. Introducing delay between broadcast production and playout can also facilitate live subtitling, closed captioning, and prevent obscenities from airing.

When Milliseconds Matter

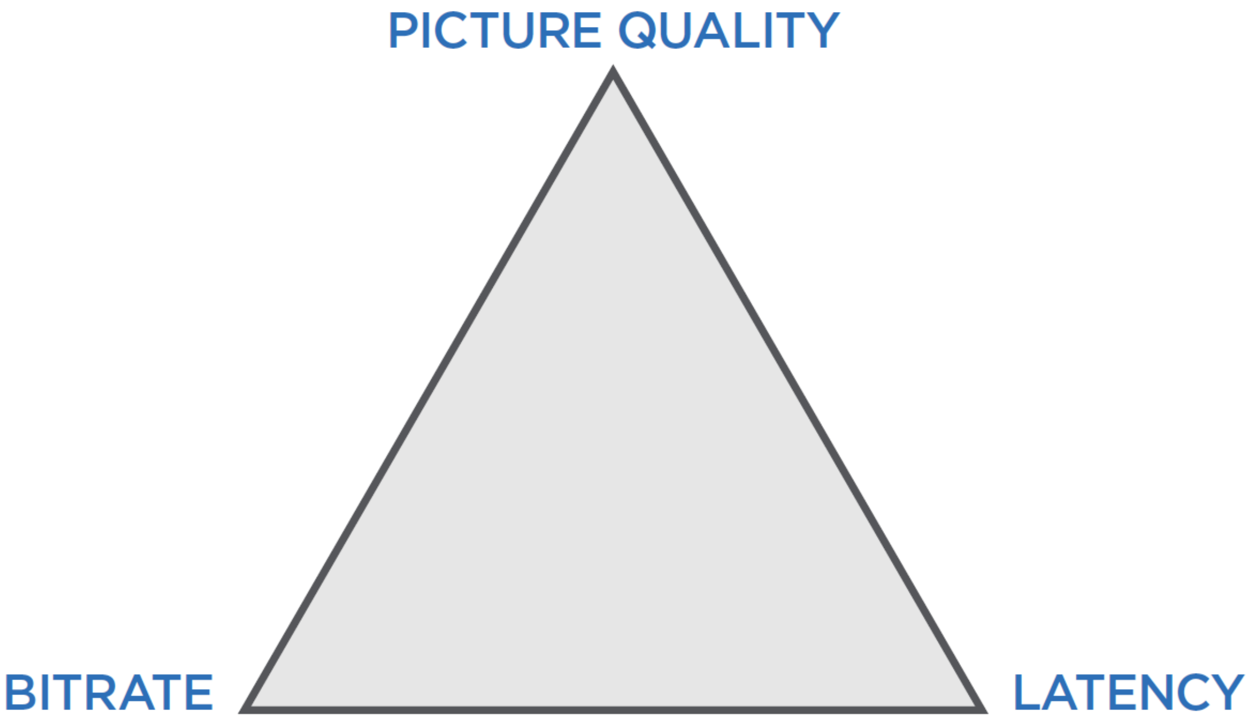
While video latency presents a serious annoyance to sports fans, low latency is also a critical requirement for surveillance missions, online auctions, and gaming applications.

How to Reduce Latency

There are a few ways to minimize video latency without having to overly compromise on picture quality.

The first is to choose a hardware encoder and decoder combo engineered to keep latency as low as possible, even when using a standard internet connection. The latest generation of video encoders and video decoders are able to maintain low latency (under 50ms in some cases) and have enough processing power to apply the HEVC codec to compress live video to extremely low bitrates (down to under 3 Mbps for HD content) all while maintaining high picture quality.

Balancing Latency with Picture Quality and Bandwidth Availability



Balancing Bitrate, Latency, and Picture Quality

These three factors need to be taken into consideration when encoding and streaming live content.

Any video encoder used for broadcast quality live video streaming should allow users to change bitrate, picture quality, and latency settings. Ultimately, the individual targeted use case will determine the best balance within this triangle of video encoding and streaming considerations.

Latency is not the only consideration in a video streaming workflow; video quality and bandwidth are also very important across each application. The three operate in a balance: when the latency is reduced, either the picture quality will be reduced, or the bitrate will increase.

Network Adaptive Encoding (NAE)

Some encoders feature Network Adaptive Encoding. NAE enables encoders to automatically adjust compression levels based on real-time network bandwidth information.

For applications where latency is critical such as video surveillance and ISR, picture quality can often be exchanged in favor of minimizing latency. However, for use cases where pristine broadcast quality video matters, latency can be increased slightly in order to support advanced video processing and error correction. By delivering the optimal combination of bandwidth efficiency, high picture quality, and low latency, viewers can enjoy a great live experience over any network.

