envivio®

TV without Boundaries™

Video Quality Optimization • Multi-rate Video Encoding

Best Practices Recommendations for Optimized Encoding Resolution

July 2010

Authors: G. Hervouet, M. Chouikrat, B. Felts

www.envivio.com

Contents

1. Introduction

The world's appetite for video is growing at an unprecedented pace as people are watching more television programming, from more sources, on an increasing variety of screens. There is a huge opportunity to capture a rising percentage of a viewer's time, if one overcomes the challenge of delivering a great viewing experience with exceptional quality to multiple devices, whatever the network conditions.

1.1 Importance of Video Resolution in Video Compression

Why is video resolution a key parameter in video compression?

While the television format has not really changed, there is now a huge variety of screens in all kinds of sizes and new devices featuring varied aspect ratios to be addressed. The one commonality in this array of equipment is that they are all capable of rescaling the video, as this is the case for the PC player going to the Full Screen mode, or the optimized rescaling on the iPhone®.

Encoding video resolution is one of the key parameters to optimize the perceived quality of a stream, while not having to match the terminal resolution. Quite often, the network constraints dictate the bandwidth that may be used for video transmission, and high resolution requires more bandwidth than low resolution. For a given bandwidth, you need to examine the tradeoff between lower resolutions (images being slightly blurred) with fewer artifacts versus higher resolutions that provide a sharper image but possibly with more artifacts.

1.2 Resolutions and Multi-rate Encoding: Context and Challenges

Adaptive streaming technology lets viewers enjoy continuous video playback despite fluctuant network conditions. Most adaptive streaming technologies are based on multi-rate encoding: the same content is encoded at multiple rates (multiple profiles) and the player device seamlessly selects profiles according to the bandwidth available. Various techniques, such as Live Bitstream Switching (used in 3GPP Mobile Streaming) or Smooth Streaming (implemented in Silverlight®), Flash® Dynamic Streaming or Apple HTTP Live Streaming (present in iPhone and iPad®) use a similar multi-rate approach.

As these devices are connected to the Internet through heterogeneous networks (Mobile, Wi-Fi, Broadband and Hi-Speed Broadband connections), they have to adapt to a large range of bit rates. For adaptive video streaming, this implies having numerous profiles to support, each being tuned for a particular bit rate.

For the operator, this means that the calibration of a multi rate encoder becomes a lot more complex than in traditional broadcast. He will have to answer the following questions:

- How many profiles are required to cover a range of bit rate?
- What should be the ideal resolutions for each profile?
- How can I draw objective comparisons on video quality across different resolutions?

Nothing will replace the experience and art of a video compressionist in order to really fine tune the parameters of an encoder, but we will try to give in this paper some guidance and recommendation on how to optimize one key parameter in multi-rate encoding: the resolution.

We first explain our methodology to compare quality across different resolutions. Then, in the second part, we provide recommendations for resolution and bandwidth in order to reach optimum video quality. Finally, we set up best practices guidelines within a multi-rate scenario.

2. Methodology

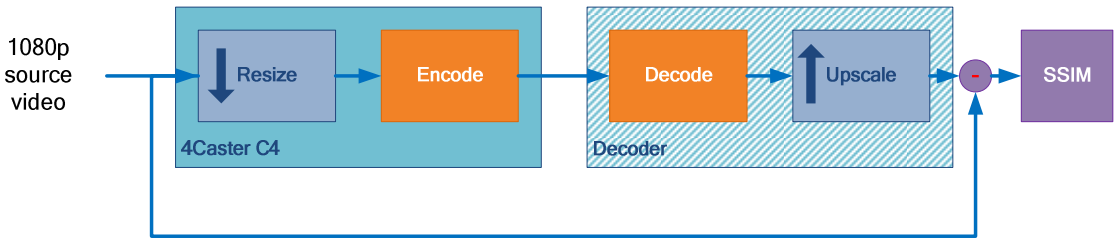
2.1 Test and Measurement Principles

The measurement setup reproduces the end-to-end treatment received by the video signal during its transmission and playback:

- Envivio 4Caster C4, multi-resolution encoder, resizes the video and encodes it
- The decoder decodes the video and scales it back to the screen resolution

In order not to make abstraction of the screen resolution and quality of the scaling processing of the playback device from the results, we compared the input and the output signal restored in its original resolution (1080p), thanks to a good quality upscaling filter[1]. The restored video signal is compared to the original using SSIM[2] measurement.

