|  |  |
| --- | --- |
|  | Moving Picture, Audio and Data Coding by Artificial Intelligencewww.mpai.community |

|  |  |
| --- | --- |
| M? | 2020/10/23 |
| Source | Roberto Iacoviello, RAI R&D, Email: roberto.iacoviello@rai.it |
| Title | Operational Guidelines |
| Target  | Contribution to MPAI-EVC |
| Purpose | Proposal |

# Abstract

The Operational Guidelines aim to facilitate the implementation of the Collaborative Evidence Conditions [1]. The Operational Guidelines are periodically revised to reflect the ongoing work of MPAI-EVC.

# Introduction

The purpose of these Operational Guidelines is to provide practical guidance to MPAI-EVC members to achieve, step by step, the common goal of starting from an existing standard and try to replace tools in that architecture with published tools that claim to perform well (Figure 1).

The current objectives are the following:

1. The starting point is the MPEG-EVC baseline (Essential Video Coding).
2. The MPAI-EVC members pick up an AI tool form the Table Tools (Table 1) upon they wish to operate.
3. The members then share their approach/results in the working group.
4. The Cross-checker then begin to express the pros and cons of each implementations
5. The members are expected to work in a common/collaborative process.
6. Each member realizes that it is in their best interests for themselves and of all the experts in the group to align processes as closely as possible.



Figure 1: Essential Video Coding scheme (Horizontal Hybrid approach)

# AI tools to add

In the following the list of deep tools to be added in the EVC schema:

**Table 1** Table Tools

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **AI-tool** | **Paper** | **R** | **Link to Source Code** | **MPAI-EVC member** | **Cross-checker** |
| AI Intra-Picture Prediction | Neural network based intra prediction for video coding | 6 |  |   |  |
| JVET-J0037-v1, Intra prediction modes based on neural networks | 2 |
| AI Cross-Channel Prediction | A hybrid neural network for chroma intra prediction | 8 |  |  |  |
| AI Inter-Picture Prediction | Generative adversarial network-based frame extrapolation for video coding | 7 |  |  |  |
| AI Probability Distribution Prediction | Convolutional neural network-based arithmetic coding of DC coefficients for HEVC intra coding | 13 |  |  |  |
| AI Transform |  |  |  |  |  |
| AI In-Loop Filtering | CNN-based in-loop filtering for coding efficiency improvement | 9 |  |  |  |
| m54991, Preliminary results of Neural Network Loop Filter | 3 |
| AI Post-Loop Filtering | The multi-scale deep decoder for the standard HEVC bitstreams | 12 |  |  |  |
| An Integrated CNN-based Post Processing Filter ForIntra Frame in Versatile Video Coding | 5 |
| AI Down and Up-Sampling | Convolutional Neural Network-Based Block Up-sampling for Intra Frame Coding | 10 |  |  |  |
| On Versatile Video Coding at UHD with Machine-Learning-Based Super-Resolution | 4 |
| AI encoding optimization |  |  |  |  |  |
| AI Motion Estimation |  |  |  |  |  |
| AI Coder control- rate control | A convolutional neural network-based approach to rate control in HEVC intra coding | 11 |  |  |  |

# Gantt

A detailed timeline planned activities is presented in the Gantt chart in Figure 2.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Contact authors scientific papers |   |   |  |  |  |  |  |  |
| Setting the environment |   |   |  |  |  |  |  |  |
| AI Intra-Picture Prediction |  |  |   |   |   |  |  |  |
| AI Cross-Channel Prediction |  |  |   |   |   |  |  |  |
| AI Inter-Picture Prediction |  |  |   |   |   |  |  |  |
| AI Probability Distribution Prediction |  |  |   |   |   |  |  |  |
| AI Transform |  |  |   |   |   |  |  |  |
| AI In-Loop Filtering |  |  |   |   |   |  |  |  |
| AI Post-Loop Filtering |  |  |   |   |   |  |  |  |
| AI Down and Up-Sampling |  |  |   |   |   |  |  |  |
| AI encoding optimization |  |  |   |   |   |  |  |  |
| AI Motion Estimation |  |  |   |   |   |  |  |  |
| AI Coder control - rate control |  |  |   |   |   |  |  |  |
| Evaluation Table |  |  |  |  |  |   |  |  |