Another important factor in keeping latency low is between the encoder and decoder – video transport, which is the subject of the next chapter.

VIDEO STREAMING

While streaming takes place after the video encoding process, it is important to consider the different applications and the best transport protocol for the job.

What Is a Transport Protocol and when is it Used?

A transport protocol is a communication protocol responsible for establishing a connection between two or more devices and delivering data across a network. Transport protocols provide data delivery guarantees that are essential for file transfers and mission-critical applications. Especially important for video streaming, different transport protocols may support a range of optional capabilities including error recovery, flow control, and support for re-transmission.

Important Factors in Choosing a Transport Protocol

- Latency: Different protocols will introduce different amounts of latency to your workflow; live applications will require a transport protocol with as low latency as possible.
- Security: Regardless if the content is for broadcast, government, or enterprise purposes, security is paramount. Video streams being sent over the public internet will need a transport protocol that enables them to be encrypted.
- Reliability: Networks like the public internet can be unreliable, which can result in the loss of data in transport ("packet loss"). This packet loss can degrade the quality of the video, unless the transport protocol in question has a means of error correction, recovering these lost packets.
- Flexibility: Some video streams need to be sent across a room; others across an ocean. Depending on how far a video must travel, it may have to cross multiple firewalls, necessitating a transport protocol that can make firewall Traversal easy.

Error Correction: FEC vs ARQ

FEC (forward error correction) and ARQ (automatic repeat request) are the most common methods of packet loss recovery in video streaming. In FEC, the encoder sends a duplicate of each packet, in case one is lost in transport. With ARQ, the encoder labels each packet, allowing the decoder to request a packet be re-sent should it be lost in transport.

TRANSPORT LAYER PROTOCOLS

The two fundamental transport protocols used by the internet for sending and receiving data include UDP and TCP. UDP is typically used for live streams while TCP is used for file-based content. Other transport protocols such as RTMP or SRT are higher layer protocols that are based on UDP and TCP.

TCP (Transmission Control Protocol,) is the most commonly used transport protocol on the internet. TCP is considered to be very reliable as packets are tracked in such a way that no data is lost or corrupted in transit. In spite of its apparent reliability, TCP is unsuitable for live video; packets are often delayed due to error recovery and they are not always received in the right order.

UDP (User Datagram Protocol,) works similarly to TCP, but is far faster, as it uses different error recovery methods. When video streaming with UDP, “datagrams” (essentially packets) are received by the decoder in the order they were sent. However, unless the application layer protocol being used has its own error recovery method, there is no recourse if packets are missing. UDP is very fast and therefore frequently used for time-sensitive applications such as online gaming or live broadcasts where perceived latency is more critical than packet loss.

COMMON TRANSPORT LAYER PROTOCOLS

RTMP (Real-Time Messaging Protocol)

RTMP is a legacy, but still commonly used protocol for live streaming within production workflows. Based on TCP, RTMP is a continuous streaming technology with packet loss recovery, though at the cost of increased latency. Because RTMP does not support HEVC, it is being phased out by broadcasters and CDNs.

SRT (Secure Reliable Transport)

SRT is an open-source streaming protocol that enables AES 128/256 bit encryption to keep streams secure. It utilizes ARQ packet recovery to maintain high quality over unreliable networks without compromising latency, but will also support FEC packet recovery for those who prefer the latter. In 2019, the protocol was awarded an Emmy for Engineering and Technology. Usually not in use in Headend-Environment’s.

RTP (Real-Time Transport Protocol)

RTP is an internet protocol for real-time transmission of multimedia data in unicast or multicast mode. RTP runs over UDP for low latency and though it does not include packet-loss recovery it has mechanisms to compensate for any minor loss of data when used in conjunction with the RTP Control Protocol (RTCP) for monitoring quality of service.

RTSP (Real-Time Streaming Protocol)

RTSP allows viewers to remotely pause, play, and stop video streams via the internet without the need for local downloads. This application layer protocol was most notably used by Real Networks RealPlayer and is still being applied for various uses including for remote camera streams, online education and internet radio. RTSP requires a dedicated server for streaming and does not support content encryption or the retransmission of lost packets as it relies on the RTP protocol in conjunction with RTCP for media stream delivery.

Delivery Protocols

HLS and MPEG-DASH are two commonly used streaming protocols for video content delivery. Both of these protocols support adaptive bitrate streaming (ABR) and are therefore ideal for delivering produced content to viewing devices over an internet connection. However their high level of latency make them unsuitable for video contribution.

Proprietary Protocols

There are a number of proprietary protocols based on TCP and UDP, which are designed for video delivery, however they do require a license which can add costs and also prevent interoperability with third-party vendor equipment.