The overall process is summarized below:



We focused our study on the quality impact of the encoding resolution at any given bit rate.

For a particular content type, measurements have been interpreted in order to provide recommendations to the following issues:

- Which resolution is best suited to a given bit rate or a given range?
- What bit rate or range best applies to a given resolution?

2.2 Quality Metrics

There is no such thing as a perfect objective video quality metrics. There is still nothing better than subjective testing. However subjective testing is quite difficult to put in place and is not well suited for testing a lot of video samples. This is why we used an objective testing approach, which allowed us to automate the tests and results.

Amongst the various quality metrics commonly used (PSNR, MSE, VQM, and SSIM), we have chosen the SSIM index, as it is considered to correlate relatively well with human perception.

---

[1] The upscaling depends on the specific playback device capacity and may affect the results. We used a good upscaling filter here.

[2] See SSIM definition on next page.

The SSIM (Structural SIMilarity) index is a subjective quality and fidelity metric that measures the similarity between the original image and the same image after processing. It is a decimal value comprised between 0 (zero correlation with original picture) and 100 (absolute similarity with original picture).[3]

2.3 Resolutions

The resolutions used in our study are summarized below:

Resolution Label	Dimensions (pixels)
QCIF	176x144
QVGA	320x240
HVGA	480x 320
VGA	640x480
WVGA	854x480
WSVGA	1024x600

Resolution Label	Dimensions (pixels)
720p 3/4	960x720 at 30fps
720p full	1280x720 at 30fps
1080p 1/2	960 x 1080
1080p 3/4	1440x1080
1080p full	1920 x 1080

2.4 Quality-optimized Scaling Filters

The 4Caster C4 supports a set of resolutions covering a wide range of use cases. Each of these resolutions is obtained thanks to high quality resizing filters.

The resizing algorithms present in the 4Caster C4 are based on hanning windowed polyphase filters, providing optimal visual results. Thanks to this method, the encoder is fed with contents of the best possible quality. Moreover, the most common resolutions have been computationally optimized. Thus 4Caster C4 delivers optimal quality at low computational cost for multi-rate encoding.

A list of resolutions currently supported is available in the Annex in section 5.

[3] More information on SSIM can be found at: http://www.cns.nyu.edu/pub/eero/wang03-reprint.pdf

2.5 Measurements

For a range of predefined bit rates, the SSIM quality indexes of a video sequence are measured at various resolutions. Such results are represented in a graph such as the one below.

