

周报

汇报人：向辉

时间分配

- 2h: 上课
- 8h: 阅读论文
- 25h: 配置以及测试DCVC-RT
- 5h: 修改论文

学习

- 和mentor一起复现DCVC-RT论文，但是模型是半精度训练，且代码去掉了所有的forward函数。首先配置了DCVC-RT的环境，其中遇到了各种c++编译出错以及c++链接不到对应库的问题。一开始尝试一个问题一个问题解决，但是发现很多问题很难找到对应的解决方法。后来尝试严格的配置DCVC-RT所对应的环境，包括cuda版本、pytorch版本、python版本，终于找到了可以编译其中inference下文件的docker镜像，跑通了test_video脚本。
- 学习DCVC-RT的代码，debug查看代码每一步。总结了DCVC-RT之所以帧率高的几个原因：
 - i. 直接对每帧使用pixel unshuffle进行八倍下采样，大大减少中间模型隐变量的分辨率。
 - ii. 替换模型中所有的卷积为cuda c++实现的卷积，并且cuda c++重写了熵模型中的mask操作使之能并行计算。
 - iii. 所有隐变量都是float16，计算速度更快。
 - iv. 去除光流部分，直接使用前一帧的feature，不用再编额外的光流。
- 由于eem测试集中有10bit，解决了DCVC-RT中无法支持10bit的bug。解决以后，在ivc集群的4090D上跑了通甲乙丙的yuv420数据集，并统计计算了其每帧编解码的实际时间，并写提案中复杂度部分。
- 在USTC-TD的数据集上测试DCVC-RT，包括yuv420和png，同时测试了DCVC-FM在png上的性能。
- 阅读了论文：
DCVC-RT_Balanced Rate-Distortion Optimization in Learned Image Compression ,
Estimating the resize parameter in end-to-end learned image compression ,
TMM_Learned_Focused_Plenoptic_Image_Compression_With_Local-
Global_Correlation_Learning ,
Compression-Aware One-Step Diffusion Model for JPEG Artifact Removal ,

Learnable_Non-uniform_Quantization_With_Sampling-based_Optimization_for_Variable-Rate_Learned_Image_Compression,

Large Language Model for Lossless Image Compression with Visual Prompts

- 修改论文，和mentor讨论了关于JPEG无损压缩论文的细节，发现了很多问题，比如在语句上的不通顺、表格规范、引用格式、语言描述、表述逻辑等诸多问题。目前修改了引用数量的问题，还有很多的问题正在修改当中。

下周计划

- 20h: 参加AVS会议
- 10h: 继续修改论文
- 5h: 阅读论文
- 5h: 继续完善DCVC-RT的代码