Myelin Whitepaper Series

VIDEO ENCODING

Rate Control Mode

12th October 2021

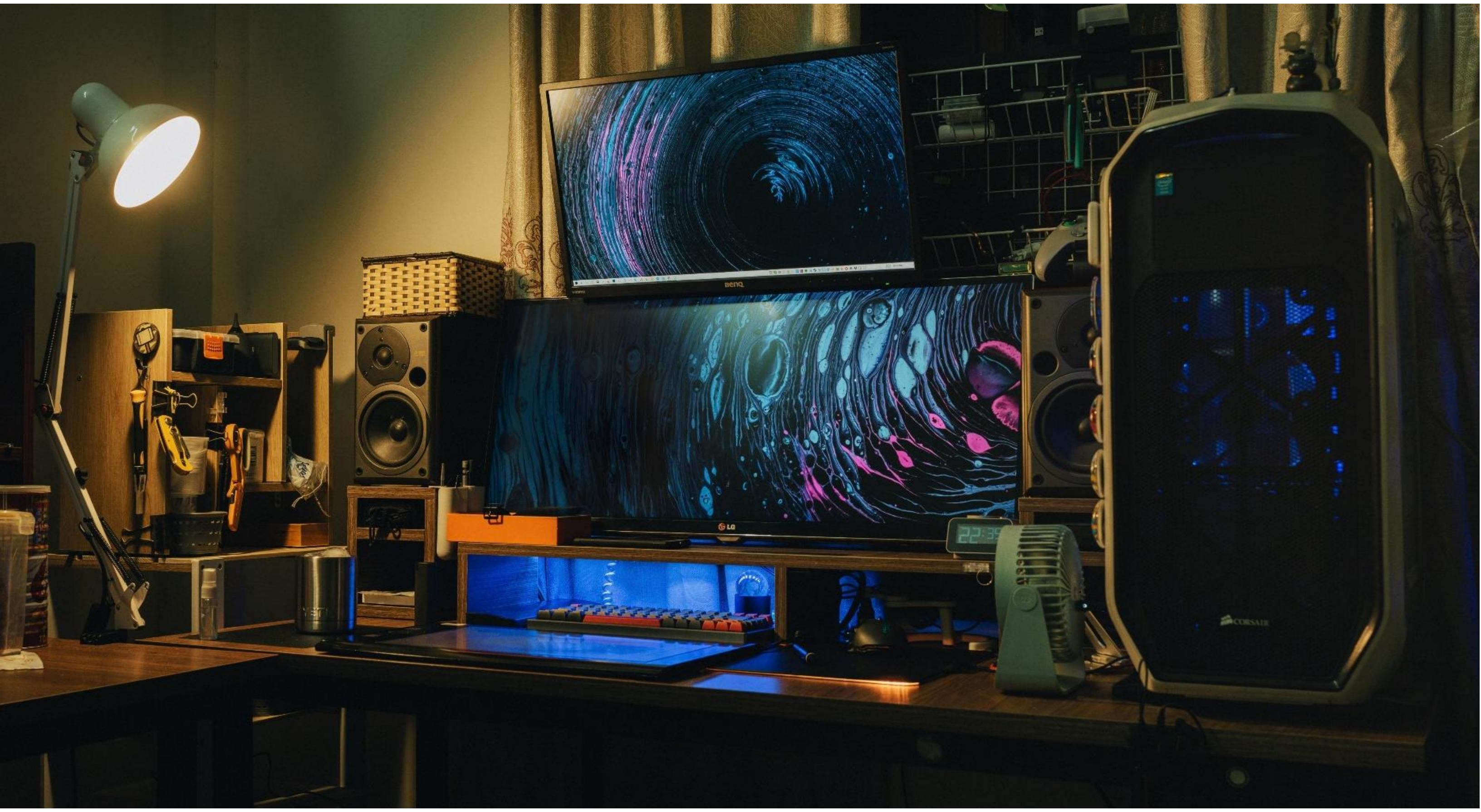
Pallabi Chandra

Product Engineer, AI & Deep Learning

Myelin Foundry®

Everything AI, Anything Creative

TABLE OF CONTENTS



WHAT ARE RATE CONTROL MODES?