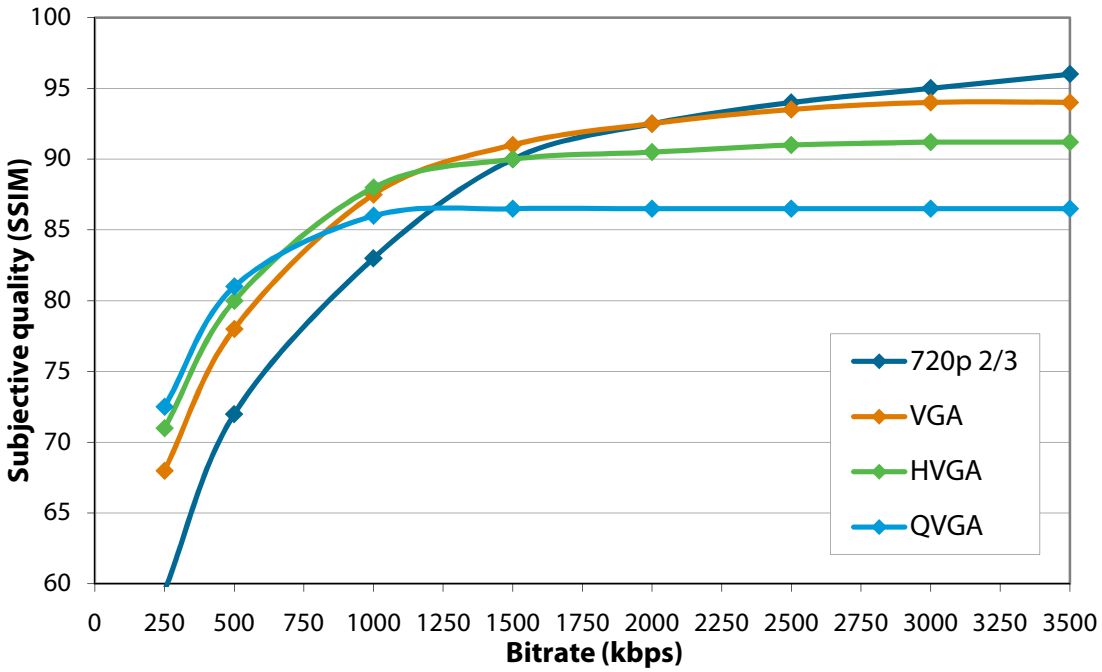


Figure 1. Video quality measures for the "Public Television" video sequence for various resolutions at various bit rates

From this graph and on this particular video sequence, we can see that the QVGA resolution has the best results at low bit rate, while 720p 2/3 had the best resolution in high bit rates.

2.6 Determination of the Optimal Resolution

For a given bit rate, the optimal resolution is the one that features the highest SSIM measurement. Link the highest points to draw a graph and you will draw an optimal resolution graph which is represented in red in the graph below:

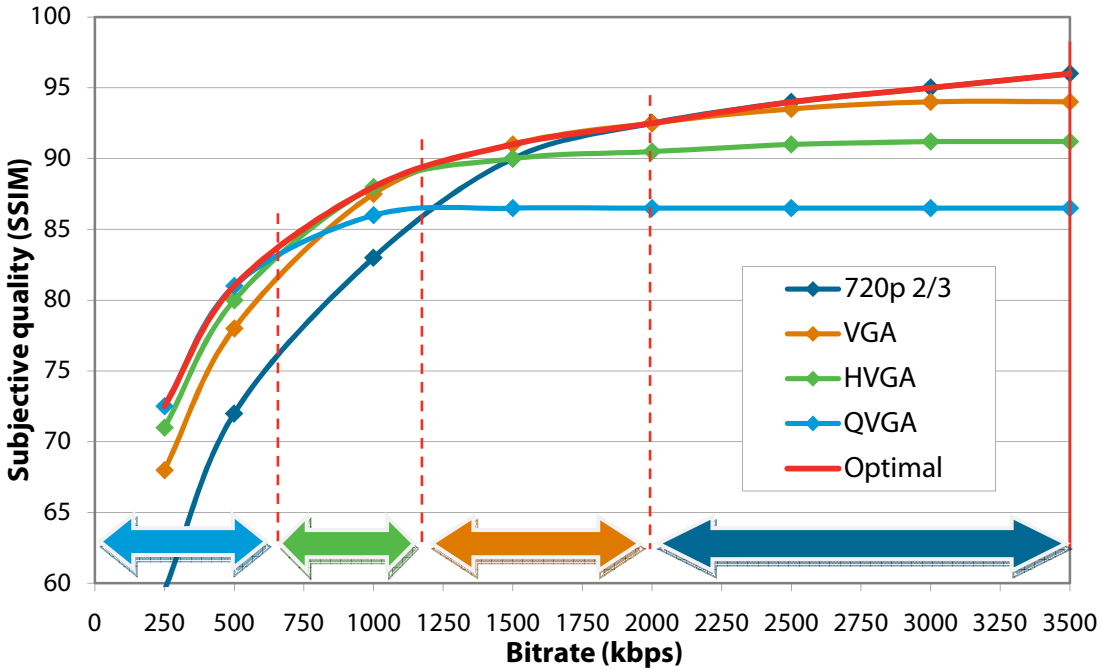


Figure 2. Optimal resolution for the sequence "Public Television" at various bit rates

For this particular sequence, we can conclude that, for each bit rate range, the optimal resolution is:

Bit-rate Range (Mbps)	Optimal Resolution
< 0.6	QVGA
0.6 to 1.2	HVGA
1.2 to 2	VGA
2 to 3.5	720p 2/3

3. Recommendations in the General Case

The results previously found applied to one video sequence. In reality, these results can vary a lot according to the type of content compressed.