Figure 2: A Gantt chart representing the detailed timeline of the activities

# Patent rights declaration(s)

**According to the Framework Licence.**

# References

|  |  |  |  |
| --- | --- | --- | --- |
| 1 | MPAI N42 Collaborative Evidence Conditions | Roberto Iacoviello | RAI |
| 2 | JVET-J0037-v1, Intra prediction modes based on neural networks | Jonathan Pfaff, Philipp Helle, Dominique Maniry, Stephan Kaltenstadler, Björn Stallenberger, Philipp Merkle, Mischa Siekmann, Heiko Schwarz, Detlev Marpe, Thomas Wiegand | Fraunhofer Institute for Telecommunications – Heinrich Hertz Institute |
| 3 | m54991, Preliminary results of Neural Network Loop Filter | Zhao Wang, Ru-Ling Liao, Changyue Ma, Yan Ye  | Alibaba |
| Shanshe Wang, Siwei Ma  | Peking University (Leonardo) |
| 4 | On Versatile Video Coding at UHD with Machine-Learning-Based Super-Resolution | Kristian Fischer, Christian Herglotz | Universitüt Erlangen-Nürnberg (FAU), Erlangen, Germany |
| André Kaup | Friedrich-Alexander-Universität Erlangen-Nürnberg, Erlangen, Germany(Leonardo) |
| 5 | An Integrated CNN-based Post Processing Filter For Intra Frame in Versatile Video Coding | Mingze Wang; Shuai Wan; Hao Gong  | Northwestern Polytechnical University, Xi'an, China |
| Yuanfang Yu; Yang Liu | Guangdong OPPO Mobile Telecommunications Corp., Ltd., China |
| 6 | Neural network based intra prediction for video coding | J. Pfaff, P. Helle, D. Maniry, S. Kaltenstadler, W. Samek, H. Schwarz, D. Marpe, T. Wiegand | Fraunhofer Institute for Telecommunications – Heinrich Hertz Institute |
| 7 | Generative Adversarial Network-Based Frame Extrapolation for Video Coding | Jianping Lin; Dong Liu; Houqiang Li; Feng Wu | CAS Key Laboratory of Technology in Geo-Spatial Information Processing and Application System, University of Science and Technology of China, Hefei, China |
| 8 | A hybrid neural network for chroma intra prediction | Yue Li; Dong Liu; Houqiang Li | CAS Key Laboratory of Technology in Geo-Spatial Information Processing and Application System, University of Science and Technology of China, Hefei, China |
| Li Li; Zhu Li;  | University of Missouri, Kansas City, MO, USA(Leonardo) |
| Jianchao Yang; Ning Xu; | Snapchat Inc., CA, USA |
| 9 | CNN-based in-loop filtering for coding efficiency improvement | Woon-Sung Park; Munchurl Kim | Department of Electrical Engineering, Korea Advanced Institute of Science and Technology, Daejeon, Korea(Miran) |
| 10 | Convolutional Neural Network-Based Block Up-sampling for Intra Frame Coding | Yue Li, Dong Liu, Houqiang Li, Li Li, Feng Wu, Hong Zhang, and Haitao Yang | IEEE Members |
| 11 | A convolutional neural network-based approach to rate control in HEVC intra coding | Ye Li; Dong Liu; Zhibo Chen | CAS Key Laboratory of Technology in Geo-spatial Information Processing and Application, System University of Science and Techonology of China, Hefei, China(Wen-Hsiao) |
| Bin Li | Microsoft Research Asia |
| 12 | The multi-scale deep decoder for the standard HEVC bitstreams | Tingting Wang; Wenhui Xiao; Mingjin Chen; Hongyang Chao | Ministry of Education, The Key Laboratory of Machine Intelligence and Advanced Computing (Sun Yat-sen University), Guangzhou, P.R. China |
| 13 | Convolutional Neural Network-Based Arithmetic Coding of DC Coefficients for HEVC Intra Coding | Changyue Ma; Dong Liu; Feng Wu | CAS Key Laboratory of Technology in Geo-spatial Information Processing and Application, System University of Science and Techonology of China, Hefei, China |
| Xiulian Peng | Microsoft Research Asia |