

数字音视频编解码技术标准工作组

AVS M8984: 2025 年 6 月

来源: 李宇琪 张昊田 李瑶 李卓元 李礼 刘东 (中国科学技术大学)
标题: EEM I 帧模型改进
状态: 输入提案

摘要

本提案在两方面改进 EEM 中的 I 帧模型：第一，优化网络结构，引入基于部分卷积的残差块网络，降低 I 帧模型复杂度并提升压缩性能。第二，引入可变码率技术，实现单模型支持多个码率点的功能。本提案提出的 I 帧模型在广泛码率范围内均表现优异，相较于 EEM6.1 的 I 帧模型，性能提升超过 10%，解码复杂度低于 200kMAC/pixel。

正文

1. 方法概述

本提案提出的 I 帧模型结构如图 1 所示，主要由变换模块与 Hyperprior 熵模型组成。模型中的变换网络和熵模型网络均采用上下采样卷积（其中 k2s2 表示卷积核大小为 2，步幅为 2）与 ResBlock 结合构建。ResBlock 内部引入了高效的部分通道空间卷积（Partial Convolution），在保持建模能力的同时有效降低了计算开销。图一中各个模块网络数量分别是 $[L1, L2, L3, L4, L5, L6] = [4, 4, 18, 2, 2, 4]$ 。

为了支持可变码率功能，本模型在编码变换、解码变换以及 Hyperprior 网络中引入嵌入层 $\text{nn.Parameter}(q, C, 1, 1)$ ，用于调制中间特征，使网络在不同码率下均能保持良好性能。

模型中设置的码率控制参数 q 取值范围为 0 至 63，共 64 个点。在训练阶段， q 值通过随机采样获得，并根据设定的码率区间 $[\lambda_{\min}, \lambda_{\max}]$ ，计算每个 q 对应的 RD 损失中的 Lagrange 乘子 λ 值：

$$\lambda_q = \exp \left(\log(\lambda_{\min}) + \frac{q}{63} \cdot (\log(\lambda_{\max}) - \log(\lambda_{\min})) \right)$$

在测试阶段，通过设定不同的 q 值即可灵活调整码率，实现单模型多码率的功能。

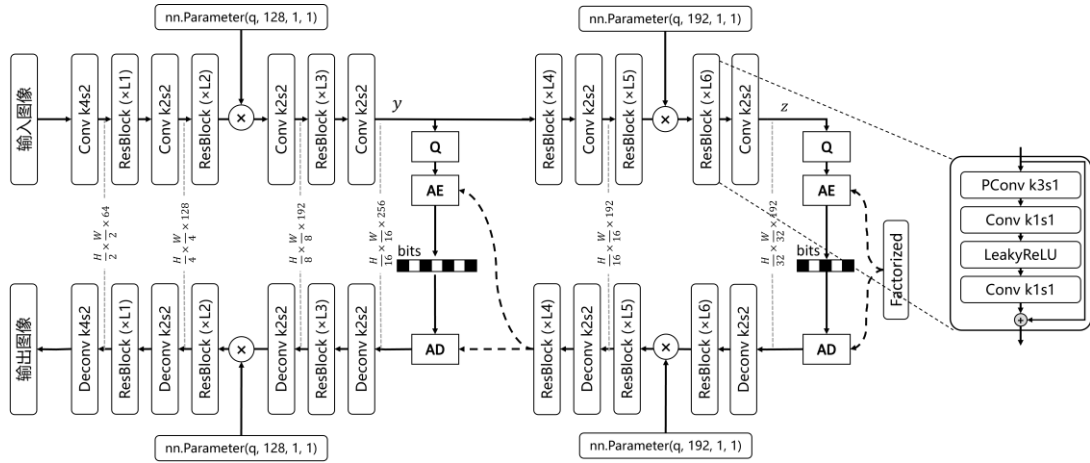


图 1. I 帧模型结构示意图

2. 实验结果

本提案提出的 I 帧模型在压缩性能与解码复杂度上均取得了显著改进。如表 1 所示，在测试序列第一帧的条件下，相较于 EEM6.1 的 I 帧模型，本提案模型在 Y、U、V 三个分量上码率节省为 7.56%、18.99%和 10.30%。此外，如表 2 所示，本提案提出的 I 帧模型解码复杂度低于 200kMAC/pixel。相比于 EEM6.1 的 I 帧模型的低码率模型与高码率模型，可以分别减少约 60k 和 380 kMAC/pixel。

