|  |  |
| --- | --- |
| **Joint Collaborative Team on Video Coding (JCT-VC)**  **of ITU-T SG 16 WP 3 and ISO/IEC JTC 1/SC 29/WG 11**  14th Meeting: Vienna, Austria, AT, July 29 – Aug 2, 2013 | Document: JCTVC-N0269 |

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **HLS: Non-significant slice segments with tiles for single layer HEVC extensions** | | |
| *Status:* | Input Document to JCT-VC | | |
| *Purpose:* | Proposal | | |
| *Author(s) or Contact(s):* | Cheung Auyeung  1730 N. First Street San Jose, CA 95112, USA | Tel: Email: | +1-408-352-4086 [Cheung.Auyeung@am.sony.com](mailto:Cheung.Auyeung@am.sony.com) |
| *Source:* | Sony Electronics Inc. | | |

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

# Abstract

The motion-constrained tile sets SEI message in Range Extensions text specification: Draft 3 can be used to signal that a CVS is comprised of one or more regions of interest in the form of independently decodable motion-constrained tile sets. A user can interactively select and decode a motion constrained tile set without decoding the other tiles. This contribution proposes syntax and semantics that support the extraction of independently decodable motion-constrained tile sets to form a new bitstream without transcoding at the CTU level for streaming of the regions of interest in applications such as interactive UHDTV application, dynamic high-quality zoom-in application, interactive on-demand, e-learning, smart surveillance, and etc.

# Introduction

In version 1 of the HEVC standard [1], when tiles\_enabled\_flag is equal to 1, each picture is constructed of rectangular regions known as tiles. In Range Extensions text specification: Draft 3, these tiles can be collected to form a collection of tile sets as regions of interest. These tile sets can be indicated as independently decodable by the motion-constrained tile sets SEI messages [2].

Such HEVC independently decodable tile structure can be applied to interactive tiled streaming [3,4,5]. In the example of tiled streaming in Figure 1, a panorama video is encoded with independently decodable tiles and in multiple resolution layers in the server side. Figure 1 shows one of the resolution layers in the server side at 8K×4K. Different viewer may request the server to send different region of interest from different resolution layers of the panorama. In Figure 1, one client requests a region to be viewed on the client’s HD mobile device from one of the resolution layer. Another client requests a different region to be viewed on the client’s 4K TV from the same resolution layer. In general, the clients may selectively to view any region of interests at one of the several spatial resolutions. Figure 1 shows user interactivity with the server at one of the resolution layer.

Tiled streaming use cases include interactive UHDTV application, dynamic high-quality zoom-in application, interactive on-demand e-learning, smart surveillance, and etc. These use cases of tiled streaming are currently being studied in the core experiments of DASH [6].

For the case that the pictures are encoded only with independently decodable tiles, this contribution proposes syntax and semantics to provide the support that a region of interest can be interactively extracted from a compressed bitstream to form a valid bitstream to send to a user without transcoding at the CTU level.

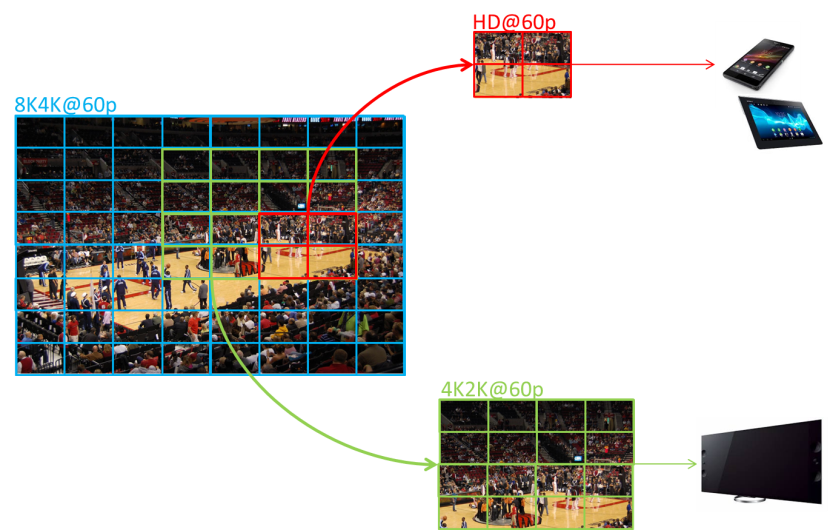


Figure Example of tiled streaming application



Figure : User requested tiles are extracted and send to a client by replacing irrelevant tiles by black tiles. It is not necessary for a client to decode the irrelevant tiles at the CTU level.

# Proposal

To support tiled streaming without transcoding at the CTU level, each resolution layer of the video is encoded with independently decodable tiles. Figure 2 shows tiled streaming from one of the resolution layer. As shown in Figure 2, users are interested to view different regions of interest. To send a region of interest to a user and to minimize transmission bandwidth, only tiles in the user’s region of interest are extracted and sent to the user. The tiles not in the region of interest are not sent to the user. From a user perspective, the tiles not in the region of interest are replaced by black tiles.

