**INTERNATIONAL ORGANISATION FOR STANDARDISATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC1/SC29/WG11**

**CODING OF MOVING PICTURES AND AUDIO**

**ISO/IEC JTC1/SC29/WG11 MPEG/m50849**

**October 2019, Geneva, CH**

|  |  |
| --- | --- |
| **Source** | **Interdigital** |
| **Status** | **proposal** |
| **Title** | **New MIV Immersive Video content proposal** |
| **Author** | Renaud Doré, Gérard Briand, Julien Fleureau |

# Context

MPEG Video group has initiated investigations through a CfP on Metadata for Immersive Video and is currently in the Core Experiment phase.

At last meeting, a new content PoznanFencing has been introduced in the CTC [1]. It is a natural content with 5 pairs of perspective cameras looking inward to the center of a scene with 2 fencers, as illustrated in figure 1 below.

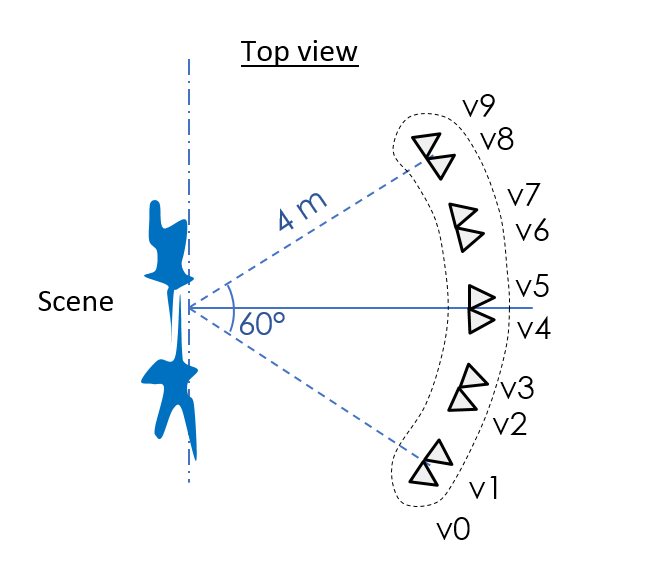


Figure : PoznanFencing rig setup

It is definitively a good thing to have such a new content for MIV investigations as it enriches the set of contents. Yet, a simple look of the depth grey video shows that the depth has big burst of inconsistency zones near the fencers, as illustrated by the figure below of a snapshot of frame 1 of v4 camera.

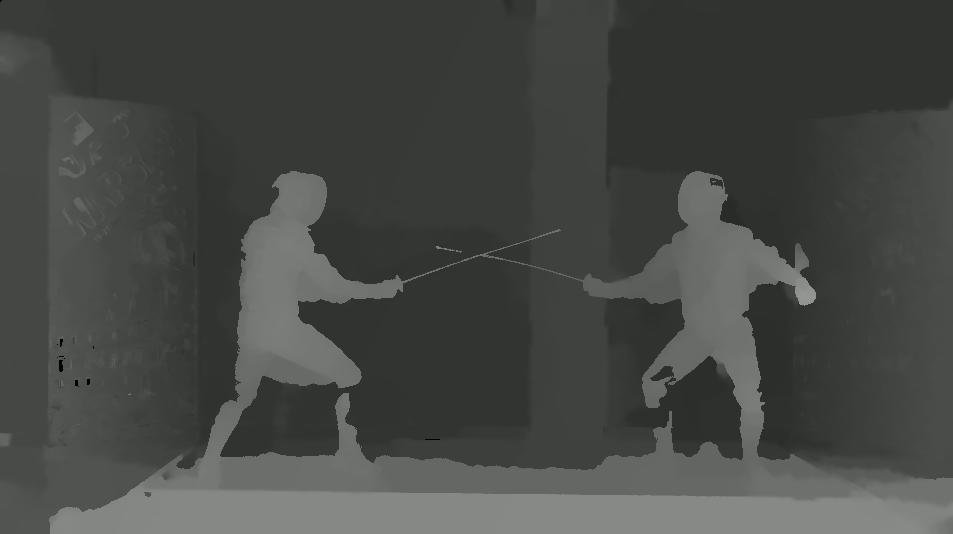


Figure : frame 1 depth of camera v4

As a matter of fact, the pose traces applied on the totality of the sources are not good at all. In these conditions, there is no possibility for the TMIV chain to output satisfactory pose traces on the decoder side, as it already hurts before MIV encoding. Cf [1] for the reference of the slight compressed video of these pose traces.

