|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  13th Meeting: Incheon, KR, 18–26 Apr. 2013 | Document: JCTVC-M0199 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **On flexible stream switching** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Patrick Lopez,  Philippe Bordes, Franck Hiron  Technicolor  975avenue des champs blancs – CS17616 35576 Cesson-Sévigné Cedex, France | Tel: Email: | +33 2 99 27 32 64 patrick.lopez@technicolor.com |
| *Source:* | Technicolor | | |

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

# Abstract

Stream switching is a traditional technique used for adapting the content to the available bandwidth. Basically, SHVC can be considered as an alternative to Stream-Switching since the content adaptation is performed by the adding/dropping of Enhancement layers NALUs. It is asserted both techniques have their pros and cons.

This contribution proposes to extend SHVC to support enhancement of already decoded pictures in the base layer DPB in order to enable stream-switching at any points (not only CRA) and hence claims to increase the stream-switching flexibility. It is proposed sequence level flags for SHVC to indicate the stream switching operation and describe the decoding process when stream switching occurs. Experiment results are provided which compare the performance of stream switching with and without the proposed method.

# Introduction

Stream switching basic technique consists in switching from one (offline encoded) stream to another to accommodate the bandwidth variation. This is a traditional technique currently used and deployed by many VOD services providers.

However, in the case of switch from low quality (LQ) to high quality (HQ), after the switch occurs, it happens the reference pictures obtained with the LQ stream (further used in inter-prediction to decode the subsequent pictures in decoding order) are not identical to the reference pictures with same POC contained in the HQ stream. Visual artifacts may appear, which disappear only when an IDR or CRA picture is inserted in the bitstream.

It is proposed to transmit additional data at switching point to increase the quality of the LQ reference pictures and reduce the drift artifacts.

It is proposed this additional data is encoded using SHVC intra-BL prediction, as previously suggested in [1] for ARC purpose.

Practically, in case of DASH protocol using pull scheme, for example, when the decoder application decides to switch to another stream, since it knows the current DPB composition (reference pictures list), it will ask for the corresponding additional data (enhancement NALUs) to up-date the reference pictures in the DPB.

Note the proposed technique supposes that both HQ and LQ streams have the same GOP structure and POC references.

# Proposal description

In the proposed solution, additional data are transmitted at switching point, aiming at updating the pictures of layer#0 in the DPB used as reference to decode coded pictures of the HQ stream.

Figure 1 shows a basic stream switching without insertion of any data.

PLQ,n-2

PLQ,n-1

PLQ,n

PHQ,n+3

PHQ,n+1

PHQ,n+2

Figure : Basic stream switching

Figure 2 shows how the proposed stream switching technique is implemented. For simplification of the drawing, a low delay P profile has been assumed here.

Layer #0

Layer #1

PLQ,n-2

PLQ,n-1

PLQ,n

PHQ,n+3

PHQ,n+1

PHQ,n+2

EHQ,n

EHQ,n-1

PLQ,n-3

Reference Pictures

In the DPB

Figure : Proposed stream switching

PLQ, n : nth picture encoded in LQ quality

PHQ, n : nth picture encoded in HQ quality

EHQ,n-j : information needed to reconstruct at HQ quality the n-jth reference picture.

The pictures EHQ,n-j are encoded in SHVC and are transmitted with layerId = 1. The coding modes for these pictures are constrained to Intra and IntraBL.

On reception of the EHQ,n-j pictures, the decoder decodes them, using the PLQ,n-j when the IntraBL mode is chosen. The decoded picture is then written in the reference picture list of Layer #0 and replaces PLQ,n-j. No picture is written on output of Layer #1 which only contains information for stream switching.