3.1 Content Categorization

We gathered a full range of high-definition video sequences at full resolution (1080p) and distributed them into four A, B, C and D representative categories. These sequences range from "very easy to encode" to "very difficult to encode" as follows:

Content Type A:
Near static

Very easy to encode
(e.g. near static shots, TV shows)

Content Type B:
Slow and uniform motion

Easy to encode
(e.g. documentary channels)

Content Type C:
Dynamic motion (soccer)

Difficult to encode
(e.g. sports, action movies)

Content Type D:
Very dynamic and complex motion

Very difficult to encode (e.g. complex transitions, frequent scene cuts and occlusions)

Sequences pertaining to each category are encoded at various bit rates and submitted to SSIM measurements, following the method described previously.

3.2 Guidelines for Optimal Resolution

Measurements obtained for sequences pertaining to each category are aggregated and synthesized into the graph below. For a given bit rate range, this graph indicates the optimal resolution determined for one category of content.

Resolutions are indicated on the vertical axis while bit rates (expressed in Mbps) are displayed on the horizontal axis.

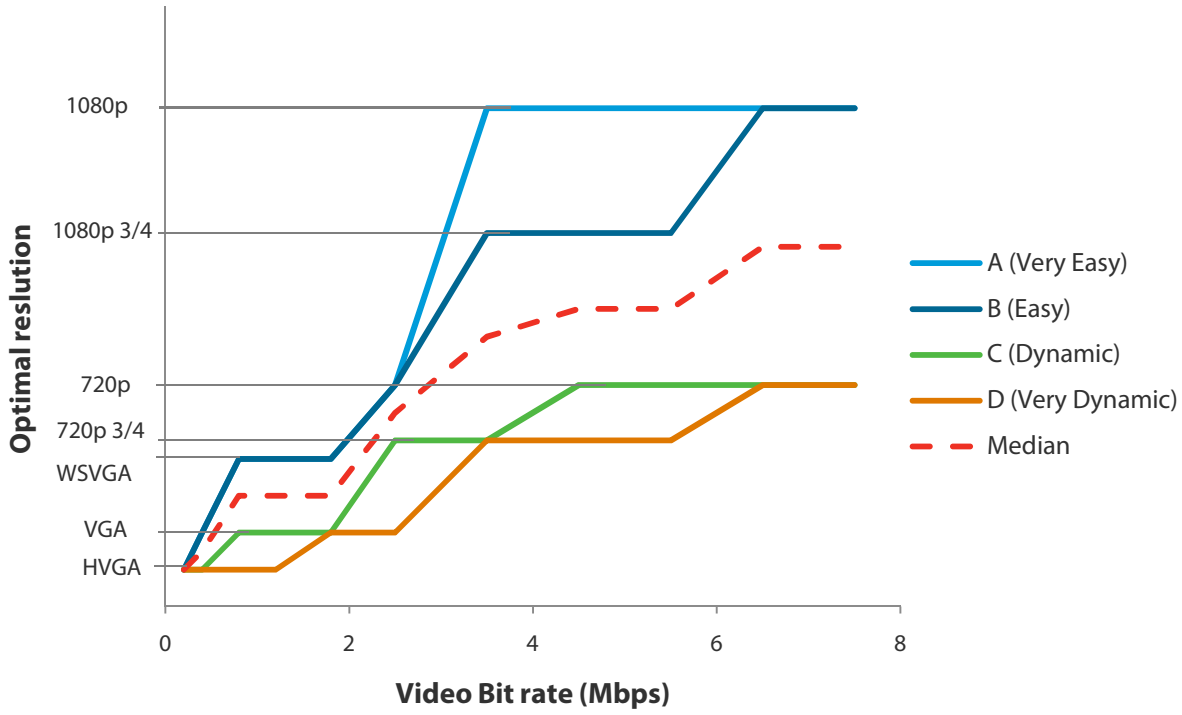


Figure 3. Optimal resolution at various bit rates for each content

To determine the optimal resolution at a given bit rate, the graph reads as follow:

- Select the content type
- Select the video bit rate
- Find the resolution at the intersection of the content curve and the bit rate

For example, the recommended resolution for sport video (category C) at 3 Mbps is 720p 3/4.

The graph can also help determining the range of bit rate applicable for a resolution. It reads as follow:

- Select the content type
- Select the resolution
- Find the corresponding range of bit rates

For example, some talking shows (category A) can be encoded at full 1080p starting at 3 Mbps.

