**INTERNATIONAL ORGANIZATION FOR STANDARDIZATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC 1/SC 29/AG 5**

**MPEG VISUAL QUALITY ASSESSMENT**

**ISO/IEC JTC 1/SC 29/AG 5 N105**

**October 2023, Hannover, DE**

|  |  |  |  |
| --- | --- | --- | --- |
| *Title:* | **Report of verification test on VVC multi-layer coding: Content layering** | | |
| *Status:* | Output document approved by JVET | | |
| *Purpose:* | Report | | |
| *Author(s) or Contact(s):* | Shunsuke Iwamura Philippe de Lagrange Mathias Wien | Tel: Email: | [iwamura.s-gc@nhk.or.jp](mailto:iwamura.s-gc@nhk.or.jp) [philippe.delagrange@interdigital.com](mailto:philippe.delagrange@interdigital.com)  [wien@lfb.rwth-aachen.de](mailto:wien@lfb.rwth-aachen.de) |
| *Source:* | Verification Test Coordinators | | |

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

# Abstract

This document reports the verification test results of VVC multi-layer coding according to the “Content layering” category of the verification test plan JVET-AD2021 that was issued in the 30th JVET meeting. Content layering is one of the useful use cases, where main content and multiple sub-content are coded with multi-layer coding into a single bitstream and delivered to the users. By selecting a layer to be decoded, the user can freely select whether to view the main content or the sub-content. The functional test using three test sequences with sub-contents was carried out using a real-time VVC decoder with multilayer coding capability, to demonstrate the effectiveness of content layering application.

# Introduction

The main design idea of the VVC standard was its high versatility as well as improvement in compression capability relative to its predecessor, the HEVC standard. One of its versatility is multi-layer coding capability. This document is the first verification report of a new application called “content layering” which is enabled by multi-layer coding. It also reports the results of a functional test demonstrated in 31st JVET meeting.

# Functional test setting

## Test sequence creation

The three test sequences used in the functional test are 4K, 60Hz and filmed by NHK. Location footage, football game, and children’s program are selected as common program. For each content, sub-content that matched the context of each program was created. For location footage, the main content was down-converted and sign language commentary was inserted. For football game, another camera footage focused on the home team and away team was overlaid in a small window on top of main content. In children’s program, dance instruction video is overlaid on the down-sized main content.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Content1: Location footage  Main: Report from library Sub: Main with sign-language | Content2: Football game  Main: Main camera Sub1:  Main with Home team view Sub2:  Main with Away team view | Content3: Children’s program  Main: Music video Sub1: Main with dance instruction by a boy Sub2: Main with dance instruction by a girl |
|  |  |  |  |
| L0 | Ein Bild, das Szene, Text, Regale, Geschäft enthält.  Automatisch generierte Beschreibung |  | Ein Bild, das Cartoon, Kleidung, Animierter Cartoon, Animation enthält.  Automatisch generierte Beschreibung |
| L1 | Ein Bild, das Text, Szene, Regal, Kleidung enthält.  Automatisch generierte Beschreibung | Ein Bild, das Gebäude, Sportplatz, Stadion, Arena enthält.  Automatisch generierte Beschreibung | Ein Bild, das Text, Kleidung, Schuhwerk, Animierter Cartoon enthält.  Automatisch generierte Beschreibung |
| L2 |  | Ein Bild, das Gebäude, Sportplatz, Fußballstadion, Stadion enthält.  Automatisch generierte Beschreibung | Ein Bild, das Kleidung, Text, Cartoon, Kleid enthält.  Automatisch generierte Beschreibung |

## Test bitstream preparation

VTM-18.0 with content layering functionality is used for generating the bitstream. The main content is coded as base layer bitstream. For the enhancement layer, the sub-content with alpha channel information is overlaid on top of the locally decoded picture of the base layer and used as the input signal of the enhancement layer. Since most of the area of the enhancement layer is exactly same as the base layer, most of the block can be coded as “skipped” without RD optimization and only the region with overlay is coded with RD optimization.

