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

|  |
| --- |
| **Public document** |
| N61 | 2020/11/18 |
| Source | MPAI-1 |
| Title | Proposed MPAI Application Note #3 R1 - MPAI-EVC AI-Enhanced Video Coding |
| Target | MPAI Members |

**MPAI Application Note #3 R1**

**AI-Enhanced Video Coding (MPAI-EVC)**

**Description**: the fact that AI technologies improve data compression more than traditional technologies stays at the foundation of MPAI. The MPAI AI-Enhanced Video Coding (MPAI-EVC) work area is based on the results of a preliminary investigation on the performance improvement of AI-enhanced HEVC, AI-enhanced VVC and End-to-end AI-based video coding [1]: by replacing and/or enhancing exis­ting sel­ected HEVC and VVC coding tools with AI-based tools, the objectively measured compres­sion performance may be improved by up to around 30%.

While reassuring, these results were obtained by combining somewhat heterogeneous data from experiments reported in the liter­ature. Therefore, MPAI is conducting the so-called MPAI-EVC Evidence Project that investigates the feasibility of improving the coding efficiency by about 25% over an existing standard with an acceptable increase in complexity using technologies reported in the literature. If the investigation will be successful, MPAI will start the MPAI-EVC Standard project with the goal to develop the MPAI-EVC standard.

At this stage MPAI conducts two parallel activities

1. Thorough development of requirements that the MPAI-EVC should satisfy (this doc­ument gives an initial list of such requirements).
2. Collaborative activity targeting a sound assessment of the improvements achieved by state-of-the-art research. To the extent possible this should be done with the participation of the authors of claimed major improvements.

**Comments:**

1. The choice of the starting point (the existing codec), starting from which an AI-enhanced video codec should be developed, is an issue because high-performance video codecs have typically many essential patents (SEP) holders. They should all be convinced to allow MPAI to extend the selected starting point with AI-based tools that satisfy the – still to be defined – MPAI-EVC framework licence. As the result of such an endeavour is not guaranteed, MPAI has picked Essential Video Coding (MPEG-5 EVC) as the starting point. EVC baseline is reported not to be encumbered by IPR and the EVC Main Profile is reported to have a limited number ofstandard essential patent (SEP) holders. The choice between the EVC baseline and main profile is TBD.
2. It may eventually turn out that, MPAI-EVC is less performing than standards developed based on FRAND declarations because it would be constrained by using IP falling under the framework licence. However, MPAI-EVC would be coming with a framework licence that can be very close to an actual licence, while other standards would come with many FRAND declarations, likely in a much larger number than we have seen so far. MPAI-EVC could later be extended with more tools and new framework licences.

**Examples**

The following figures represent the block diagrams of 3 potential configurations to be adopted by the MPAI-EVC standard



*Figure 1* – *A reference diagram for the Horizontal Hybrid approach*

The green circles of *Figure 1* indicate traditional video coding tools that could be enhanced or replaced by AI-enabled tools. *Figure 1* is at the basis of the collaborative activity men­tioned above.

MPAI is also aware of ongoing research targeted at hybrid schemes where AI-based technologies are added to the existing codecs as an enhancement layer without making any change to the base-layer codec itself, thus providing backward-compatible solutions [2]. Some MPAI members are conducting research in this area and a coordinated MPAI activity could be kicked off soon. *Figure 2* shows a traditional video codec enhanced by an AI Enhancement codec.



*Figure 2* – *A reference diagram for the Vertical Hybrid approach*

Investigation [1] also showed that encouraging results can be obtained from new types of AI-based coding schemes – called end-to-end. These schemes, while promising, still need substantial more research.



*Figure 3 – End-to-end deep video compression scheme*

Even though MPAI considers the end-to-end approach of *Figure 3* not mature for standardis­ation, MPAI should not add any constraints on the technology that will be submitted in response to the MPAI-EVC Call for Technology other than satisfaction of the MPAI-EVC requirements [6].

MPAI is currently engaged in the MPAI-EVC Evidence Project with the goal to verify that AI-based technologies improve coding efficiency. It has produced the Operational Guidelines for MPAI-EVC Evidence Project [7] to provide practical guidance to achieve, step by step, the collaborative goal of starting from an existing standard and trying to replace tools in that architecture with published AI-tools that claim superior performance compared to traditional tools. The first tools planned to be replaced are Super resolution and In loop filter.

This project is being conducted in two parallel activities

1. Integrating EVC software with the neural network frameworks (e.g., Tensorflow, Python and Torch) via a web socket approach, thus building an abstraction layer agnostic to the frame­work.
2. Porting code developed in activity 1. to FPGA boards that are more effective than generic processors in terms of performance with low latency and high throughput [3,4,5]

**Requirements**

MPAI has already developed a consistent set of requirements [6]. Further revisions of the document are expected in the future.

**Object of standard**: Syntax and semantics of a bitstream entering a video decoder.

**Benefits**: Gradual introduction of AI-based technologies will allow a transition from technologies used in traditional signal processing to a common base of technologies used for information proc­essing.

**Bottlenecks**: The computational costs of AI-based tools for video compression should be assessed under common test conditions.

**Social aspects**: A simplified access to the technologies underpinning the MPAI-EVC standard will offer end users undelayed use of the latest video compression technologies.

**Success criteria**: MPAI becomes the bridge between traditional video codecs and fully AI-based video codes.

**References**:

1. Roberto Iacoviello; Analysis of performance of AI based video codecs, October 2020, submitted to MPAI incentive to use AI
2. W. C. Lee, C. P. Chang, W. H. Peng, and H. M. Hang, "A Hybrid-based Layered Image Compressor," *IEEE International Workshop on Multimedia Signal Processing (MMSP)*, Sep. 2020.
3. Luca Marchese, “The Internet Search Engines Based on Artificial Neural Systems Implemented in Hardware would Enable a Powerful and Flexible Context-Based Research of Professional and Scientific Documents”, 2015
4. NeuroStack, <https://www.general-vision.com/documentation/TM_NeuroStack_Hardware_Manual.pdf>
5. <https://www.analyticsinsight.net/why-fpga-is-better-than-gpus-for-ai-and-deep-learning-applications/>
6. N68 - MPAI-EVC Use Cases and Requirements, MPAI public document N68
7. Operational Guidelines for MPAI-EVC Evidence Project, MPAI public document N70