|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG16 WP3 and ISO/IEC JTC1/SC29/WG11**  6th Meeting: Torino, IT, 14-22 July, 2011 | Document: JCTVC-F292 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Recommendations for evaluation of scalable coding** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Informational | | |
| *Author(s) or Contact(s):* | Jill Boyce, Adeel Abbas, Danny Hong, Wonkap Jang 433 Hackensack Ave. Hackensack, N.J. 07601 USA | Tel: Email: | +1 (201) 289-8597  jill@vidyo.com |
| *Source:* | Vidyo, Inc. | | |

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

# Abstract

The contribution provides input to the definition of test and evaluation conditions for the planned scalability extension to HEVC, for coding efficiency comparison to simulcast and single layer coding. The same raw experimental data provided in the JCTVC-F290 contribution is presented using three different analysis methods. The contribution recommends one analysis method, in which the simulcast high resolution and scalable enhancement layer Y-PSNR values are matched.

# Introduction

Some discussion has taken place on the mpeg-hevc-ext reflector regarding scalable coding assessment methods. An interest has been expressed in comparing scalable coding performance with simulcast and single-layer coding. The draft requirements document [1] incorporates a method proposed by R. Sjöberg, originally described in [2], for computing scalability gain and cost vs. simulcast. That method is directed towards computations of bitrate reduction and coding gain for individual operating points, rather than a summary over a range of rates, as is done in the BD-PSNR and BD-rate calculations.

During the development of SVC, experimental test conditions typically specified required bitrates for particular resolution/frame rate operating points. JCT-P205 provides an example of a core experiment definition. The experimental results would frequently be presented as bitrate/PSNR curves for the proposed method and the scalability anchor. Comparisons with simulcast and single layer coding were not provided in individual contributions, and were performed using subjective testing at the end of the process.

The current HEVC test conditions [3] provide an Excel spreadsheet which does not include bitrate/PSNR curves, but presents tables of BD-rate and BD-PSNR values for each of the three color components, Y, U, and V, calculated using 4 QP operating points per sequence. Running a complete set of test results using the HM software [4] is very time and CPU consuming, given the large number of test conditions and sequences, including some at sequences at very high resolutions. The HM software does not currently contain rate control functionality.

# Discussion

For simplicity, discussion will be limited to the 2-layer spatial scalability case, although any resolution ratios between the layers can be considered.

Until anchors for scalable coding can be established (and perhaps even after that point) it is most instructive to compare the performance of scalable coding to simulcast and single layer encoding.

In the scalable case, we refer to a base layer and an enhancement layer. In the simulcast case, we refer to a low resolution and a high resolution, where the low resolution is the same resolution as the base layer, and the high resolution is the same resolution as the enhancement layer which is also the same as the single-layer resolution.

## Rate control/QP selection

The lack of rate control in the HM, as well as the amount of time required to run experiments for the entire common test condition test set, make setting target bitrates impractical.

We recommend to maintain a fixed QP differential between the base and enhan layers, and suggest using the specific value of the base layer QP value 2 below the enhancement layer’s QP. Using a good quality base layer tends to make encoding of the enhancement layer more efficient, improving the overall scalability efficiency. The impact on relative bitrates and PSNR values with a fixed QP delta between the layers will be highly sequence dependent, but using a QP delta of 2 tends to make the enhancement layer bitrate equal to or higher than the base layer bitrate in the experiments.

We recommend using the identical coded scalable base layer bitstream as the simulcast low resolution layer bitstream. We recommend using the same QP for the scalable enhancement layer and simulcast high resolution.

## BD-rate calculation

When performing the BD-rate/BD-PSNR calculations with 4 QP operating points for scalable and simulcast coding, it is not clear what are the appropriate bitrate and PSNR inputs to the calculations.

In our scalability experiments, in the vast majority of cases, when a scalable enhancement layer and simulcast high resolution are encoded at the same QP, the simulcast high resolution bitstream has both a higher PSNR and a higher bitrate than the scalable enhancement layer. In [1], performance metrics are proposed where the quality of the scalable enhancement layer and the simulcast high resolution are the same.

We describe several possible methods of using the BD-rate calculation for simulcast and scalability comparisons, and make a recommendation of a preferred method.

### Method 1: Use the same base layer and same QP for both simulcast and scalability, compute BD-rate/PSNR using only enhan/high res

In this method, only the scalable enhancement layer and simulcast high resolution layer experimental data points and their bitrates and PSNR values are used in the BD-rate/BD-PSNR calculations, and the base layer is neglected. The potential impact of the difference in quality between the scalable enhancement layer and simulcast high resolution on the base layer is neglected.

Considering only the enhan/high res layer bitrates and not including the base layer bitrate in the calculations tends to magnify the coding gains/losses, because a smaller denominator is used. However, the changes in bitrate do correspond more directly to the changes in PSNR of the higher resolution layer.

Table 1 shows results for the experiments in JCTVC-F290 using this method, where the inputs to the BD-rate calculation are the PSNR values of the scalable enhancement layer and simulcast high resolution, and the bitrates of the scalable enhan layer and the full res layers.

However, this method cannot be used to compare with single layer coding, because in the scalable case the enhancement layer only bitrate does not represent a decodable bitstream. So, we recommend not using this method, and instead utilizing the total bitrates of the two layers, as described in Methods 2 and 3.