The two major factors affecting video quality are-resolution and bitrate. Differing video quality can be attributed to the different principles based on which bits are allocated to the video packets, called rate control modes. Rate control modes define how the encoding engine uses the buffer model and how bits flow into the engine.

In this paper, we will demonstrate how the distribution of bits is inherently dependent on the spatial and temporal characteristics of the input video itself. The video quality in this study is represented in terms of VMAF (Video Multimethod Assessment Fusion), a perceptual video quality metric (scored out of 100) developed by Netflix. Higher the VMAF score, better the perceptual quality of the video. All the analysis presented was done using the AVC codec.

SCENE COMPLEXITY

In this paper, we will use the qualitative terms 'simple' and 'complex' scenes very frequently. Thus, it is important to understand their meaning.

- **Simple Scene**: A scene with low intra-frame spatial variation and low inter-frame temporal variation. For example, a slow-moving panoramic shot of the sky. In such a shot, the spatial and temporal variation will be low since it will consist of a uniform blue background interspersed with clouds across multiple frames, with very low variation in textures.

- **Complex Scene**: A scene with high intra-frame spatial variation and high inter-frame temporal variation. For example, a performance in a reality show with flashing lights, multiple background dancers with swift movements, rapid shot transitions between performance and audience, etcetera. The intra-frame spatial variation will be high since the performers and the stage will have wide chromatic contrast as well as rich texture details in their attire. The inter-frame temporal variation will be high due to rapid shot transitions between stage and audience.

CONSTANT V/S VARIABLE BITRATE

Bitrate is essentially the number of bits assigned to a video encode output per second. Usually, there are two ways in which bits are distributed across an output video- at a constant rate or at a variable rate.

A constant bitrate would imply that a chosen number of bits during the encode would be uniformly distributed over intervals of a second. For example, in a 25fps video, it would be distributed over 25 frames, such that the total number of bits assigned to these frames is cumulatively the bitrate value chosen during the encode. If this is the method adopted while encoding, the bit usage would be inefficient, but the encoding time would be quicker and many QOE (Quality of Experience) parameters such as rebuffering, latency issues would be mitigated. However, the overall perceptual quality of video might take a hit in case there are many complex scenes.

A variable bitrate, on the contrary, would imply that different magnitudes of bits are distributed throughout the video in intervals of one second. The principle behind the variable distribution of bits can differ. It may be distributed according to a prescribed quality requirement, a range of maximum and minimum bitrate, an average bitrate value, a quantization parameter, etcetera. Typically, the variable bitrate encodes take greater encoding time and may impact rebuffering or latency negatively. However, the perceptual quality of the video is likely to improve since the complex scenes will have a higher number of bits assigned to it as compared to the simple scenes.

In this paper, we will mainly speak about rate control modes offered by the following encoders- encoders from AWS such as MediaLive and MediaConvert and the open-source encoder FFMPEG. AWS encoders are highly popular among several video streaming platforms for large scale encoding. The main rate control modes available on AWS encoders are Constant Bitrate (CBR), Variable Bitrate (VBR) and Quality-Defined Variable Bitrate (QVBR). Some companies operating at a smaller scale use FFMPEG as part of their encoding pipeline, since it is free and easier to customize for niche requirements. The main rate control modes available in FFMPEG are Constant Rate Factor (CRF), Constant Quantization Parameter (CQP), Constant Bitrate (CBR), and Average Bitrate encoding. These modes can be coupled with the Video Buffering Verifier (VBV) option to constrain encoding in different ways.

RATE CONTROL MODES ON AWS

As discussed above, the three major rate control modes offered by AWS MediaLive / MediaConvert are:

- **Constant bitrate (CBR)**: Maintains constant bitrate throughout the video irrespective of scene complexity. It wastes many bits in the portions where the scenes are simple and do not need as many bits as the CBR encode prescribes. More complex portions of the video that need higher bits than the designated CBR are of poor quality, since a large number of bits have already been allocated to simple scenes.

