|  |  |
| --- | --- |
| **ITU – Telecommunications Standardization Sector**STUDY GROUP 21 Question 6**Video Coding Experts Group (VCEG)**76th Meeting: 27 March – 4 April 2025, by teleconference | Document VCEG-BX20-v1 |

|  |  |
| --- | --- |
| Question: | 6/21 (VCEG) |
| Source: | Bart KroonChris Varekamp | Email: | bart.kroon@philips.comchris.varekamp@philips.com |
| Title: | **On the immersive requirements for the next generation video codec** |
| Purpose: | Requirements proposal |

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

# Abstract

The *Preliminary draft of use cases and requirements for potential next-generation video coding standard beyond VVC capability* [1], not yet available as [2] at the moment of writing, includes use cases and requirements for immersive applications. This contribution aims to improve those requirements that relate to immersive use cases.

# Introduction

The draft [1] has two immersive use cases that are of our interest, and we believe that they require novel view synthesis:

* Immersive applications, such as virtual, augmented and mixed reality for communication, interaction and entertainment.
* Representation of content to allow the rendering of a viewpoint from a limited specified perspective

There are also some adjacent use cases that are of our interest:

* Single- and multiview representations for medical applications
* Low-delay 3D representation for remote communication

With these use cases in mind, the following content types in [1] are relevant:

* Stereoscopic and multi-view content.
* Content with associated sample-based maps (e.g. for depth or alpha).
* Coding of alpha and/or depth components together with tristimulus colour components.

# Motivation

In our experience, it is useful to be able to have *multiple* ***partial******views******or planes*** *with optional depth information and camera parameters*.

There are two main arguments for the support of partial views or planes. First, from an efficiency perspective, when the number of cameras increases, it becomes less relevant to fully represent each of the views. Second, video regions often have a semantic meaning (object instance/category) that is useful for immersive playback. Partial views allow the representation of such semantically meaningful objects.

The additional information offered by an extra view or plane is concentrated in disocclusion regions, non-Lambertian (e.g. specular) regions and/or an increased field of view.

The combination of multiple image regions of multiple views or planes into a video frame is often called a video atlas. A next generation video codec with improved support for the coding of multiple disjoint image regions into a single frame has our interest in view of the mentioned use cases.

# Proposed changes

## Specific content types to be addressed

We suggest the term "occupancy" for a 1-bit alpha-like component, indicating that a pixel is present (occupant) or should be ignored (non-occupant):

 • Coding of alpha, occupancy and/or depth components together with tristimulus colour components.

We also suggest to add the following content type:

* Video atlases: video frames consisting of multiple distinct image regions from the same or different sensors.

## Bit depth and colour sampling

We suggest to add an example for "auxiliary components" that is relevant in some immersive video applications:

 • Stereoscopic 3D, multi-view, depth, alpha, and auxiliary components (e.g. object instance ID maps)

## Random access and "trick mode" support

We suggest to add "video atlases" as an enabling application:

 • The standard shall have support for spatial random access in the video bitstream for functionality such as extraction, transmission, and decoding of only one or more rectangular regions of the pictures, for efficiently enabling applications such as virtual reality, 360° video streaming, video atlases, and region-of-interest stream adaptation.

# References

[1] Igor Curcio, *Update on Preliminary draft of use cases and requirements for potential next-generation video coding standard beyond VVC capability*, MPEG/M71662, January 2025, Geneva (CH).

[2] *Update on Preliminary draft of use cases and requirements for potential next-generation video coding standard beyond VVC capability*, ISO/IEC JTC 1/SC 29/WG 02 N 0429, January 2025, Geneva (CH).

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