CHECKLIST: CHOOSING A VIDEO ENCODER

There are several important factors involved in choosing a video encoder. For those who are new to video streaming, it can feel a little overwhelming, while individuals who are accustomed to certain systems may be nervous about switching components.
Regardless of the workflow, there are important considerations to bear in mind when selecting a video encoder, as we outline in our checklist.

Use Case

The single most important question to ask when looking for a video encoder. By establishing the purpose of the streamed video content, you can prioritize the elements that a video encoder can bring to your workflow.

Form Factor

The form factor of a video encoder is very important to a streaming setup, whether you have the space for a full hardware video encoder, or simply a blade. Some situations require ruggedized video encoders for harsh environments, especially when operating outdoors.

Codecs

It is important that the video encoder you are using is capable of supporting the video codec you want to use, in most cases HEVC or H.264. Some video encoders support both of these codecs, which is an excellent way of future- proofing your streaming setup.

Transport Protocol

Choose a video encoder that supports the video transport protocols best suited to your streaming application; for streaming encrypted video at low latency for broadcast contribution or live event production, consider an encoder that supports SRT.

Latency

To keep latency low in a video streaming workflow, you need to start from the beginning. If the video encoder is adding latency, there won’t be a way to “catch up” on that delay later in the streaming process. Therefore, choosing a video encoder with low latency is crucial for minimizing total delay.

Interoperability

A video encoder is part of a larger workflow or setup; it is important that it fits with the other pieces. Ensure that your video encoder is compatible with the other elements of your workflow including your camera and decoder.

Ease of Use and Support Availability

This is more a function of who is operating the video encoder. While many broadcast engineers will be familiar with a variety of encoders and hardware, IT managers at corporate offices or volunteers at a religious or community organization may need a more user-friendly device with support available to them.

Reliability

Depending on the situation, some video encoders may have to contend with unpredictable networks. In order to ensure the reliability of the video stream, some video encoders have options like adaptive bitrate encoding to ensure that the best quality video possible is always available.

Quality

The video quality needed will greatly influence the kind of encoder required. There are video encoders that can stream in HD, and 4K UHD – although not all workflows prioritize that level of quality.

Security

Keeping video streams secure helps protect intellectual property and prevents unauthorized access to videos. Choosing a video encoder with security options like AES 128/265-bit encryption will ensure that streams are safe from the start.

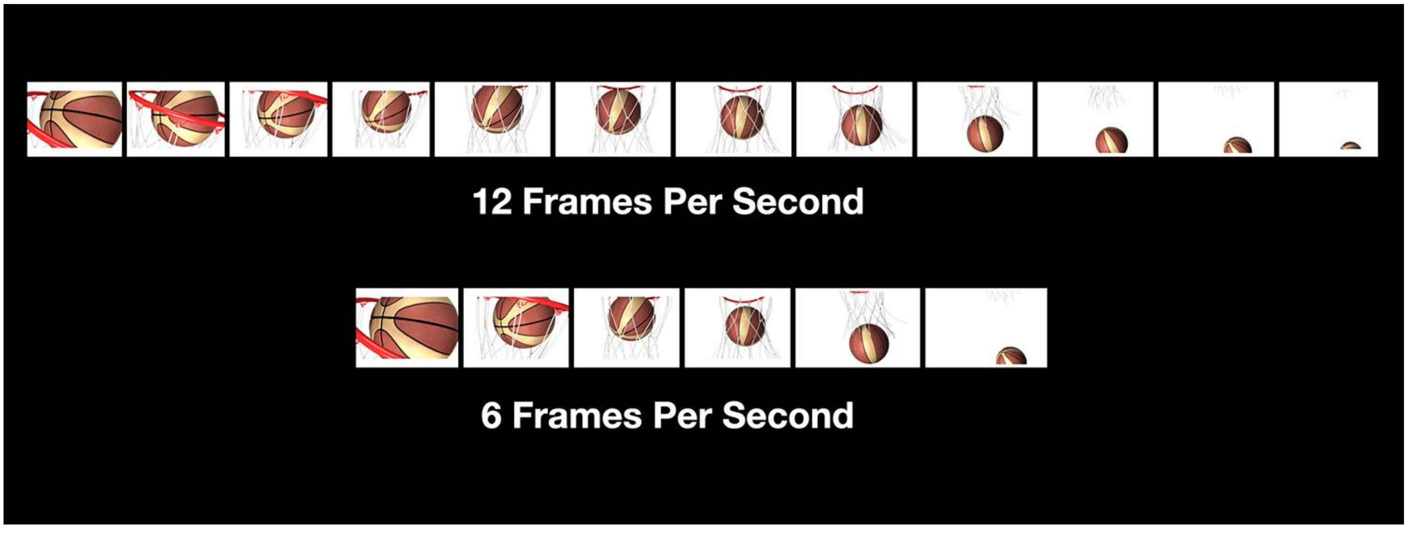
What is Frame Rate and How to Set the FPS for Your Video

Mar 09 2021

What is Frame Rate?