- **Variable bitrate (VBR)**: Accepts a value for an average and maximum bitrate across the video. In simpler portions of the video, the bits are distributed such that the minimum bitrate is at least the average value specified. Thus, if the video is still less complex and needs lesser bits than the average bitrate, some bits are wasted. In complex portions, the maximum bits allowed is the maximum bitrate value, keeping the constrain of the lower limit being the average bitrate each second. Thus, if the video is still more complex and needs more bits than the maximum bitrate allowed, the complex portions are still of poor quality and may contain artifacts.

- **Quality-defined Variable bitrate (QVBR)**: Accepts a maximum bitrate for variable bitrate across the video to meet prescribed quality level. The quality levels range from 1-10. Higher the magnitude of the quality level chosen, better the perceptual quality. In simpler portions of the video, the bitrate might in certain cases go significantly lower than the maximum bitrate specified, if it can meet the quality level satisfactorily. In complex portions of the video, a value close to maximum bitrate is used, keeping the constrain of the required quality level to be met. Since the simpler portions have the lowest possible bit allocation, meeting quality for the complex portions becomes within a maximum bitrate threshold.

The following is a comparison between different rate control modes in terms of important parameters-

Criterion	CBR	VBR	QVBR
Bit usage Efficiency	Wastes many bits	Wastage lower than CBR	Most efficient bit usage
Encoding speed	Highest	Second Highest	Lowest
Packet Size Variation	Lowest	Second Lowest	Highest
Compute Resources	Lowest	Second Lowest	Highest

Impact on Quality

The following chart in **Figure 1** represents the perceptual quality comparison between QVBR and CBR rate control modes across different genres of content, for the same average bitrate. In order to carry this out, several source clips were first transcoded to 360p resolution using QVBR and a fixed maximum bitrate. Following this, the outputs from the encodings were examined to determine the average allocated bitrate, which was found to be 460 kb/s. Using this average bitrate, the same source clips were transcoded using CBR.

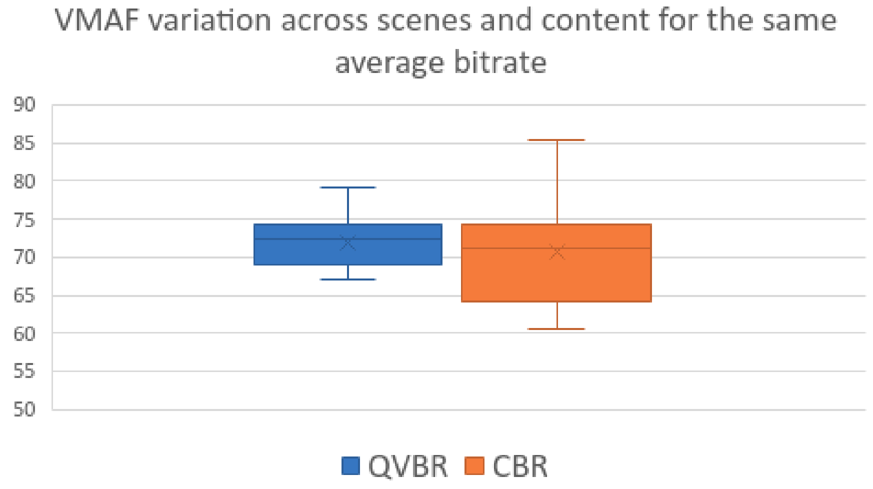


Figure 1

An attempt was made to study the impact of differing QVBR quality levels on perceptual quality. To this end, a source clip was transcoded to different QVBR quality levels, keeping the maximum bitrate and resolution(360p) constant as depicted in **Figure 2**. The maximum bitrate paired with the quality level was chosen such that the average allocated bitrates were approximately equally spaced apart, resembling those of an ABR ladder.

