

数字音视频编解码技术标准工作组[©]

AVS M8526: 2024 年 08 月 20 日

来源: 唐传波 卞逸凡 盛锡华 李卓元 李礼 刘东 (中国科学技术大学)
标题: 一种用于稳定 EEM 训练的验证集 USTC-TD
状态: 输入提案

摘要

现有 EEM2.1 版本训练分为 pretrain 和 finetune 两个阶段，在 pretrain 阶段会训练 4 个相同码率的高码率点模型，后在 finetune 阶段使用 4 个不同的 λ 分别对 pretrain 得到的 4 个模型进行微调，从而得到 4 个码率点模型。然而，这样的训练方式会由于 pretrain 阶段部分模型的训练不稳定，导致最终 finetune 阶段获得的对应码率点模型性能较差。为了提升 EEM 训练的稳定性和训练性能，本提案提出在 pretrain 的最后阶段和 finetune 阶段加入 USTC-TD 视频验证集（包含 8 个 96 帧 1080P 序列）对模型进行挑选。具体而言，在 pretrain 阶段使用验证集从 4 个相同码率点模型中挑选出验证损失最小的模型作为预训练模型，在 finetune 阶段对挑选出的预训练模型进行微调，并用验证集挑选出每个码率点验证损失最小的模型。实验测试在 LDP 配置下与 EEM2.1 (short 训练策略) 进行性能对比，结果显示在 Y_PSNR, U_PSNR 和 V_PSNR 上分别取得 2.38%，1.79%，和 2.88% 的性能增益。

正文

1. 方法

我们从验证集 USTC-TD（包含 8 个 96 帧 1080P 序列）中每个序列随机挑选出连续的 7 帧进行验证推理，使用的验证损失函数：

$$Loss_{val} = \frac{1}{seq_num} \sum_{seq=1}^{seq=8} \frac{1}{frame_num} \left(\sum_{t=1}^{t=7} \lambda \cdot D + R \right) \quad (1)$$

其中，seq_num 为 8，frame_num 为 7，pretrain 阶段四个 λ 都为 2048，finetune 阶段四个 λ 分别为 2048，512，256，64。我们在 pretrain 和 finetune

[©] 本文件和所包含技术信息的知识产权归提出者所有，仅供“数字音视频编解码技术标准工作组”工作范围内应用，并非公开发表，提出者保留申请专利或在公开出版物发表等相关权利。对于侵犯上述权利的行为，文档提出者和“数字音视频编解码技术标准工作组”保留公开谴责或提起诉讼的权利。

阶段调用验证集对模型进行挑选。

a) Pretrain 阶段

在 pretrain 阶段的 stage3 的最后 2 个 epoch，我们调用验证集进行验证，先在每个码率点下选出当前验证损失最小的模型，最后在四个相同码率点层级挑选出验证损失最小的模型作为 finetune 阶段的预训练模型。

b) Finetune 阶段

在 finetune 训练的所有阶段，我们调用验证集进行验证，挑选出每个码率点中验证损失最小的模型作为最终的训练模型。

2. 测试

本方法在 AVS 通测序列上进行测试，在 LDP 配置下与 EEM2.1（short 训练策略）进行性能对比，结果如下：

相较于 EEM2.1（short 训练策略），通甲在 Y_PSNR, U_PSNR 和 V_PSNR 上的 BD-Rate 分别为：-3.39%，-0.44%，-3.49%；通乙在 Y_PSNR, U_PSNR 和 V_PSNR 上的 BD-Rate 分别为：-2.13%，-2.43%，0.46%；通丙在 Y_PSNR, U_PSNR 和 V_PSNR 上的 BD-Rate 分别为：-1.62%，-2.51%，-5.61%。

	Low delay P		
	VS EEM2.1		
	Y	U	V
通甲 4K	-3.39%	-0.44%	-3.49%
通乙 1080P	-2.13%	-2.43%	0.46%
通丙 720P	-1.62%	-2.51%	-5.61%
Overall	-2.38%	-1.79%	-2.88%

3. 验证集选取及介绍

验证集序列选取 USTC-TD 视频验证集中 8 个 96 帧 1080P 视频（如图 1 所示），该数据集及其相关的评估测试数据结果已公布于开源网址（<https://esakak.github.io/USTC-TD/>，如图 2 所示），所采用的验证集已连续两年应用于国际会议的端到端图像/视频压缩竞赛。相比先前的编码测试集，该数据集在多个数据集评价指标下，验证了其数据验证覆盖范围更广，且能够更全面体现被测试编码器的性能。



图 1 验证集序列图

USTC-TD: USTC Test Dataset for Image and Video Coding in 2020s

Zhuoyuan Li, Junqi Liao, Xihua Sheng, Haotian Zhang, Yuqi Li,
Chuanbo Tang, Yifan Bian, Yao Li, Xinmin Feng, Changsheng Gao,
Li Li, Dong Liu

University of Science and Technology of China (USTC),
Intelligent Visual Data Coding Laboratory (iVC Lab)



Figure 1. Examples of USTC Video Test Dataset.

