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

**Public document**

|  |  |
| --- | --- |
| **N786** | 2022/07/19 |
| **Source** | Requirements (EEV) |
| **Title** | Requirements (EEV) progress report and plans |
| **Target** | MPAI-22 |

Requirements (EEV) mainly investigates the fine-tuning of the reference model using the BVI-DVC dataset to seek further coding gain on the top of the 2nd version of EEV reference model. However, the BVI-DVC dataset is a huge one with various spatial resolutions from 270P to 2160P, which results in a very low training/fine-tuning speed. The available experiments show that directly train the EEV model from scratch might not be a good solution. In addition, EEV has found a very interesting paper published in Neurips 2021 by Microsoft Research Asia, entitled “Deep Contextual Video Compression”, which proposed a novel DCVC framework for end-to-end video coding. The coding performance is encouraging. The main idea of DCVC is to transform the inter-frame prediction into deep feature domain, in which it calls contextual information embedded in the reference frame. Another important contribution is that the DCVC also models motion in the deep feature domain.

Since it is difficult to directly use their model due to the licence problem (it is not possible to as Microsoft for the permission for code and intellectual property). One practical way is to absorb some effective idea or solutions from DCVC and use its design for EEV.

In the next step, the requirements (EEV) is planned to investigate more neural network based coding tools into the reference software, such as multiple reference frame based EEV, or in-loop enhancement network models, such that further coding gain can be realized.