One of the essentials you should know is the "Frame Rate" to learn the process of video production. Before talking about the frame rate, we must first understand the principle of animation (video) presentation. The videos we watch are formed by a series of still images. Since the difference between each still image is very small, when those images viewed at a certain speed, the fast-flashing still images give the appearance on the human eye's retina which results the video we watch. And each of those images is called a "frame."

"Frame Per Second" or the so-called "fps" means how many still images frames in the per-second video. For example, 60fps implies that it contains 60 frames of still images per second. According to the research, the human visual system can process 10 to 12 still images per second, while more frames per second are perceived as motion. When the frame rate is higher than 60fps, it is hard for the human visual system to notice the slight difference in the motion image. Nowadays, most movie production applies 24fps.



What are the differences of NTSC System and the PAL System?

When television comes to the world, the television also changed the video frame rate format. Since the monitor presents images by lighting, the frame rate per second is defined by how many images can be scanned within one second. There are two ways of image scanning-"Progressive Scanning" and "Interlaced Scanning."

Progressive scanning is also referred to as noninterlaced scanning, and it is a format of displaying in which all the lines of each frame are drawn in sequence. The application of interlaced scanning is due to the limitation of the signal bandwidth. The interlaced video applies the traditional analog television systems. It has to scan the odd-numbered lines of the image field first and then to the even-numbered lines of the image field. By quickly changing the two "half-frame" images make it look like a complete image.

According to the above theory, the "p" means Progressive Scanning, and the "i" represents Interlaced Scanning. The "1080p 30" means Full HD resolution (1920x1080), which is formed by 30 "full frames" progressive scan per second. And "1080i 60" means the Full HD image is formed by 60 "half-frames" interlaced scan per second.

To avoid the interference and noise generated by current and TV signals at different frequencies, the National Television System Committee (NTSC) in the USA has developed the interlaced scanning frequency to be 60Hz, which is the same as the alternating current (AC) frequency. This is how the 30fps and 60fps frame rates are generated. The NTSC system applies to the USA and Canada, Japan, Korea, the Philippines, and Taiwan.

If you are careful, do you ever notice some video devices note 29.97 and 59.94 fps on the specs? The odd numbers are because when the color TV was invented, the color signal was added to the video signal. However, the frequency of the color signal overlaps with the audio signal. To prevent the interference between video and audio signals, American engineers low 0.1% of the 30fps. Thus, the color TV frame rate was modified from 30fps to 29.97fps, and the 60fps was modified to 59.94fps.

Compare to the NTSC system, the German TV manufacturer Telefunken has developed the PAL system. The PAL system adopts 25fps and 50fps because the AC frequency is 50 Hertz (Hz). And many European countries (except France), the Middle East countries, and China apply the PAL system.

Today, the broadcast industry applies 25fps (PAL system) and 30fps (NTSC system) as the frame rate for video production. Since the frequency of AC power is different by region and country, so be sure to set the right corresponding system before shooting the video. Shoot video with the wrong system, for example, if you shoot the video with the PAL system frame rate in North America, you will find that the image flicking.

The Shutter and the Frame Rate

The frame rate is highly associated with the shutter speed. The "Shutter Speed" should be double the Frame Rate, resulting in the best visual perception to human eyes. For instance, when the video applies 30fps, it suggests that the camera's shutter speed is set at 1/60 seconds. If the camera can shoot at 60fps, the shutter speed of the camera should be 1/125 second.

When the shutter speed is too slow to the frame rate, for example, if the shutter speed is set at 1/10 second to shoot the 30fps video, the viewer will see blurred movement in the video. On the contrary, if the shutter speed is too high to the frame rate, for example, if the shutter speed is set at 1/120 second for shooting 30fps video, the movement of objects will look like robots as if they were recorded in stop motion.

Shutter Speed = 1/ (Frame x 2)

How to Use the Suitable Frame Rate

The frame rate of a video dramatically impacts how the footage looks, which determines how realistic the video appears. If the video production subject is a static subject, such as a seminar program, lecture recording, and video conference, it is more than enough to shoot video with 30fps. The 30fps video presents the natural motion as the human visual experience.

If you want the video to have a clear image while playing in slow motion, you can shoot the video with 60fps. Many professional videographers use the high frame rate to shoot video and apply lower fps in post-production to produce the slow-motion video. The above application is one of the common approaches to create an aesthetically romantic atmosphere through slow-motion video.

If you want to freeze the objects in high-speed motion, you have to shoot a video with 120fps. Take the movie "Billy Lynn in the Middle" for example. The movie was filmed by 4K 120fps. The high-resolution video can vividly present the very details of images, such as the dust and splattering of debris in the gunfire, and the spark of fireworks, providing the audience an impressive visual perception as if they were personally on the scene.