### Method 2: Use the same base layer and same QP for both simulcast and scalability, compute BD-rate/PSNR using both layers

In this method, the total bitrates of both layers are considered, and the difference in quality between the scalable enhancement layer and simulcast high resolution is neglected. Table 2 shows results for the experiments in JCTVC-F290 using this method, where the inputs to the BD-rate calculation are the PSNR values of the scalable enhancement layer and simulcast high resolution, and the bitrate of the scalable base + enhan layer and the simulcast low + full res layers.

This method tends to disadvantage scalable coding, because it uses the same base layer, even though the high resolution qualities differ, and hence the low res layer quality/bitrate would also be expected to differ, corresponding to the quality changes of the high res layer.

The U and V BD-rate values sometime have strange results, showing losses even with the Y BD-rate vs. shows large gains. Even in non-scalable coding, comparing entire sequence bitrate changes with changes in chroma PSNR is somewhat strange. In the scalable coding case, it is not clear that the U and V BD-rate values are particularly meaningful.

### Method 3: Use the same base layer, adjust the simulcast high res bitrate along the simulcast high res curve to match the scalable enhan layer PSNR

This method makes an adjustment to match the quality of the simulcast high resolution and scalable enhancement layers, which the authors consider to be most appropriate when an identical scalable base/ simulcast low res layer is used.

Because the scalable enhancement layer is directly dependent on the base layer, and the simulcast high resolution bitstream is independent of the low res bitstream, we recommend to adjust the simulcast high resolution operating point for input to the BD-PSNR calculation, by curve fitting, to match the simulcast high resolution layer Y-PSNR to the scalable enhancement layer Y-PSNR.

It has been observed in discussions on the jct-vc reflector about BD-rate calculations that for many sequences, bitrate and Y PSNR have a relationship that is approximately linear between log(bitrate) and Y PSNR. We propose estimating intermediate (bitrate, PSNR) points in the simulcast curve by linear interpolation in the log(bitrate) vs Y PSNR domain.

Figure 1 shows an example, for the Kimono Sequence, Low Delay LoCo from JCTVC-F290. Figure 1(a) show the normal bitrate/PSNR curves for the scalable enhan layer in blue and simulcast high res layer in red, as well as the intermediate estimated points on the simulcast curve to match the scalable PSNR points, shown as a separate curve in green. Figure 1(b) plots the same curves on a log(bitrate) scale. The adjusted simulcast bitrate points were calculated by a linear interpolation between the two closest experimental points, using a piece-wise linear approximation.

The four adjusted simulcast points were used in the BD-rate/BD-PSNR calculation, with the base layer bitrate added to both the scalable and simulcast points. Table 3 shows the experimental results from JCTVC-F290 calculated using this method.

Because the simulcast bitrate curve fitting was done for the Y PSNR, neglecting the U and Vcomponents, the U and V BD-rate calculations are not very meaningful, and almost match the Y BD-rate values.

The comparison of scalable coding with single-layer coding is unchanged from Method 2, as only the simulcast operating points are adjusted.



**Figure 1(a). Bitrate vs. PSNR curves for Kimono, Low Delay LoCo from JCTVC-F290**

****

**Figure 1(b). Log(Bitrate) vs. PSNR curves for Kimono, Low Delay LoCo from JCTVC-F290**

.



**Table 1. Method 1 Experimental Results**

****

**Table 2. Method 2 Experimental Results**

**Table 3. Method 3 Experimental Results**

### Results of the methods

Which method is used to determine the inputs to the BD-rate/BD-PSNR calculations can impact the reported bitrate savings of a proposed scalability method vs. simulcast. For example, for the Random Access HE results, Method 1 shows an average bitrate reduction of 23.8%, vs. 14.8% for Method 2, and 16.1% for Method 3. No reasonable comparison with single-layer coding can be made for Method 1. Methods 2 and 3 yield the same scalable coding loss of 20.5% for Random Access HE.

# Recommendation

We recommend using Method 3 for comparing scalable coding performance with simulcast and single-layer coding. The experimental results in JCTVC-F290 were presented using this method. The most accurate comparison can be made when the differences in quality of the scalable enhancement layer and simulcast full resolution layers are considered and adjusted accordingly. Because of long run-times for the HM software and common test conditions, it is impractical to expect the encoder to match the quality of the scalable enhancement layer and simulcast full resolution layers precisely, so estimation methods may be employed.

# References

[1] A. Luthra, “WG11 AHG on requirements for extension of HEVC in the area of scalable coding” “Recommended Requirements and Test Points for Scalable Coding Extensions”, version 7, distributed on mpeg-hevc-ext reflector 23 June, 2011.

[2] Thomas Rusert, Rickard Sjöberg, Clinton Priddle, “Analysis of H.264/AVC Scalable Video Coding for Video Delivery to Heterogeneous Terminals”, MobiMedia 2010.

[3] F. Bossen, “Common Test Conditions and software reference configurations,” JCTVC-E700, Geneva, CH, March, 2011.

[4] T. Wiegand, W.-J. Han, B. Bross, J.-R. Ohm and G. J. Sullivan, “WD3: Working Draft 3 of High-Efficiency Video Coding,” JCTVC-E603, Geneva, CH, March, 2011.