

引用格式:于海,杨磊,高阳,等. 基于块编码特点的压缩视频质量增强算法[J]. 北京工业大学学报, 2024, 50(9): 1069-1076.  
YU H, YANG L, GAO Y, et al. A compressed video quality enhancement method based on block coding features[J].  
Journal of Beijing University of Technology, 2024, 50(9): 1069-1076. (in Chinese)

## 基于块编码特点的压缩视频质量增强算法

于海<sup>1</sup>, 杨磊<sup>1</sup>, 高阳<sup>1</sup>, 刘枫琪<sup>1</sup>, 刘鹏宇<sup>2,3,4</sup>, 孙萱<sup>2,3,4</sup>, 张悦<sup>2,3,4</sup>

(1. 河南九域恩湃电力技术有限公司, 郑州 450000; 2. 北京工业大学信息学部, 北京 100124;  
3. 先进信息网络北京实验室, 北京 100124; 4. 计算智能与智能系统北京市重点实验室, 北京 100124)

**摘要:** 针对现有压缩视频质量增强算法未能充分利用压缩视频特点的问题, 研究了视频编码与压缩视频质量增强任务之间的本质关系, 并针对性地设计了一种基于三维卷积神经网络(3D convolutional neural network, 3D-CNN)的非对齐压缩视频质量增强算法。实验结果表明: 相较于高效视频编码(high efficiency video coding, HEVC)标准 H. 265, 所提算法在低延迟(low delay, LD)配置下且量化参数(quantization parameter, QP)为 37 时, 峰值信噪比(peak signal-to-noise ratio, PSNR)提升了 0.465 2 dB; 相较于数据压缩会议(data compression conference, DCC)中提出的多帧视频质量增强方法(multi-frame guided attention network, MGANet), 该算法 PSNR 的增长量提升了 15.1%。

**关键词:** 视频编码; 高效视频编码标准(high efficiency video coding, HEVC); 压缩视频质量增强; 深度学习; 卷积神经网络(convolutional neural network, CNN); 三维卷积神经网络(3D convolutional neural network, 3D-CNN)

中图分类号: TN 919. 81

文献标志码: A

文章编号: 0254-0037(2024)09-1069-08

doi: 10.11936/bjtxb2022080003

## A Compressed Video Quality Enhancement Method Based on Block Coding Features

YU Hai<sup>1</sup>, YANG Lei<sup>1</sup>, GAO Yang<sup>1</sup>, LIU Fengqi<sup>1</sup>, LIU Pengyu<sup>2,3,4</sup>, SUN Xuan<sup>2,3,4</sup>, ZHANG Yue<sup>2,3,4</sup>

(1. Henan Jiuyu EPRI Electric Power Technology Co., LTD, Zhengzhou 450000, China;

2. Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China;

3. Beijing Laboratory of Advanced Information Network, Beijing 100124, China;

4. Beijing Key Laboratory of Computational Intelligence and Intelligence Systems, Beijing 100124, China)

**Abstract:** To solve the issue that existing compressed video quality enhancement algorithms do not fully utilize the characteristics of compressed videos, the intrinsic relationship between video encoding and the task of compressed video quality enhancement was studied and a targeted non-aligned compressed video quality enhancement algorithm was designed contrapuntally, utilizing a three-dimensional convolutional neural network (3D-CNN). Experimental results show that compared with the high efficiency video coding (HEVC) standard, the peak signal-to-noise ratio (PSNR) of the proposed method is improved to 0.465 2 dB under low delay (LD) configuration and quantization parameter (QP) is 37. Compared with MGANet proposed in data compression conference (DCC), the increase of proposed algorithm PSNR is

收稿日期: 2022-08-08; 修回日期: 2022-12-14

基金项目: 青海省重点研发与转化计划资助项目(2022-QY-205)

作者简介: 于海(1971—), 男, 高级工程师, 主要从事电力工程技术与电力信息化方面的研究, E-mail: 15939001869@126.com

通信作者: 刘鹏宇(1979—), 女, 副教授, 主要从事多媒体信息处理方面的研究, E-mail: liupengyu@bjut.edu.cn

improved by 15.1%.