Finally, we would like to remind readers must use the same frame rate to shoot videos in the same project. The technical team must check that every camera applies the same frame rate while performing the EFP workflow. If Camera A applies 30fps, but Camera B applies 60fps, then the intelligent audience will notice the video's motion is not consistent.

So, we need to talk about COLOURs:

What are 8-bit, 10-bit, 12-bit, 4:4:4, 4:2:2 and 4:2:0

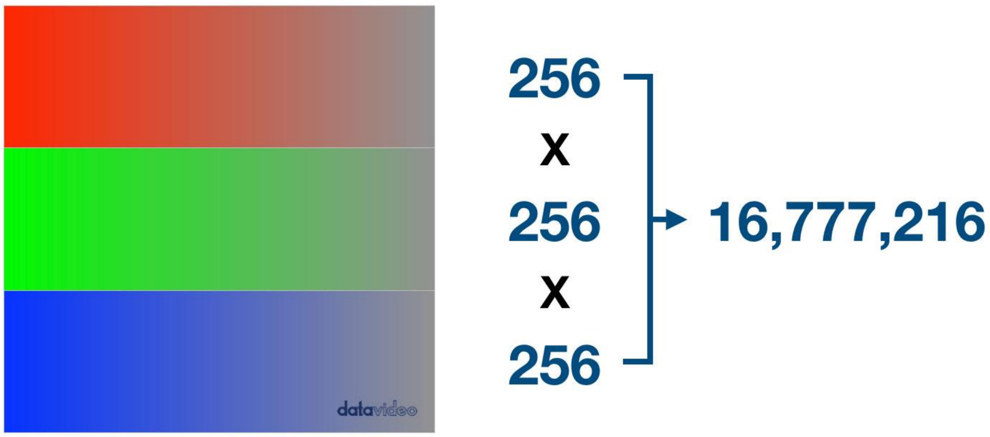
Jan 07 2020

When it comes to digital video production, we often see 8-bit, 10-bit or even 12-bit presenting the spec of image processing. Sometimes you also find the numbers like 4:4:4, 4:2:2 and 4:2:0 on recording devices. What exactly do these numbers mean and how they affect the image quality and colors? We will answer all of your questions in this article.

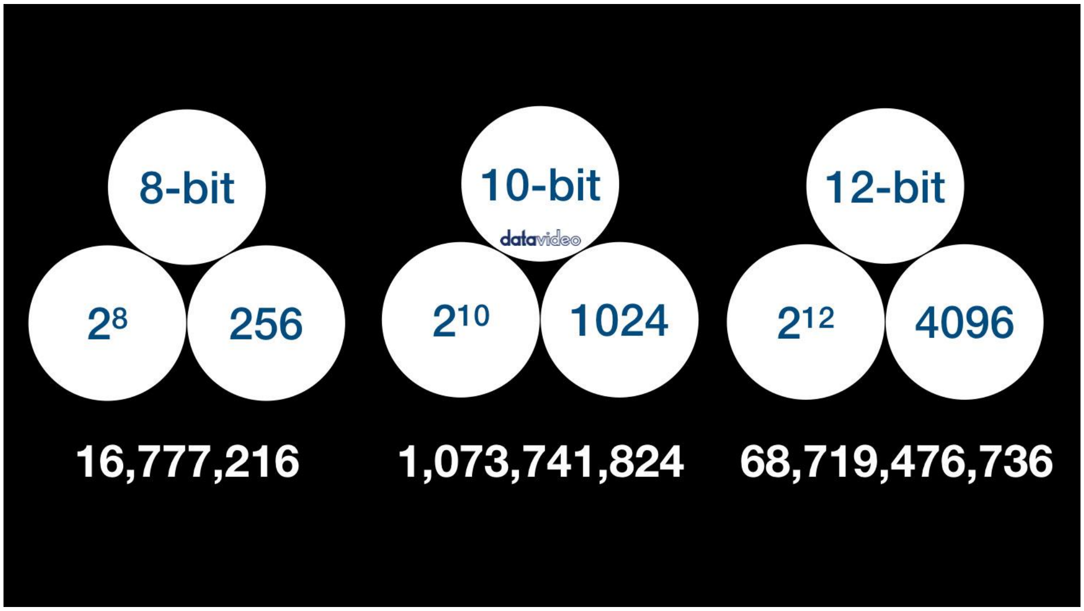
What is 8-bit, 10-bit and 12-bit color depth?

Color depth is also known as bit-depth which refers to the number of bits used to define the color channels, red, green or blue, for each pixel.

In most RGB systems, there are 256 shades per color channel. If you know binary system well enough, this number 256 should sound very familiar to you. The number, 256, is 2 raised to the 8th power or the 8-bit color depth. This means that each of the RGB channels has 256 shades so there are 256x256x256 or 16,777,216 colors in total in this 8-bit RGB system.