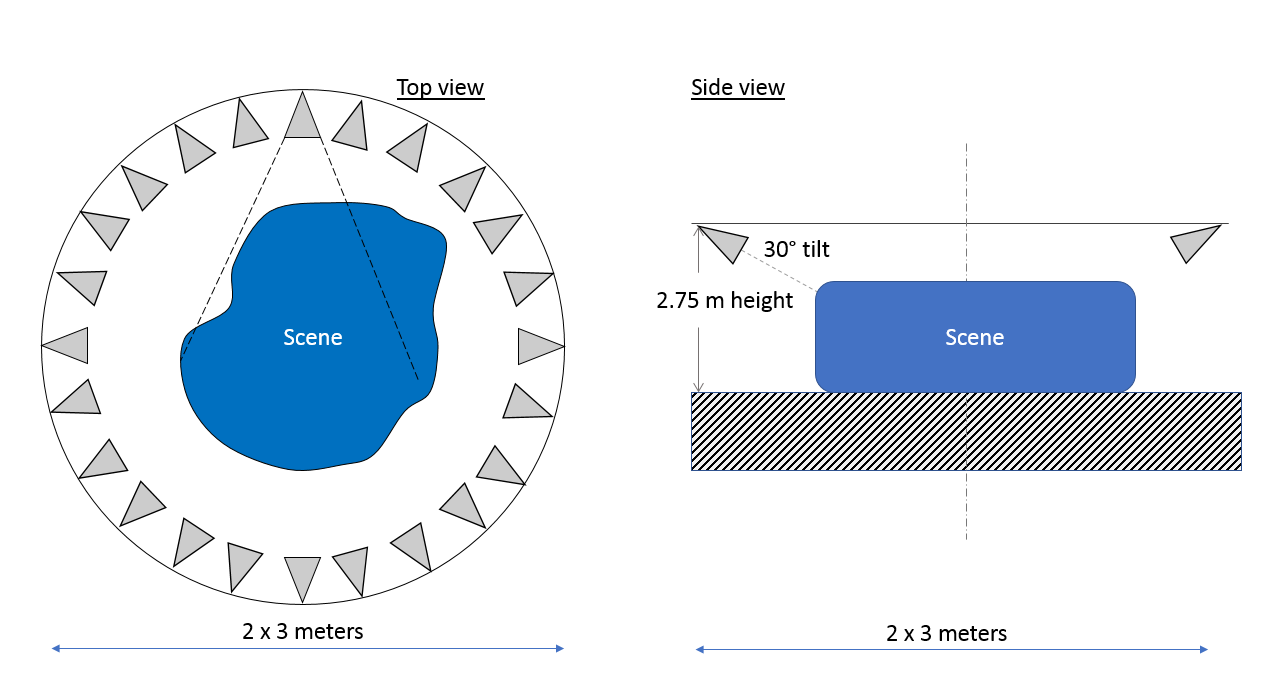
If we want the MIV to present acceptable results, the depth of PoznanFencing should be ideally updated, but it should be checked if this is possible.

Another possibility is to have another content with the same rig disposition with a perfect depth. It will be therefore possible to investigate the specificity of this rig by making abstraction of the depth quality - which is not the topic of this MIV working group - , and this is what we propose here.

## Proposing MuseumInward Content

In order to enable some investigations, we have made a new rig configuration of the semi-synthetic “TechnicolorMuseum” scene content and regenerated MVD inputs accordingly.

The rig is a ring of 24 half-ERP 2048x2048 cameras, equally disposed on a ring of 3m radius at height of 2.75m, and looking inward with a downward tilt of 30°.



This leads to the following image examples (for the texture):



Figure 3:from left ot right, V0, V1, V2



Figure : from left to right: v12, v16, v20



Figure :from left to right, v21,v22,v23

With such a content, as it is a linear camera disposition, it is possible to adjust the number of source views, and to adjust the json file to take only the first 6, 7, 8, 9 , 12, etc.. source views, corresponding to an arc of different length.

Our primary objective was at least to mimic PoznanFencing rig, considering that in this content, each cameras v0, v2, v4, v6, v8 is almost uselessly doubled by its stereo companion: this PoznanFencing rig is almost just like 5 cameras for a 60° arc only, meaning 1 camera every 15°. In MuseumInward, the cameras are also disposed every 15°.

# Conclusion

We recommend pushing this new content into the CTC for the next meeting iteration after having decided the number of successive source views. Interdigital will upload on MPEG server the number of source views which is decided for the MIV investigations. At minimum, we recommend taking at least 6 source views in order to be compared with PoznanFencing result, so that the specificity of this interesting arc-shape rig can be investigated while making abstraction of the depth quality.

# Reference

[1] J. Jung, B. Kroon, J. Boyce, Common Test Conditions for Immersive Video, ISO/IEC JTC1/SC29/WG11 MPEG/N18563, July. 2019, Goeteborg, Switzerland.