**** **ISO/IEC JTC 1/SC 29/WG 2 N54**

**ISO/IEC JTC 1/SC 29/WG 2**

**MPEG Technical requirements   
Convenorship: SFS (Finland)**

**Document type:** Output Document

**Title:** Common Test Conditions and Evaluation Methodology for Video Coding for Machines

**Status:** Approved

**Date of document:** 2021-01-15

**Source:** ISO/IEC JTC 1/SC 29/WG 2

**Expected action:** none

**Action due date:**  none

**No. of pages:** 3 (Including cover page)

**Email of Convenor:** igor.curcio@nokia.com

**Committee URL:** <https://isotc.iso.org/livelink/livelink/open/jtc1sc29wg2>

**INTERNATIONAL ORGANISATION FOR STANDARDISATION**

**ORGANISATION INTERNATIONALE DE NORMALISATION**

**ISO/IEC JTC 1/SC 29/WG 2**

**MPEG TECHNICAL REQUIREMENTS**

**ISO/IEC JTC 1/SC 29/WG 2 N54**

**Online – January 2021**

**Title: Common Test Conditions and Evaluation Methodology for Video Coding for Machines**

**Source: WG 2 MPEG Technical requirements**

**Status: Approved**

**Serial number: 20161**

**Editor(s): Shan Liu (Tencent),**

**M. Rafie (GTI),**

**Yuan Zhang (China Telecom)**

# Introduction

Common test conditions and evaluation methodology have been discussed for Video Coding for Machines. An Excel template for reporting and evaluating results for Video Coding for Machines was developed and refined during the last a couple of MPEG meeting cycles. The common test conditions, evaluation methodology and reporting template are finalized during the MPEG 133 meeting, taking into account the decisions and comments made during a set of AHG meetings. The common test conditions and evaluation methodology are presented in this document and the Excel reporting template is attached with this document. They will be used for the Call for Evidence (CfE) of Video Coding for Machines, and the following standard development activities.

# Common Test Conditions

Five machine tasks will be evaluated with corresponding datasets, shown as follows:

|  |  |
| --- | --- |
| Machine Task | Evaluation Dataset |
| Object detection | OpenImageV6  FLIR (IR dataset) |
| Instance segmentation | OpenImageV6 |
| Object tracking | HiEve-10 |
| Pose Estimation | HiEve-10 |
| Action Recognition | HiEve-10 |

# Evaluation Methodology

BD-rate, BD-mAP, BD-fmAP, or BD-MOTA are used to evaluate the proposed solution against the anchor. They are computed using the R-D curves from the proposed solutions compared with the R-D curve from the anchor. Here a RD curve can be either a mAP vs BPP curves, a mAP vs. bitrate, or a MOTA vs. bitrate curve depending on machine tasks. Details refers to the Excel reporting template attached to this document.

# References

1. S. Liu, W. Gao, X. Xu and C. Hollmann, “[VCM] Evaluation Methodology and Reporting Template”, MPEG-m56252, Online, January 2021.
2. MPEG Requirements, “Draft of Evaluation Framework for Video Coding for Machines”, MPEG WG 2-N00019, Online, October 2020.