Figure 1 illustrates the overview of encoding procedure of the content layering. Note that alpha channel information does not have to be transmitted to the decoder side since the input signal of the enhancement layer is the main content with overlay.



Figure 1. Content layering using multilayer coding

Figure 2 shows the alternative solution to realize the same functionality by using systems layer composition. Here, the main content, sub-content, and alpha channel information is separately encoded as single layer bitstreams and transmitted to the decoder (with position information if needed). At the decoder side, the main and sub-content are decoded separately and composited using the position information and alpha channel information.



Figure 2. Content layering using system layer composition

For some test sequences (Content1 and Content3), the sub-content is not just overlaid on top of the main content. The main content is down-scaled, and L-shaped sub-content is inserted in the margin. Inter-layer prediction can be effectively performed with scaling window functionality enabled. Down-scaling filter is the same as the filter used in RPR so that inter-layer prediction is perfectly performed in a similar way to the case of normal overlaying sub-content shown in figure 1. Figure 3 illustrates the overview of coding procedure of content layering for L-shaped sub-content.



Figure 3. Content layering with scaling window

## Encoding parameters

The following test conditions were used for the functional test.

1. Rate control
   1. The rate control is enabled for base and enhancement layer. The allocated bitrate for BL and EL is 20Mbps and 2Mbps.
2. Coding structure
   1. Random access, RA
      1. Intra refresh at approximately 1 second intervals.
      2. GOP size 32
      3. Picture reordering allowed.
      4. Inter-layer and intra-layer prediction allowed.

## Equipment setting

The coordination of the functional tests was done by the chairs of JVET AHG4 and AG5. Block diagram of the equipment for the test is illustrated in Figure 4. Multilayer bitstream capsuled in CMAF over MMT is output from the streamer and VVC decoder supporting multilayer decodes the bitstream in real time. The brief specification of VVC multilayer decoder is shown in Table 1.

Viewers can select which contents (main or sub) to play out by controlling by Tablet PC. VVC multilayer real-time decoder and CMAF/MMT streamer, signal converter for the functional test are prepared by NHK and 4K Monitor is prepared by ITU-T SG16.



Figure 4. Equipment for the functional test

Table 1. Specifications of multilayer VVC decoder

|  |  |
| --- | --- |
| Profile | Main 10, Multilayer Main 10 |
| Level | 6.3 |
| Number of supported layers | 4 layers |
| Input Format | CMAF (ISO/IEC 23000-19) on MMT (ISO/IEC 23008-1) |
| Output Format | Resolution: 1080p, 2160p, 4320p  Frame rate [Hz]: 59.94, 60, 119.88, 120 |

# Demonstration and discussion

Demonstration was conducted on July 19th at “M03b” in Mont-brilliant building. Approximately 20 participants attended the demonstration. A simple remote controller application was used so that participants could freely switch the content by tapping the buttons on the application in the tablet PC.

It was asked how rate control is handled. Rate control was reportedly done independently per layer. This aspect is considered to be future work to assert compliance with limits for the overall stream.

It was asked if how multiple versions of such a scheme would be supported, e.g., in a HTTP streaming scenario. This is considered to be a challenging encoding task and has not been investigated.

It was asked ho HDR metadata would be handled. In this case, the metadata would need to be updated for the respective enhancement layers to reflect the parameters for the final composition of the scene.

グラフィカル ユーザー インターフェイス, Web サイト

自動的に生成された説明

Figure 5. 4K monitor and remote-control-like tablet PC



Figure 6. Demonstration of content layering

# References

1. Recommendation ITU-T H.266 | International Standard ISO/IEC 23090-3 (2022), *Versatile Video Coding*.
2. S. Iwamura, P. de Lagrange, and M. Wien, “Verification test plan for VVC multilayer coding,” JVET-AD2021, June 2023.
3. S. Iwamura, D. Arai, S. Nemoto, and A. Ichigaya, ”AHG4: Preparation for content layering functional test,” JVET-AE0235, July 2023.

# Acknowledgements

The verification test coordinators wish to thank the organizations who contributed to and supported the VVC verification tests, particularly including NHK and ITU-T SG16.