图 2 验证集开源示例

4. 总结

本提案针对 EEM 的训练提供一套验证集，用于在训练中根据验证损失挑选模型。在通测序列数据集上， Y_PSNR ， V_PSNR 和 U_PSNR 指标平均获得 2.38%，1.79% 和 2.88% 的码率节省。建议标准组采纳提供的验证集以及利用验证集挑选 pretrain、finetune 模型的方式。

中国数字音视频编解码技术标准工作组

会员提案专利披露与许可承诺表

根据《中国数字音视频编解码技术标准工作组知识产权政策》第十四、十五、十六条等相关规定，AVS 会员在向工作组各专题组提交技术提案时应填写本《会员提案专利披露与许可承诺表》，作为该提案的必要组成部分同时提交。

专题组名称： 音频 视频 系统 DRM

提案 AVS 文档编号： M8XXX 提案日期： 2024 年 8 月 21 日

提案标题： 一种用于稳定 EEM 训练的验证集 USTC-TD

提案会员名称： 中国科学技术大学

提案代表姓名（印刷体）： 唐传波 Email: cbtang@mail.ustc.edu.cn

提案代表通讯地址： 安徽省合肥市中国科学技术大学高新校区 邮编 230026

电话： 13308039177 传真：

提案会员应当通过选中表 A 或者通过选中并填写表 B 相关部分完成此表。表 C 可以自愿填写。下列表格均可根据实际需要增加表格行。

表 A：提案会员在其实际知晓的范围内已获知本提案不涉及提案会员和他人的专利、专利申请和专利计划。

表 B-1：在中华人民共和国已获得授权的专利和/或已公开的专利申请

如果本提案中包含提案会员或其关联者在中华人民共和国已获得授权的专利和/或已公开的专利申请，提案会员应当填写下表：

序号	专利名称	申请号 或 专利号	申请日 或 授权日	由提案会员自愿给出的专利主要技术 内容介绍及其与标准内容相关性说明 (可另加附件)	许可承诺(请选择) (a) RAND 免费许可或 (b) 加入 AVS 专利池
1					
2					

表 B-2：在中华人民共和国未公开的专利申请

如果提案会员的缺省许可义务不是 RAND-RF 或者 POOL, 当提案会员或其关联者有与此提案相关的未公开的中华人民共和国专利申请时，提案会员必须选中此表。

提案会员可在下表中自愿披露上述未公开的专利申请的具体情况并选择许可承诺：

序号	专利名称	申请号 或 受理号	申请日	专利主要技术内容介绍及其与标准内 容相关性说明(可另加附件)	许可承诺(请选择) (a) RAND 免费许可或 (b) 加入 AVS 专利池
1					(b) 加入 AVS 专利池
2					

表 B-3: 在中华人民共和国之外已获得授权的专利和/或已公开的专利申请

如果本提案中包含提案会员或其关联者在中华人民共和国之外已获得授权的专利和/或已公开的专利申请, 提案会员应当填写下表:

序号	专利名称	申请号 或 专利号	申请日 或 授权日	申请或者 获得授权 的国家 /地区	由提案会员自愿给出的专利主要技术 内容介绍及其与标准内容相关性说明 (可另加附件)	许可承诺(请选择) (a) RAND 免费许可或 (b) 加入 AVS 专利池或 (c) RAND
1						
2						

表 B-4: 在中华人民共和国之外未公开的专利申请

如果提案会员的缺省许可义务不是 RAND-RF 或者 POOL, 当提案会员或其关联者有与此提案相关的在中华人民共和国之外的未公开的专利申请时, 提案会员必须选中此表。

提案会员可在下表中自愿披露上述未公开的专利申请的具体情况 & 选择许可承诺:

序号	专利名称	申请号 或 受理号	申请日	申请的 国家 /地区	专利主要技术内容介绍及其与标准内 容相关性说明(可另加附件)	许可承诺(请选择) (a) RAND 免费许可或 (b) 加入 AVS 专利池或 (c) RAND
1						
2						

表 C: 自愿填写

提案会员自愿在其实际知晓的范围内披露的本提案可能涉及到的他人专利和专利申请:

序号	专利名称	申请号 或 专利号	申请日 或授权 日	申请或者 获得授权 的国家/地 区	专利的主要技术内容介绍及其与标 准内容相关性说明(可另加附件)	申请人或 专利权人
1.						
2.						