# Proposed syntax and Semantics

The syntax for description of this mechanism can be implemented at the VPS level [2]. Two bits are added into the vps\_extension which indicate whether :

* all non-base layer VCL NALUdon’t use temporal prediction,
* the base layer DPB must be updated with upper layer decoded content

|  |  |
| --- | --- |
| vps\_extension( ) { |  |
| … |  |
| for( i = 1; i <= vps\_max\_layers\_minus1; i++ ) |  |
| for( j = 0; j < i; j++ ) |  |
| **direct\_dependency\_flag**[ i ][ j ] | u(1) |
| **intra\_or\_intrabl\_pred\_only\_flag** | u(1) |
| **base\_layer\_update\_for\_non\_base\_layer\_flag** | u(1) |
| } |  |

**intra\_or\_intrabl\_pred\_only\_flag** equal to 1 indicates that any coded picture with layer\_id greater than 0 does not use inter prediction, nor inter-layer prediction except IntraBL. intra\_or\_intrabl\_pred\_only\_flag equal to 0 indicates that any picture with layer\_id greater than 0 may use inter prediction as well as any inter-layer prediction mechanism.

**base\_layer\_update\_for\_non\_base\_layer\_flag** equal to 1 indicates that any coded picture with layer\_id greater than 0 must be decoded, and replaces the picture which has the same POC into the DPB of layer\_id equal to 0. base\_layer\_update\_for\_non\_base\_layer\_flag equal to 0 indicates that the DPB of picture with layer\_id equal to 0 is not modified when decoding picture with layer\_id not equal to 0.

# Experiments

The following Figure 3 shows the result obtained on ClassB BasketballDrive sequence. Random Access case has been considered, with hierarchical GOP of length 8 and Intra Period of 128. QP of LQ and HQ are respectively 32 and 26. The switch occurs after the POC #32.

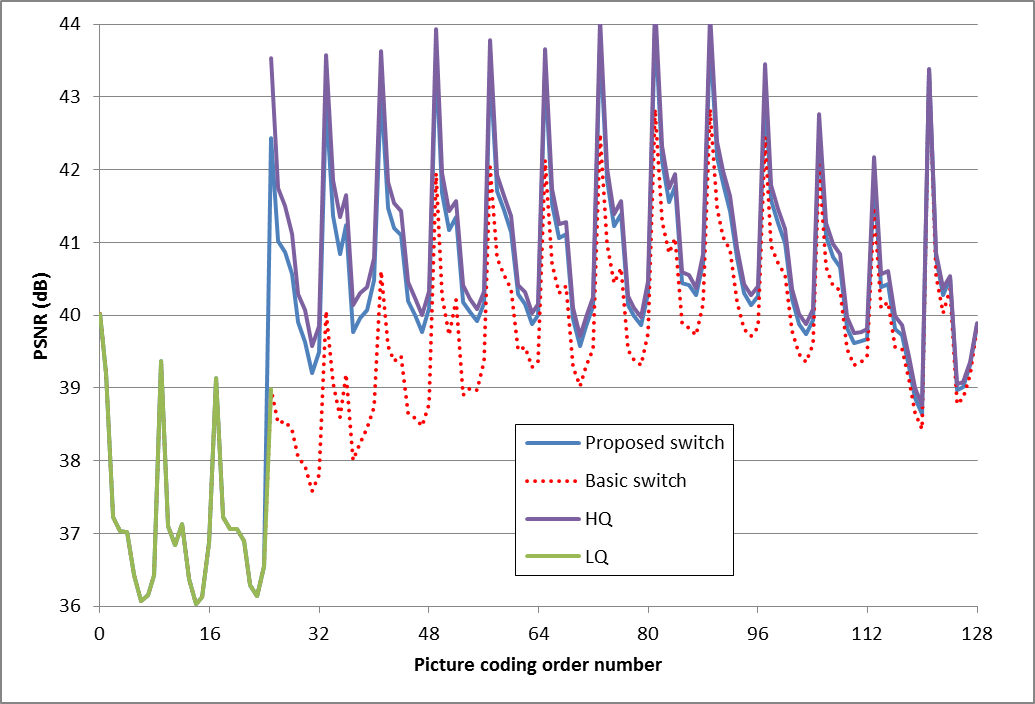


Figure : PSNR with the 2 switching techniques

Figure 4 shows the degradation of the obtained PSNR in both cases, the basic switch with no additional data and the proposed switch. The expected “ideal” PSNR is the LQ PSNR until the POC#32 and then the HQ PSNR.