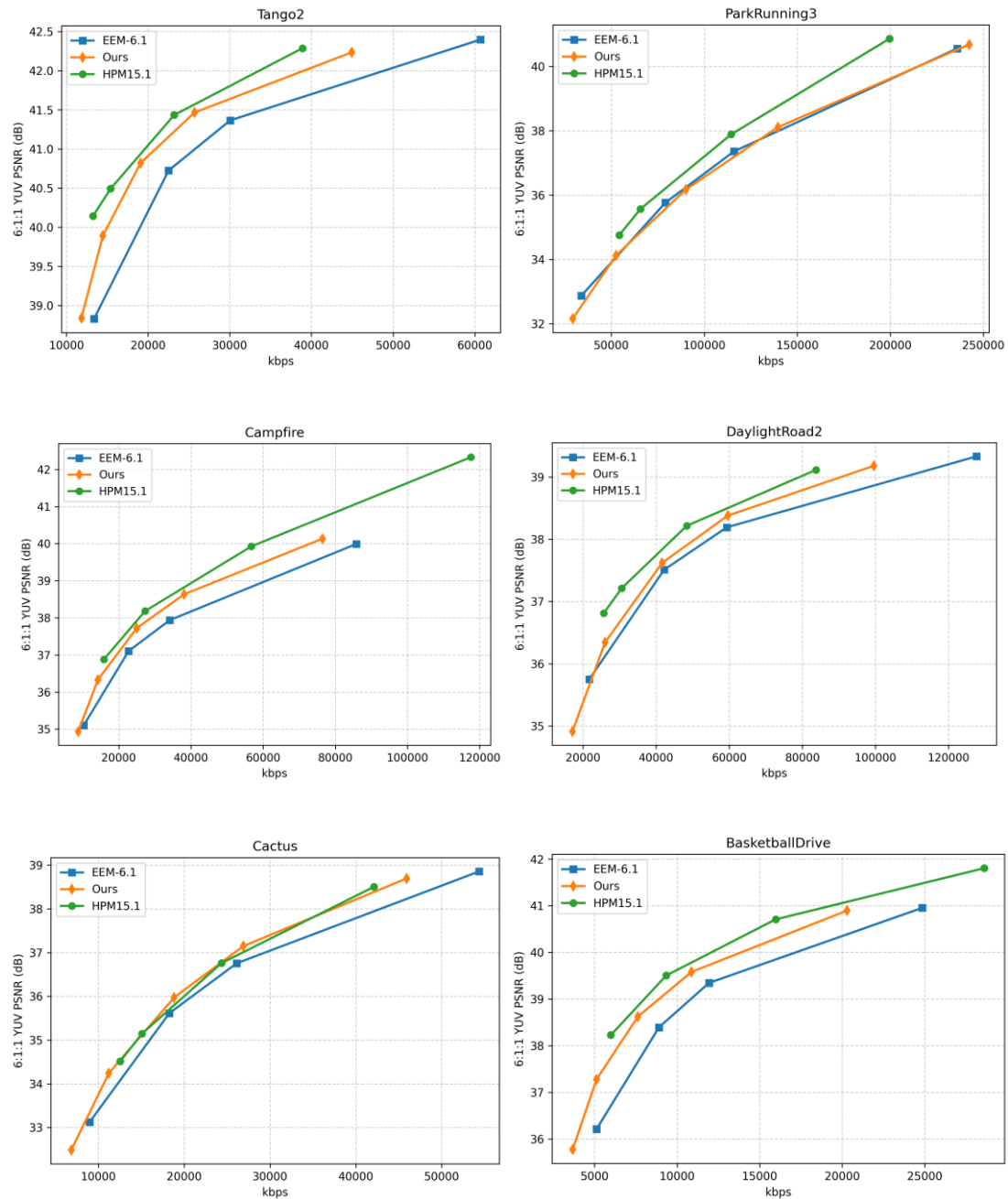
	Low delay P					
	Over EEM6.1			Over HPM15.1		
	Y	U	V	Y	U	V
Tango2	-12.84%	-34.76%	-20.22%	18.50%	-7.66%	-3.61%
ParkRunning3	7.94%	-31.05%	-24.00%	13.47%	45.79%	40.08%
Campfire	-7.20%	-42.11%	-28.38%	17.77%	19.10%	-8.62%
DaylightRoad2	-7.06%	-13.08%	-2.65%	17.29%	-10.56%	18.87%
Cactus	-8.64%	-15.09%	-5.72%	3.80%	-28.66%	-20.32%
BasketballDrive	-16.69%	-26.02%	-20.64%	26.51%	-12.01%	-19.31%
MarketPlace	1.36%	-0.30%	12.00%	8.16%	-18.24%	-20.54%
RitualDance	-9.14%	-8.85%	-18.42%	14.52%	5.39%	-11.74%
City	-0.06%	0.05%	13.95%	17.34%	4.70%	-3.30%
Crew	-13.33%	-26.07%	-9.58%	22.40%	-27.06%	-14.38%
vidyo1	-12.26%	-10.60%	-9.78%	-0.44%	-34.84%	-33.36%
vidyo3	-12.77%	-19.96%	-10.18%	9.87%	-21.29%	-12.57%
通甲	-4.79%	-30.25%	-18.81%	16.76%	11.67%	11.68%
通乙	-8.28%	-12.57%	-8.20%	13.25%	-13.38%	-17.97%
通丙	-9.60%	-14.15%	-3.90%	12.29%	-19.62%	-15.90%
Overall	-7.56%	-18.99%	-10.30%	14.10%	-7.11%	-7.40%