An 8-bit color system is capable of producing over 16 million colors. This may look humungous, but when it compared to 10 bit, this is actually nothing. In a 10-bit system, you can produce 1024 x 1024 x 1024 = 1,073,741,824 colors which is 64 times of the colors of the 8-bit. What is more shocking is that a 12-bit system is able to produce a whopping 4096 x 4096 x 4096 = 68,719,476,736 colors! As a result, increasing the color depth will enable you to better represent your colors.



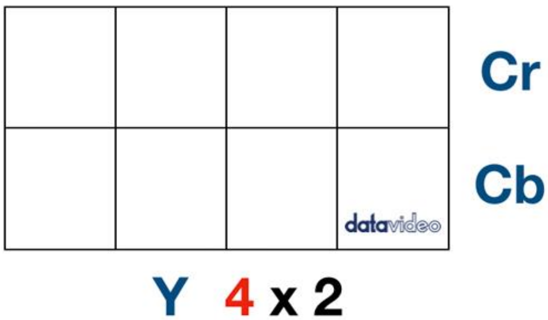
What are chroma sampling and the numbers 4:4:4, 4:2:2 and 4:2:0?

We often see numbers 4:4:4, 4:2:2 and 4:2:0 on recording devices and these are known as chroma subsampling. Have you ever wondered how does chroma subsampling affects the colors of the image? And what exactly do these numbers 4:4:4, 4:2:2 and 4:2:0 mean?

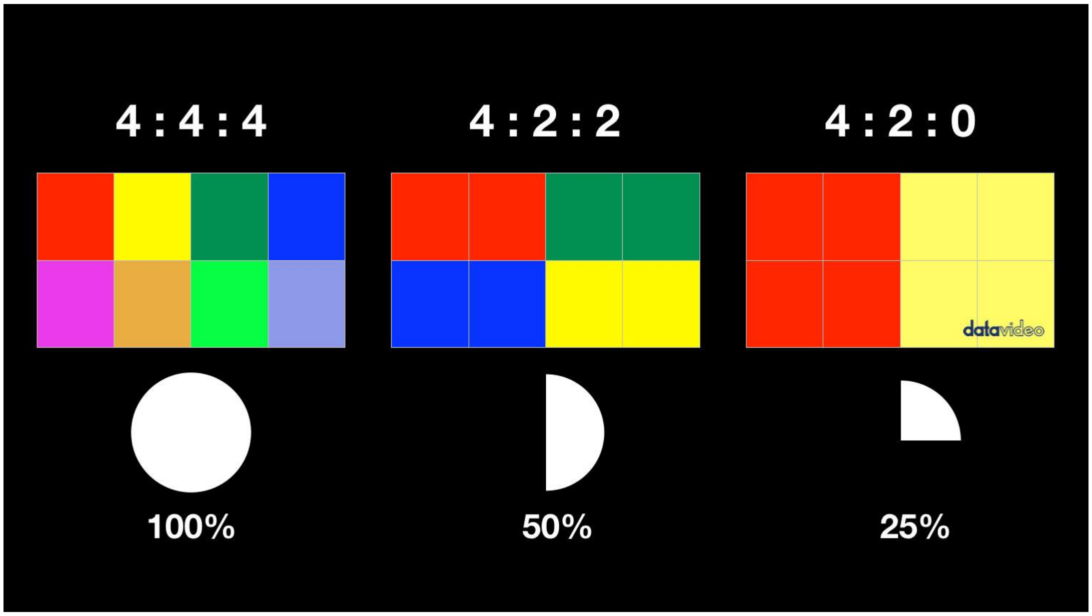
Before we dive into chroma subsampling, let’s first talk about the image pixel. The image pixel is defined by luma and chrominance components. Without the chrominance components, the luma of each pixel produces a greyscale representation of the image. Also, study indicates that human eyes are more sensitive to light or luminance than colors.

YCbCr is a family of color spaces used as a part of the color image pipeline in video and digital photography systems. Y refers to the brightness of the pixel and shares 1/3 the amount of signal. The brightness signal is always retained without compressing. Cb and Cr are the two chroma signals which share 2/3 amount of signal. The chroma signals can be compressed for saving the amounts of data loading.

Y	Cr	Cb
The Numbers of Across Pixels 4 x 2	First Row The Numbers of Color	Second Row The Numbers of Color



Let's take 4:4:4 for the example. The first 4 represents the number of pixels across we are subsampling. The second 4 means 4 colors yield in the first row of chroma sampling, and the third 4, again, means 4 colors yield in the second row of chroma sampling. Technically speaking, 4:4:4 represents each pixel has its' own color value which includes all the chroma information, so it isn't chroma subsampling. Now let's take a look for 4:2:2. The second 2 means two chroma subsampling in the first row. And the third 2 means two chroma subsampling in the second row, too. Therefore, a 4:2:2 image only retains a half of the chroma samples that a 4:4:4 image does. As to 4:2:0, it indicates two chroma subsampling in the first row, and no chroma subsampling in the second row, so the pixels in the second row copy the same chroma value of the first row. As a result, a 4:2:0 image retains only a quarter of the color subsampling that a 4:4:4 image does.