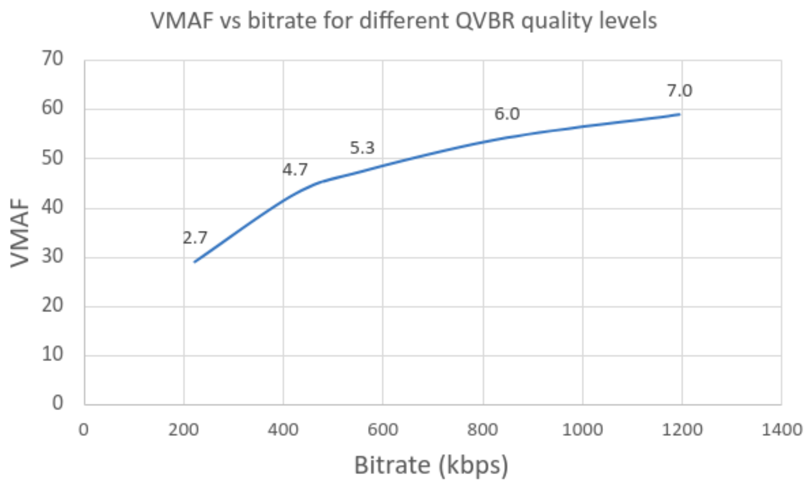


Figure 2

Important Observations

- QVBR provides a more consistent perceptual quality across multiple videos of different genres as compared to CBR for the same average bit usage, as is indicated by the boxplots in case of VMAF of the two.
- With increase in quality level, the average allocated bitrate as well as the perceptual quality of a video rises.

Recommendations

Thus, based on the above findings, here are some generic examples for different use cases-

<table>
<tr><th colspan="2">Live Streaming
E.g., News and Sports</th><th colspan="2">Video On-Demand
E.g., Films and TV Series</th></tr>
<tr><th>Lower Bitrate Range, Low Resolution</th><th>High Bitrate Range</th><th>Drama Content</th><th>Action Content</th></tr>
<tr><td>• Use QVBR for a boost in quality.
• Latency issues might be incurred, and a tradeoff must be maintained.
• Maintain the quality level at 6.</td><td>• Use CBR since sufficient bits are available.
• Slight hit in quality may be incurred.
• Low latency.</td><td>• Use CBR.
• Consists of simple scenes for which available bits are sufficient</td><td>• Use QVBR to get a boost in quality for complex scenes.
• Maintain the quality level at 7 for a mobile screen.
• Maintain the quality level at 9 for a TV screen.</td></tr>
</table>

These parameters can further be fine-tuned based on the specific customer’s use case and its nuances.

RATE CONTROL MODES ON FFMPEG

As discussed above, the major rate control modes offered by FFMPEG are:

- **Constant Quantization Parameter (CQP)**: The encode works based on a quantization parameter (QP) which controls quantization for every block in each video frame. The QP values range from 0 to 51-a higher value indicates more compression and worse perceptual quality.

- **Constant Bitrate (CBR)**: Forces the encode to maintain a constant bitrate throughout the video. For this purpose, the minimum and maximum bitrate parameters should be set to identical values. Its bit usage is inefficient since many bits may be wasted in simpler portions of the video. This encode is similar to the CBR encode in case of AWS encoders

- **Average Bitrate**: Takes only an average bitrate value and allows the encoder to arrive at it how it deems fit. It usually yields poor quality results. It is similar to the VBR encode of AWS encoders.

- **Constant Rate Factor (CRF)**: It maintains consistent prescribed quality throughout, but the distribution of bits and file size can’t be controlled by directly specifying a bit rate. The CRF values range from 0 to 51 where a higher value indicates more compression and worse perceptual quality. It has efficient bit usage. This mode is similar to QVBR mode in AWS. Just as the QVBR mode accepts a maximum bitrate, we can constrain the CRF encoding with a maximum bitrate using the Video Buffering Verifier (VBV) option.

Impact on Quality

In this section, we will analyze the constant rate factor mode in more detail. Figure 3 depicts the effect of changing constant rate factor modes on a video transcoding, keeping the resolution constant. In order to construct this curve, a source clip was transcoded to 360p and the CRF values were chosen such that the average bitrates were approximately equally spaced apart. The maximum bitrate (using VBV) chosen was the same in the case of Figure 2, so that adequate parallels could be drawn between the two.