We also indicated on a "median" curve the intermediate resolution that could be applicable to heterogeneous content.

The results are also summarized in the table below:

Bit rate Range (Mbps)	**Category A** Very Easy	**Category B** Easy	**Category C** Dynamic	**Category D** Very Dynamic
0.15 to 0.25	HVGA	HVGA	HVGA	HVGA
0.25 to 0.5	VGA	VGA	HVGA	HVGA
0.5 to 1	WSVGA	WSVGA	VGA	HVGA
1 to 1.5	WSVGA	WSVGA	VGA	HVGA
1.5 to 2	WSVGA	WSVGA	VGA	VGA
2 to 3	720p	720p	720p 3/4	VGA
3 to 4	1080p	1080p 3/4	720p 3/4	VGA
4 to 5	1080p	1080p 3/4	720p	VGA
5 to 6	1080p	1080p 3/4	720p	720p 3/4
6 to 7	1080p	1080p	720p	720p 3/4
7 to 8	1080p	1080p	720p	720p

4. Conclusion

In this paper, we described an objective measurement of the video quality in a multi-rate and multi-resolution environment, and derived recommendations for the optimal resolutions at any given bit rate for various kinds of HD sources.

With optimized scaling filters and resolution, combined with Extreme compression, Envivio 4Caster C4 is the best platform for multi-rate and multi-profile video delivery. Coupled with the support of HTTP streaming technologies, Envivio solution delivers a truly Premium Internet TV experience.

5. Appendix: Multi-resolution in 4Caster C4

Envivio keeps adding optimized resolution to optimize the visual experience on new devices.

A list of supported profiles is indicated below. The resolutions have been grouped in categories of similar properties, corresponding to the ones used in the study.

Output Resolution Group	Resolution	Max Frame Rate	Resolution Name
720p60/1080i	1920x1080	30	1080i
720p60/1080i	1440x1080	30	
720p60/1080i	1280x1080	30	
720p60/1080i	1280x720	60	720p
720p60/1080i	960x720	60	
720p30	854x720	60	
720p30	848x720	60	
720p30	640x720	60	
720p30	960x1080	30	
720p30	1280x720	30	
WSVGA	960x720	30	720p30 3/4
WSVGA	854x720	30	
WSVGA	960x640	30	
WSVGA	848x720	30	
WSVGA	1024x576	30	
WSVGA	960x544	30	
WSVGA	960x540	30	
WSVGA	640x720	30	
WVGA / SD	854x480	30	
WVGA / SD	704x576	30	
WVGA / SD	832x480	30	
WVGA / SD	800x480	30	
WVGA / SD	720x480	30	
WVGA / SD	480x720	30	
WVGA / SD	704x480	30	
WVGA / SD	544x576	30	
VGA	640x480	30	VGA
VGA	528x576	30	
VGA	704x396	30	
VGA	480x576	30	
VGA	544x480	30	

Output Resolution Group	Resolution	Max Frame Rate	Resolution Name
VGA	528x480	30	
VGA	640x360	30	
VGA	480x480	30	
VGA	352x576	30	
HVGA	480x360	30	
HVGA	480x352	30	
HVGA	352x480	30	
HVGA	480x320	30	HVGA
HVGA	512x288	30	
HVGA	480x270	30	
HVGA	384x288	30	
HVGA	320x320	30	
HVGA	352x288	30	CIF
HVGA	416x240	30	
HVGA	400x224	30	
HVGA	360x240	30	
HVGA	352x240	30	
HVGA	384x216	30	
QVGA	320x240	30	QVGA
QVGA	320x192	30	
QVGA	320x180	30	
QVGA	240x240	30	
QVGA	320x176	30	
QVGA	304x168	30	
QVGA	240x192	30	
QVGA	240x176	30	
QVGA	220x176	30	
QVGA	240x160	30	SQVGA
QVGA	192x192	30	
QCIF	176x144	30	QCIF
QCIF	176x128	30	
QCIF	176x120	30	
QCIF	144x140	30	
QCIF	160x120	30	
QCIF	160x90	30	
QCIF	128x96	30	SQCIF
QCIF	144x80	30	
QCIF	96x96	30	