The proposed switch exhibits an average gain of 1 dB vs the basic switch, when the PSNR is computed from the switching point to the next RAP (POC #128). This gain is higher on the first frames after the switch : 2.1 dB and 1.6 dB respectively over the first and the two first GOP after the switch.

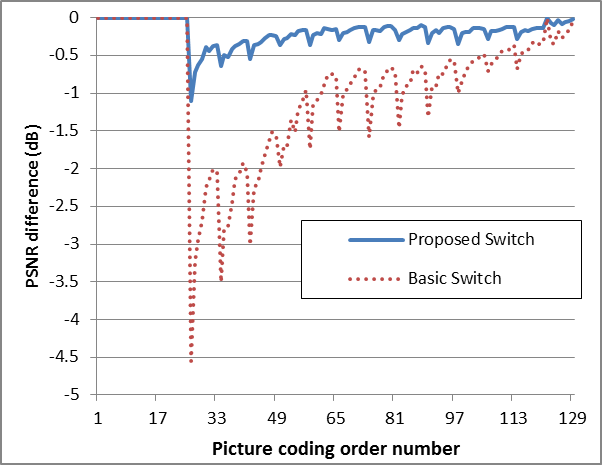


Figure : Degradation of PSNR after switch

The cost of the additional data inserted for the switch must be compared to the alternative which consists in regular insertion of IDR/CRA frames.

The following picture shows that a gain of 1 dB over the basic switch scheme is obtained when the cost of additional data is equivalent to 1.8 CRA frames.

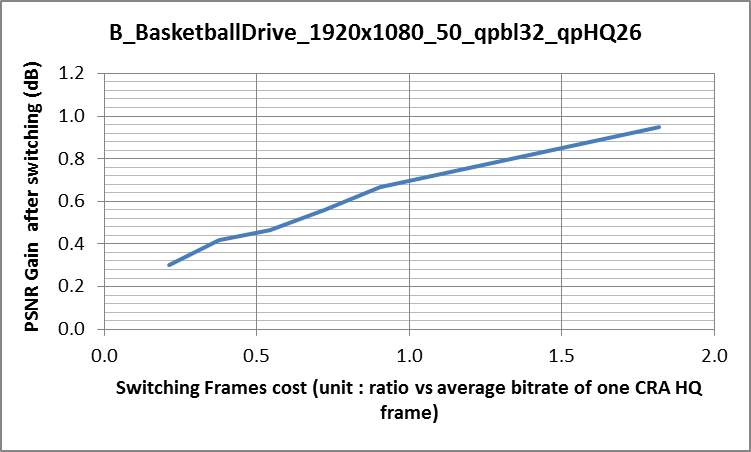


Figure : Rate distorsion curve of switching frames

# Conclusion

A technique of flexible stream switching has been presented. It relies on the usage of SHVC with two layers, the enhanced layer being used to convey the high quality coded pictures used as references in the DPB.

This technique has been tested on a limited set of sequences and QP values. A gain of 1 dB is reported versus a basic switch. For this gain of 1dB , the cost is equivalent to 1.8 CRA frames.

It is proposed the syntax enabling the flexible stream switching is added to the HEVC design.

# References

1. K.Ugur and al, “Adaptive resolution change with SHVC”, JCTVC-L0119, Geneva, January 2013.
2. J. Chen and al, “SHVC Draft Text 1”, JCTVC-L1008, Geneva, January 2013.

# Patent rights declaration(s)

**Technicolor may have current or pending patent rights relating to the technology described in this contribution and, conditioned on reciprocity, is prepared to grant licenses under reasonable and non-discriminatory terms as necessary for implementation of the resulting ITU-T Recommendation | ISO/IEC International Standard (per box 2 of the ITU-T/ITU-R/ISO/IEC patent statement and licensing declaration form).**