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

|  |  |
| --- | --- |
| **N271** | 2021/06/09 |
| **Source** | Video group |
| **Title** | Status report on MPAI-EVC Evidence Project |
| **Target** | MPAI Members |

MPAI has created a collaborative environment that will allow testing independently developed AI tools on a common EVC code base. The goal is to enhance EVC to reach at least 25% improvement over EVC baseline profile.

The group is working on three phases: database building, learning phase and inference phase. All these phases will be developed in parallel.

The state of the art is:

* Intra prediction: EVC enhanced code, EVC-socket communication working, database building:
	+ 1.5M pictures database of 32x32 intra prediction
	+ Neural network trained for 32x32 intra prediction
	+ Tested on 1st frame of Class B sequences *BasketballDrive*, *BQTerrace*, *Cactus*, *Kimono1*, *ParkScene*
	+ Computed BD rate over QPs = {32, 37, 42, 47}
	+ average BD-Rate: -0.6% (Figure 1)



Figure 1: BD-rate over 5 test sequences

* + To further improve the performances, we decided to extend the intra prediction enhancement to 16x16 block size. We are following two parallel tracks:
		- low hanging fruit: using the same parameters of the previous work (32x32 training phase) in order to don’t retrain the network



Figure 2: DP (Decoded Predictor) format (inference phase)

We will use a 16x16 intra predictor size embedded into a 48x48 context (Figure 2). The DP format will maintain the same size of the previous one (MPAI Dcoument N240)

* + - in case we don’t get satisfying performance we will retrain the neural network on a new data set
* Dataset building
	+ Training dataset of 32x32 intra prediction block created
	+ Training dataset 16x16: we have already written the code to extract the ODP (Original Decoded Predictor) format (MPAI Document N240). We are now producing a large dataset (> Million pictures)
* Super resolution track: we have selected the super resolution as a second AI tool. We have tested few neural networks available on Internet and selected the Densely Residual Laplacian Super Resolution (DRLN) as a candidate one.
* We have built a dataset to train the super resolution network: 2000 pictures
	+ 3 use cases:
		- SD to HD
		- HD to 4K
		- SD to 4K
	+ 3 resolutions:
		- 4k - 3840x2160
		- HD - 1920x1080
		- SD - 960x540 (actually 544…)
	+ 3 QP values:
		- 15
		- 30
		- 45
	+ Two options:
		- Deblocking enabled
		- Deblocking disabled
	+ 3 resolutions, 3 quality values, 2 coding tool sets
		- 18 versions, 170 GB dataset
* The training phase of the Super resolution network will require substantial computing resources. The group has decided to distribute the workload between members building an excel file with the hyper parameters to train.

The next steps are:

|  |  |  |
| --- | --- | --- |
|  |  |  |
| **Date** | **Topic** | **Who** |
| 1 meeting cycle | AI-intra next step: Test 16x16 block size in a 64x64 DP format  | Attilio, Roberto |
| 1 meeting cycle | AI-intra next step: Build a database of 16x16 block size | Alessandra, Roberto |
| 2 meeting cycles | AI-intra next step: deliver the results on further experiments on training | Alessandro, Francesco (Kebula), Giuseppe (Kebula), Francesco (CNR) |
| 2 meeting cycles | Measure the performances after training (BD-Rate)  | Attilio |
| 2 meeting cycles | AI-SR next step: Perform other experiment to understand the pre-condition of training (e.g. cropping type, which data augmentation, training procedure)  | Alessandro, Francesco CRN, Francesco and Giuseppe Kebula |
| 2 meeting cycles | Produce a short tutorial to deploy and train the network  | Francesco and Giuseppe |
| 2 meeting cycles | Produce an Excel file with the people that will work on the training | Roberto, Alessandro, Francesco (Kebula), Giuseppe (Kebula), Francesco (CNR) |