Why is broadcast level video camera so powerful?

Pixels are very tiny dots of color, so it is very hard to find out the noticeable visual difference whether the video is recorded in 4:4:4, 4:2:2 or 4:2:0. However, 4:4:4 is able to record more color information than 4:2:2 and 4:2:0, thus the 4:4:4 chroma subsampling model still has advantages over 4:2:0 and 4:2:2 in terms of the color quality.

Most of the available DSLR and mirrorless cameras on the market use 4:2:0 chroma subsampling model to compress the video files. Even though you can yield good image quality from the 4:2:0 video, you might still encounter problems when doing chromakeying or post-editing because of the low resolution for chroma information. Compared to 4:4:4 images, it will be more difficult and time-consuming to achieve a clean chromakey result with 4:2:0 videos. This is why professional video producers still prefer working with 4:4:4 or 4:2:2 video, which contains more chroma information facilitating the post-edits, only the final video is compressed in 4:2:0 for saving the size of file. This production procedure is like that a professional photographer always shoot photos with RAW files and then output the post-edits pictures in JPG format for the subsequent applications.

By knowing the theory of chroma subsampling, audience should by now know why only the professional broadcast level video equipment is able to present a very high-quality image, and why they are expensive than the consumer digital cameras and mobile phones. Let's take the example of the BC-100 Interchangeable Lens Video camera. The BC-100 is a broadcast level video camera designed for virtual studio. The camera is equipped with a 12-bit image processing sensor, capable of capturing a massive amount of color information and presenting the finest color differences. The rich colors and sharp image quality are not only for visual enjoyment but also essential properties for attaining clean and pure objects from background by employing chromakeying. With the advanced technique, you can easily chroma key the challenging objects such as a glass or hair, presenting the finest details combined with virtual background. Additionally, the High Dynamic Range (HDR) technique allows video camera to record the details of the bright and dark parts of the image in high contrast conditions, making images more real as seen by the human eye.

The Techniques to Master Correct Exposure

Mar 08 2021

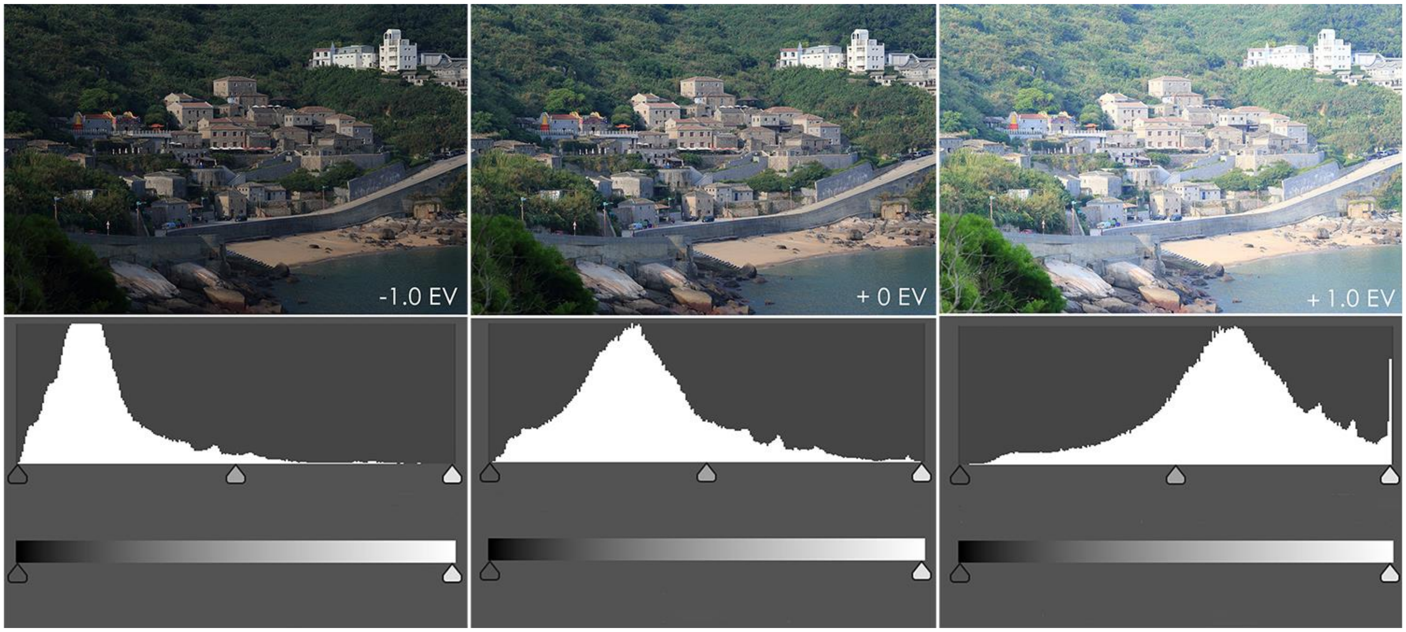
Have you ever looked at the LCD screen of a camera in a bright room and thought that the image was very dim or under-exposed? Or have you ever seen the same screen in a dark environment and thought the image was over-exposed? Ironically, sometimes the resulting image is not always what you think it will be.