表 1. 本提案提出的 I 帧模型性能对比

模型	Decoder kMAC/pixel
EEM6.1 低码率 I 帧模型(ckpt_1~ckpt_3)	256.80
EEM6.1 高码率 I 帧模型(ckpt_4~ckpt_7)	575.25
本提案模型	197.82

表 2. 不同 I 帧模型解码复杂度对比

HPM15.1、EEM6.1 与本提案提出的 I 帧模型在通甲、通乙、通丙三类共 12 个测试序列上的 RD 曲线如图 2 所示,其中本提案提出的 I 帧模型取 $q=[0, 15, 31, 47, 63]$ 共五个码率点。可以看到, 本提案提出的 I 帧模型可以单模型覆盖较大地码率范围。



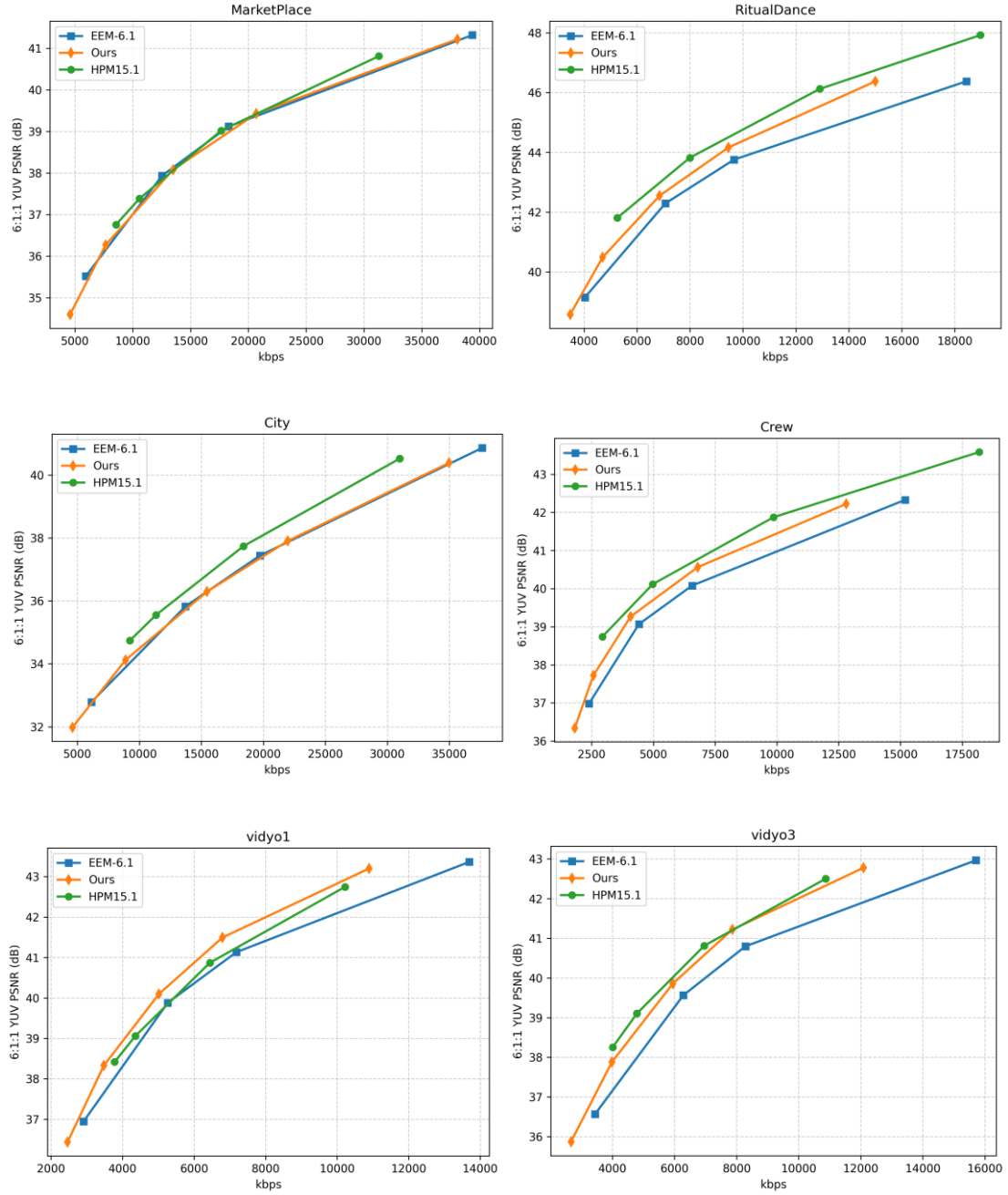


图 2. I 帧模型 RD 曲线

后续尝试将该 I 帧模型与 P 帧模型进行联合训练。联合训练时,为对齐 EEM 中原有的 7 个 I 帧模型配置,取 $q = [15, 23, 31, 39, 47, 55, 63]$ 得到共 7 个固定码率 I 帧模型,用以替代原有 I 帧参与联合训练。实验结果如表 3 所示,联合本提案的 I 帧模型训练 P 帧后整体性能出现一定程度下降。这可能是由于目前简单地将该可变码率模型分成 7 个固定码率模型进行训练时,可能存在 I 帧码率与 P 帧 λ 无法对齐的情况,导致模型码率分配失衡,最终影响整体压缩性能。后续工作可进一步探索联合训练策略,以更好发挥可变码率 I 帧模型的潜力。

	Over EEM6.1		
	Y	U	V
Tango2	3.25%	10.46%	2.51%