**Key words:** video coding; high efficiency video coding (HEVC); compressed video quality enhancement; deep learning; convolutional neural network (CNN); 3D convolutional neural network (3D-CNN)

随着多媒体信息技术的持续发展和智能终端技术的逐渐成熟,广大民众在信息可视化方面的现实需求日趋增多。视频作为内容信息传递的主要媒介,已全面融入人民生活的诸多方面,成为推动社会经济蓬勃发展的强大动力。从发展规模的角度分析,视频流量已成为全球互联网流量的核心,视频数据规模呈现出爆发式增长的发展趋势,视频的高效传输与存储面临着前所未有的严峻考验。

视频编码技术的目标就是减少视频冗余,进而实现视频的高效存储和传输。由此看来,视频编码技术作为视频产业的关键一环,其技术发展水平尤为重要。由于网络带宽和存储容量的限制,着眼于更高压缩效率和更高画面质量的高效视频编码(high efficiency video coding, HEVC)标准 H. 265<sup>[1]</sup>应运而生。H. 265 标准集成的众多先进技术使其与当今的视频数据规模和特点天然契合,这为其在视频通信等应用领域的大规模实用化提供了关键推动力。

H. 265 标准中编码单元划分、量化等相关编码技术使其在本质上属于一种基于块的有损压缩方式。因此,视频编码技术本身会不可避免地产生大量的压缩失真,如块效应和振铃效应。虽然 H. 265 标准在编码端设置了环路滤波器来应对这一问题,但其只能在一定程度上缓解块效应和振铃效应的影响,更无法处理其他的压缩失真。因此,重建视频的主客观质量仍然有很大的提升空间。此外,随着显示设备制造业的蓬勃发展,4K、8K 等超高清显示终端已走进千家万户。重建视频的压缩失真在超高清显示设备上会被进一步放大,严重影响消费者的主观感受。因此,如何有效地提高压缩视频的质量是当前迫切的现实需求。

近些年来,随着深度学习的蓬勃发展,国内外的相关研究机构都在积极尝试将视频编码技术与深度学习相结合。卷积神经网络(convolution neural network, CNN)在图像识别、图像分类等计算机视觉任务中取得了显著的成果,这为图像恢复任务提供了新的研究方向。CNN 拥有学习退化图像到原始图像之间的非线性映射的强大能力<sup>[2-3]</sup>。因此,在图像超分辨率、图像去模糊和图像去噪等图像恢复任

务上,CNN 相较于传统方法取得了很大的进步。

在 H. 265 标准中,后处理环节的目标是最小化编码后视频的失真,提高编码视频质量和压缩效率。后处理环节可以利用终端强大的计算能力,在不增加传输消耗的情况下,有效地恢复编码视频的图像质量,有利于充分利用视频的帧内和帧间信息,平衡编解码压力。因此,可以将后处理视为图像恢复类的任务,而 CNN 的应用可以有效地提高解码视频的质量。

根据输入帧的不同,压缩视频质量增强技术可分为基于单帧输入的方法和基于多帧输入的方法。基于单帧输入的方法是只使用帧内的像素信息来恢复该帧<sup>[4-6]</sup>。Wang 等<sup>[7]</sup>提出了一种基于 CNN 的针对环路滤波后重建图像进行增强的方法。由于不同量化参数(quantization parameter, QP)影响压缩图像的失真程度,因此,需要针对不同的 QP 分别训练模型。研究人员分别针对帧内编码失真和帧间编码失真提出了 CNN 质量增强(quality enhancement with convolutional neural network, QE-CNN)方法<sup>[8]</sup>。虽然这些方法表面上处理的是压缩视频,但本质上是循环输入单帧,只利用了帧内信息,不能有效地挖掘视频中特有的帧间信息。事实上,帧间信息的有效利用对视频增强任务是非常重要的。从理论上讲,如果放弃对帧间信息的利用,视频质量增强任务的性能将因为原始信息的先天缺乏而受到严重限制。