The server send the region of interest requested by a user by replaceing the slice segments outside the region of interest by non-significant slice segments to from a new bitstream to the user. In this way, transcoding at the CTU level is avoided.

To enable non-significant tile for tiled streaming, this contribution proposes the following modifications and coding conditions for non-significant slice segment.

# Set tiles\_enabled\_flag in pic\_parameter\_set\_rbsp() to 1.

* Each tile is comprised of one or more slice segments.

# Set slice\_segment\_header\_extension\_present\_flag in pic\_parameter\_set\_rbsp() to 1.

* Motion-constrained tile sets SEI messages are used to indicate that every tile is independently decodable.

# Modify slice\_segment\_header() to enable the functionality of non-significant slice segment with tiles.

# Modify slice\_segment\_layer\_rbsp() to enable the functionality of non-significant slice segment with tiles.

## Proposed modification to slice segment header

Yellow indicates proposed changes from the HEVC standard [1] and the Range Extensions text specification: Draft 3 [2].

|  |  |
| --- | --- |
| slice\_segment\_header( ) { | Descriptor |
| … |  |
| … |  |
| if( slice\_segment\_header\_extension\_present\_flag ) { |  |
| if( tiles\_enabled\_flag ) { |  |
| **non\_significant\_slice\_segment\_flag** | u(1) |
| if( non\_significant\_slice\_segment\_flag ) |  |
| **slice\_segment\_size\_minus1** | ue(v) |
| } |  |
| **slice\_segment\_header\_extension\_flag** | u(1) |
| if (slice\_segment\_header\_extension\_flag) { |  |
| **slice\_segment\_header\_extension\_length** | ue(v) |
| for( i = 0; i < slice\_segment\_header\_extension\_length; i++) |  |
| **slice\_segment\_header\_extension\_data\_byte**[ i ] | u(8) |
| } |  |
| } |  |
| byte\_alignment( ) |  |
| } |  |

If non\_significant\_slice\_segment\_flag is not signaled in a bitstream, it has a default value of 0.

**slice\_segment\_size\_minus1** is the number ofCTU in the non-significant slice segment minus 1.

## Proposed modification to slice segment layer

When a slice segment is non-significant, the slice\_segment\_data( ) syntax structure is not present in the coded bitstream.

|  |  |
| --- | --- |
| slice\_segment\_layer\_rbsp( ) { | Descriptor |
| slice\_segment\_header( ) |  |
| if (!non\_significant\_slice\_segment\_flag) |  |
| slice\_segment\_data( ) |  |
| rbsp\_slice\_segment\_trailing\_bits( ) |  |
| } |  |

## Proposed decoding process for non-significant slice segment with tiles

All pixels in all CTU in a non-significant slice segment are not required to be decoded. When it is decoded, it shall have the value of 0.

# Conclusion

This contribution proposes extensions to the slice segment header to define non-significant slice segment with tiles. Non-significant slice segment with tiles can be used in tiles streaming to send a region of interest (ROI) extracted from a compressed bitstream encoded with independently decodable tiles by replacing the slice segments outside the ROI by non-significant slice segments.

# Reference

1. B. Bross, W.-J. Han, J.-R. Ohm, G. J. Sullivan, Y.-K. Wang, and T. Wiegand (Eds.), “High efficiency video coding (HEVC) text specification draft 8”, JCTVC-L1003\_v34, 12th Meeting: Geneva, CH, 14–23 Jan. 2013.
2. D.Flynn, J.sole, T.Suzuki, “High Efficiency Video Coding (HEVC) Range Extensions text specification: Draft 3”, JCTVC-M1005, 13th Meeting: Incheon, KR, 18–26 Apr. 2013.
3. Franck Denoual, Hervé Le Floch, Frédéric Mazé, Eric Nassor, Nael Ouedraogo, Cyril Concolato, Jean Le Feuvre, “Interactive ROI streaming with DASH”, MPEG2013/M29232, Incheon, South Korea, April 2013.
4. Emmanuel Thomas, Rob Koenen, “Spatially segmented content description”, MPEG2013/m28883, Incheon, South Korea, April 2013.
5. Sally Hattori, Ohji Nakagami, Teruhiko Suzuki, “**H**LS: Extensions to Motion-constrained tile sets SEI message”, JCTVC-Nxxxx, 14th Meeting: Vienna, Austria, AT, July 29 – Aug 2, 2013.
6. A.Giladi, T.Stockhammer, EE coordinators, “Descriptions of Core Experiments on DASH Amendement”, N13499, MPEG2013, Incheon, Korea, April 2013.

# Patent rights declaration(s)

**Sony Corp. and Sony Electronics Inc. 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).**