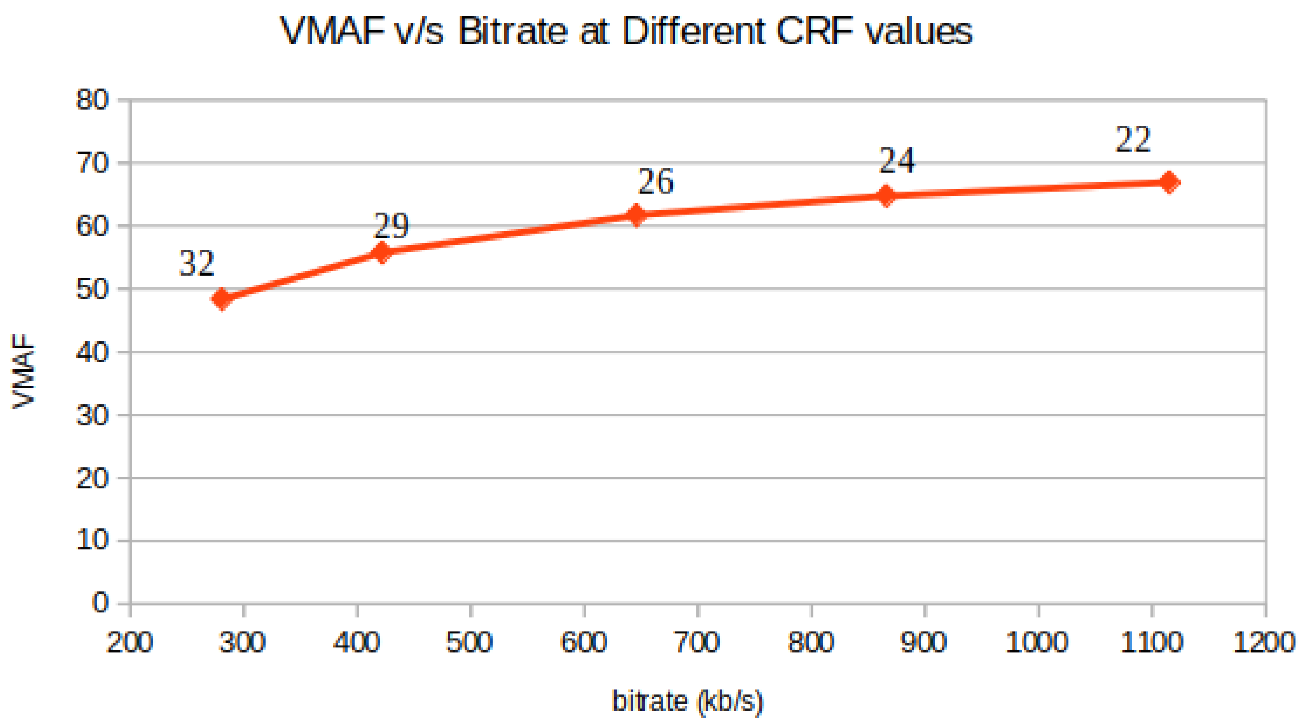


Figure 3

Conversely, Figure 4 depicts the effect of changing resolutions of the video transcode output, keeping the CRF value constant. In order to construct this curve, a source clip was transcoded to separate resolutions with CRF 23, and the resolutions were chosen in accordance with industry standards for ABR ladders.

