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

|  |
| --- |
| **Public document** |
| N36 | 2020/10/21 |
| Source | MPAI-1 |
| Title | Proposed MPAI Application Note #3 - MPAI-EVC AI-Enhanced Video Coding |
| Target | MPAI Members |

**MPAI Application Note #3**

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

**Description**: MPAI has carried out an investigation on the performance improvement of AI-enhanced HEVC, AI-enhanced VVC and End-to-end AI-based video coding [1]. Preliminary evidence offered by the investigation suggests that 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%. These results were obtained by combining somewhat heterogeneous data from experiments reported in the liter­ature.

The reported initial results, however, do indicate that AI can bring significant im­prov­ements to existing video coding technologies. Therefore, MPAI is investigating the feasibility of improving the coding efficiency by about 25% to 50% over an existing standard with an acceptable increase in complexity using technologies reported in the literature. If the investigation is successful, MPAI will develop a standard called MPAI AI-Enhanced Video Coding (MPAI-EVC).

The investigation showed that encouraging results can beobtained from new types of AI-based coding schemes – called end-to-end. These schemes, while promising, still need substantial more research.

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].

At this stage MPAI conducts two parallel activities

1. Collaborative activity targeting a scientifically 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 major improvements
2. Thourough development of requirements that the MPAI-EVC should satisfy (this doc­ument gives an initial list of such requirements).

**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 is planning to pick Essential Video Coding (MPEG-5 EVC) as the starting point. EVC baseline is reported not to be encumbered by IPR. Additionally, EVC Main Profile is reported to have a limited number of SEP holders. As an EVC patent holder has announced the release of a full implem­entation of EVC as Open Source Software (OSS), the choice of EVC as the starting point would also make available a working code base. 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, much larger 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. This will be taken as the basis of the collaborative activity men­tioned above.



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

In *Figure 2* a traditional video codec is enhanced by a Deep Enhancement codec.



*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 submitted in response to a Call for Technology other than satisfaction of the requirements.

**Requirements**

These are some initial requirements. Work on requirements will begin in earnesr once MPAI-EVC will be promoted to stage 2 Functional requirements.

* The standard shall be usable for multiple applications/use cases:
	+ Entertainment content distribution
	+ Live TV (soccer, ...)
	+ Off-line encoded content streaming (large delay and high complexity possible)
	+ Game content distribution, including Cloud Gaming (encoding with very very low delay)
	+ Videoconferencing (multipoint-to-multipoint, computed generated pictures)
	+ Video contribution (studio-to-studio) (low latency, less compression, higher-quality) 4:2:2 and 4:4:4, shorter GOP)
	+ High-quality video and images
	+ High-speed moving objects (defence, ...)
	+ Drones for monitoring (expand...)
	+ Medical video (up to 20k images/s) for ultra-fast, ultra-precise (5 micron)
	+ Telemedicine (point-to-point), higher quality than VC
* The standard shall support
	+ Coding of content of interest
* The standard shall be
	+ Friendly to architectures that have vector processing units, e.g. CPU + GPU (or FPGA), CPU+GPU+NPU
	+ Capable of exploiting a variable number of CPU cores
	+ Capable of encoding with varying degrees of delay
	+ Capable of balancing complexity between encoder and decoder as needed by the application
	+ Capable of providing about 25% to 50% more compression than EVC Main Profile
	+ Capable of providing backward compatibility with EVC

**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 traditionally used in 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.