"Exposure" is one of the essential skills for shooting videos. Though users can use image-editing software to make adjustments in post-production, managing correct exposure can help the videographer get high-quality images and avoid spending excess time in post-production. To assist videographers in monitoring image exposure, many DSLRs have built-in functions to monitor exposure. For example, the Histogram and Waveform are handy tools for professional videographers. In the following article, we are going to introduce the standard functions for getting correct exposure.

Histogram

Histogram Scope is composed of an "X-axis" and a "Y-axis." For the "X" axis, the graph's left side represents the darkness, and the right side represents the brightness. The Y-axis represents the pixel intensity distributed throughout an image. The higher the peak value, the more pixels there are for a specific brightness value and the larger area it occupies. If you connect all the pixel value points on the Y axis, it forms a continuous Histogram Scope.

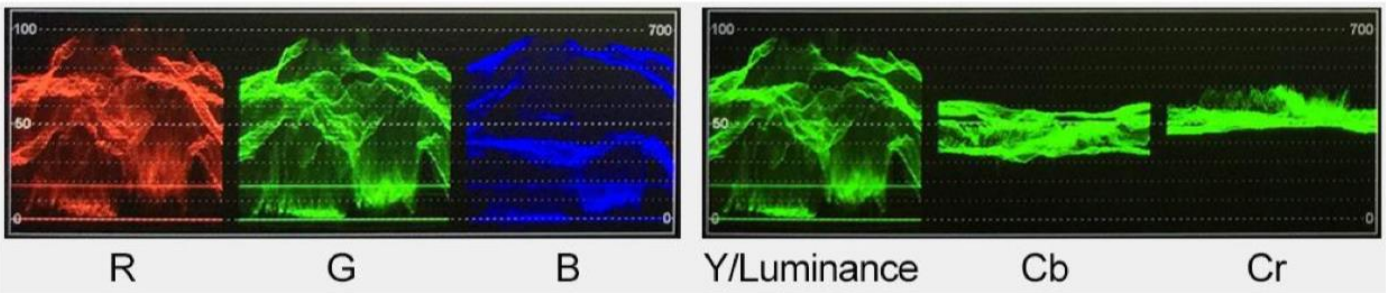
For an overexposed image, the histogram's peak value will be concentrated on the right side of the X-axis; conversely, for an underexposed image, the histogram's peak value will be concentrated on the left side of the X-axis. For a properly balanced image, the histogram's peak value distributes evenly on the center of the X-axis, just like a normal distribution chart. Using the Histogram Scope, the user can evaluate whether the exposure is within the correct dynamic brightness and the color saturation range.



Waveform Scope

The Waveform Scope shows the luminance and RGB & YCbCr values for the image. From the Waveform Scope, users can observe the brightness and darkness of the image. The Waveform Scope converts the bright level and the dark level of an image to a waveform. For example, if the "All Dark" value is "0" and the "All Bright" value is "100", it will warn users if the dark level is lower than 0 and the brightness level is higher than 100 in the image. Thus, the videographer can better manage these levels while shooting video.

Currently, the Histogram function is available on entry-level DSLR cameras and field monitors. However, only the professional production monitors support the Waveform Scope function.



False Color

The False Color is also called "Exposure Assist." When the False Color Function is on, an image's colors will be highlighted if it is over-exposed. So, the user can examine the exposure without using other expensive equipment. To fully realize False Color's indication, the user must understand the color spectrum shown below.

For example, in areas with an exposure level of 56IRE, the false-color will be shown as pink color on the monitor when applied. Therefore, as you increase the exposure, that area will change color to grey, then yellow, and finally to red if overexposed. Blue indicates under exposure.

Zebra Pattern

The "Zebra Pattern" is an exposure-assisting function that is easy to understand for new users. Users can set a threshold level for the image, available in the "Exposure Level" option (0-100). For example, when the threshold level is set to "90", a zebra pattern warning will appear once the brightness in the screen reaches above "90", reminding the photographer to be aware of the image's overexposure.

Coming Next: HDR