ParkRunning3	3.89%	6.90%	8.94%
Campfire	2.01%	-3.04%	2.01%
DaylightRoad2	9.81%	22.48%	10.51%
Cactus	3.33%	5.61%	6.75%
BasketballDrive	4.87%	13.27%	3.41%
MarketPlace	5.13%	5.51%	8.44%
RitualDance	3.56%	7.04%	3.17%
City	10.86%	12.54%	14.38%
Crew	1.64%	5.09%	-0.49%
vidyo1	-2.76%	-2.50%	-3.71%
vidyo3	1.64%	5.57%	-3.57%
通甲	4.74%	9.20%	5.99%
通乙	4.22%	7.86%	5.44%
通丙	2.84%	5.17%	1.66%
Overall	3.94%	7.41%	4.36%

表 3. 本提案提出的 I 帧模型联合训练 P 帧模型的性能对比

3. 总结

本提案提出了一种更先进的 I 帧模型，通过优化网络结构、引入可变码率技术，该 I 帧模型可以单模型支持较宽的码率范围，并且相比于 EEM6.1 中的 I 帧模型取得 10% 以上的性能增益，解码复杂度低于 200kMAC/pixel。鉴于该模型在 I 帧上的突出表现，建议采纳所提出的方案替换原有的 I 帧模型，并进一步研究 I 帧和 P 帧联合训练，特别是 P 帧可变码率模型的训练

会员提案专利披露与许可承诺表

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.						