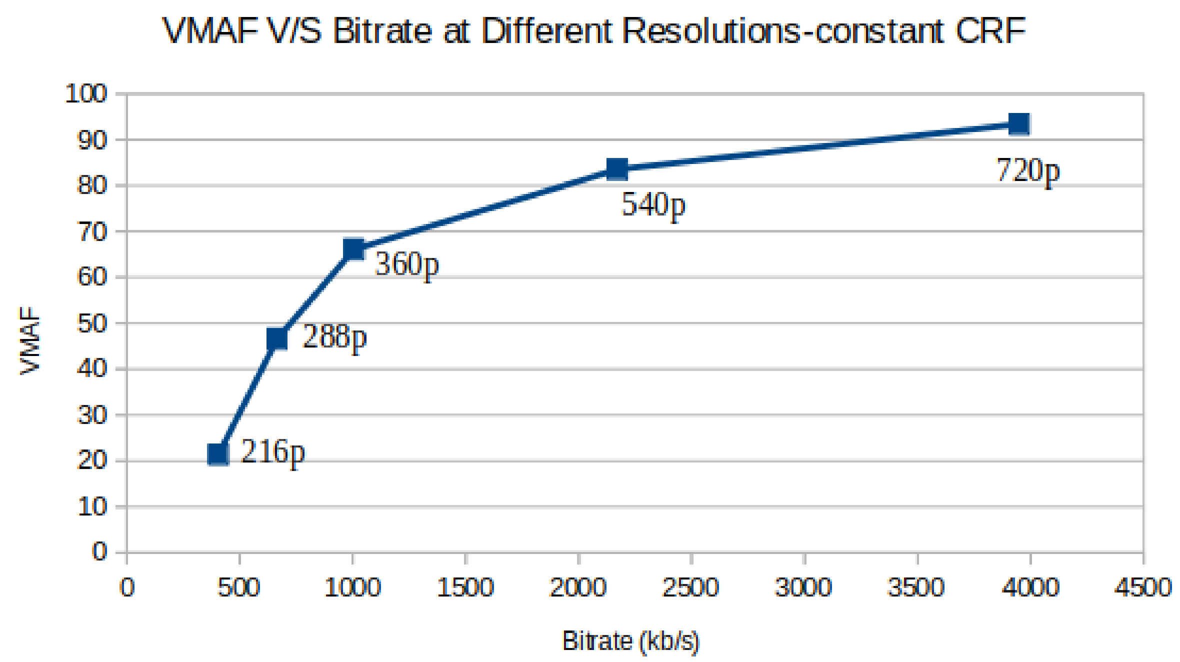


Figure 4

We also studied the impact of CRF encodings on separate types of content. The following chart in Figure 5 represents the VMAF at the same CRF levels (360p) for different types of content such as sports and drama. There was no maximum bitrate constraint used for these encodings.

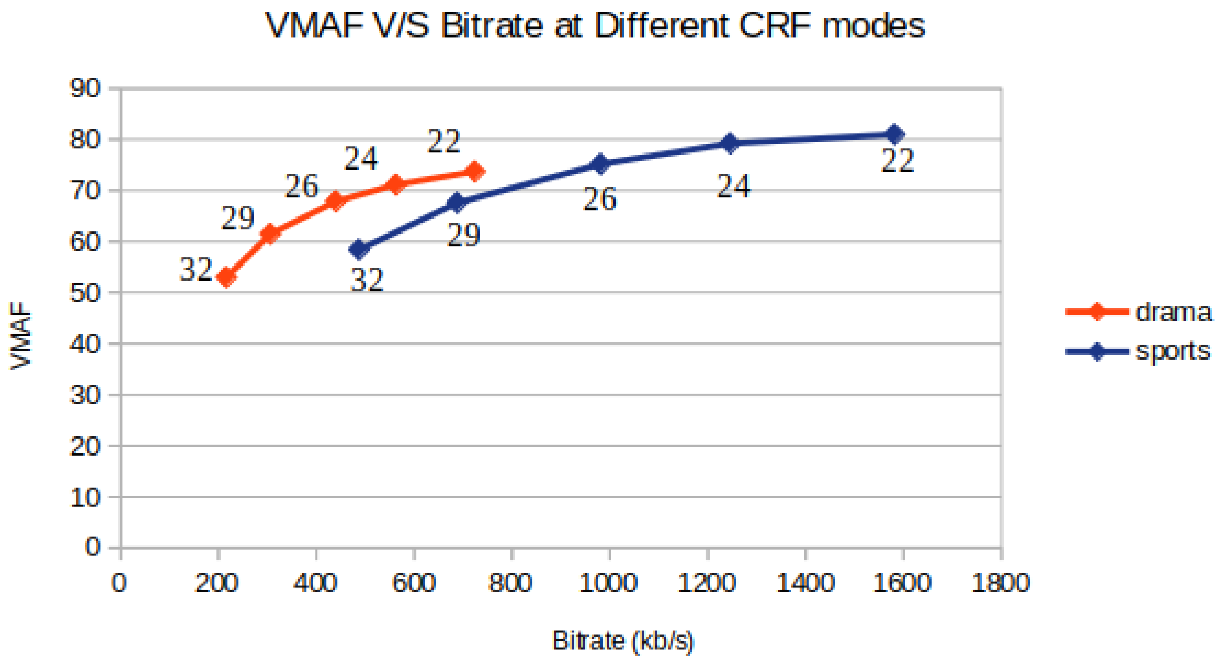


Figure 5

Important Observations

- With decrease in CRF values at the particular resolution, the perceptual quality of a video increases. However, the increase is less steep as we go to lower CRF values. Furthermore, the allocated average bitrates also increase with decrease in CRF value, as we move closer to lossless compression.

- With increase in resolution at a particular CRF value, the perceptual quality increases and so does the average allocated bitrate.

- The behavior of QVBR quality levels is similar to CRF values with a maximum bitrate constraint applied, as is indicated by the shape of the VMAF v/s bitrate curve for the two.

- Different genres of content may have different perceptual quality even if parameters such as CRF values and resolution are identical.

Recommendations:

Thus, based on the above findings, here are some generic examples for different use cases-

Live Streaming E.g., News and Sports		Video On-Demand E.g., Films and TV Series	
Lower Bitrate Range, Low Resolution	High Bitrate Range	Drama Content	Action Content
• Use CRF for a boost in quality. • Latency issues might be incurred, and a tradeoff must be maintained. • Maintain the CRF mode at 23 for AVC and 28 for HEVC.	• Use CBR since sufficient bits are available. • Slight hit in quality may be incurred. • Low latency.	• Use CBR. • Consists of simple scenes for which available bits are sufficient	• Use CRF to get a boost in quality for complex scenes. • Maintain the CRF mode at 23 for AVC and 28 for HEVC.

These encoding parameters can be further fine-tuned based on the specific customer's use case and its nuances.

CONCLUSION

From the above analysis, it's clear that the selection of a rate control for video encoding greatly impacts video quality and the distribution of bits across the video.

- Using **perceptual video quality** assessments through VMAF, for a given bitrate and resolution, the best order of video output perceptual quality and efficiency of bit usage would be **QVBR > VBR ~ CBR** for AWS and **CRF > Average Bitrate > CBR** for FFMPEG

- Conversely, there is a tradeoff to be considered in terms of video packet size according to the content being encoded, where the order of variation would be **QVBR > VBR > CBR** for AWS and **CRF > Average Bitrate > CBR** for FFMPEG

- **QVBR / CRF + VBV** provides a more consistent perceptual quality across multiple videos of different genres as compared to CBR for the same average bit usage

- **QVBR** and **CRF** on AWS encoders and FFMPEG both use constant video quality as the guiding principle behind bit distribution and produces similar encoding results

The selection of a rate control mode is incumbent on several factors including perceptual quality, bandwidth savings, user QOE (Quality of Experience) and latency demands, etc. Thus, the particulars of a use case are very important in rate control mode selection.

REFERENCES

1. Understanding Rate Control Modes (2017, March 1) Werner Robitza

   https://slhck.info/video/2017/03/01/rate-control.html

2. AWS Rate Control Modes (2021) AWS Elemental MediaConvert Guide

   https://docs.aws.amazon.com/mediaconvert/latest/ug/cbr-vbr-qvbr.html

3. Toward a Better Quality Metric for the Video Community (2020) Netflix

   https://netflixtechblog.com/toward-a-better-quality-metric-for-the-video-community-7ed94e752a30

ACKNOWLEDGMENT

- **Anusha Rammohan**, Senior Technology Leader, Myelin Foundry
- **Anshul Gupta**, Product Marketing Lead, Myelin Foundry
- **Dinesh Babu Sukumar**, Product Marketing Manager, Myelin Foundry