针对多帧输入的压缩视频增强的研究<sup>[9-11]</sup>是近几年才开始的。GUAN 等<sup>[12]</sup>、Yang 等<sup>[12]</sup>提出了使用多帧作为输入的多帧卷积神经网络(multity frame convolutional neural network, MF-CNN)用于视频增强,但其网络结构复杂,性能较差。Wang 等<sup>[14]</sup>建立了可利用时域信息的多分辨率 CNN 来提高增强效果。近年来也有相关的研究尝试利用编码辅助信息,如块划分结构、量化系数等,提升压缩视频复原的性能<sup>[15-18]</sup>。Meng 等<sup>[19]</sup>提出了 MGANet,该方法利用变换单元的先验划分信息来引导模型进行学习,但参数较多,训练效率较低。He 等<sup>[15]</sup>设计的神经网络不仅输入解码后的图像,还将压缩过程中产生的块划分信息作为输入,从而提取有效的边界信息,辅助压缩失真的恢复。上述方法采用多帧输入的视

频质量增强是因为连续视频帧的内容具有强相关性,即相邻的视频帧可能包含关于同一对象或场景的有效信息;此外视频编码的本质是选择性地替换相邻帧之间的像素块,相邻帧很可能包含当前帧缺失的信息。因此,利用多帧信息可以有效地提高压缩视频的质量。但是,现有的重建视频画质恢复方法通常仅将自身定位于普通的视频恢复任务,并没有深度探究视频因编码带来的深入影响,这会显著影响重构视频的图像质量恢复效果。

### 1 算法提出动机

考虑到经过编码的视频具有与普通视频不同的特点,本算法并没有将压缩后的视频质量增强任务简单地视为视频去噪或去模糊任务,重点剖析了视频编码与压缩视频质量增强任务之间的本质关系。通过对视频编码过程的反思,总结出压缩视频的2个独有特征。

#### 1.1 质量波动特点

本算法针对 H. 265 标准中的低延迟(low delay, LD)配置下视频质量展开研究。对于 LD 压缩配置下,尽管 QP 直接被设置为 37,实际上视频压缩时每帧都有 QP 偏移。理论上讲,QP 越大,量化步长越大,编码越粗糙,视频帧的编码质量越差;QP 越小,量化步长越小,编码越细致,视频帧的编码质量越好。不同的压缩配置决定了每帧压缩时的真实 QP,而 QP 的值与帧的质量直接相关。因此,不同的

压缩配置导致了压缩视频的质量波动。

在 LD 配置下,BasketballPass 序列的编码质量与 QP 之间的关系如图 1 所示,纵坐标为峰值信噪比(peak signal-to-noise ratio, PSNR)。可以看出,压缩后的视频在 QP 为 38 处存在高质量压缩帧(high quality Frame, HQF),在 QP 为 39、40 处存在低质量压缩帧(low quality frame, LQF),但是视频帧编码的客观质量与 QP 之间并未体现出完全的负相关性,例如,尽管图 1 中第 42 帧的 QP 小于第 41 帧的,但是第 42 帧的客观质量却比第 41 帧的更低。根本原因在于,LD 配置下采用的是“1+x”的帧间编码方案,即在对当前帧进行帧间编码时会利用已编码的最临近的 1 个参考帧和 x 个 HQF,而帧间编码的本质是利用参考帧信息消除视频的时域冗余,因其编码过程不是无损编码,故会引入一定程度的误差累积。第 41 帧作为第 42 帧的帧间编码参考帧,由于其本身已经是 LQF,存在一定误差,故第 42 帧将进一步累积误差,从而导致第 42 帧的客观质量比第 41 帧更低。因此,LD 配置下需要有规律地引入 HQF 来缓解误差的累积。

总体来讲,LD 配置下的压缩视频的确存在着 2 个 HQF 中夹杂 3 个 LQF 的规律且 3 个 LQF 也的确存在着客观质量顺序降低的情况。以上实验仅针对 BasketballPass 序列进行具体分析,目的在于从编码原理的角度对质量波动特点进行详尽的剖析,并不完全具有统计性。本文继而针对场景丰富、分辨率

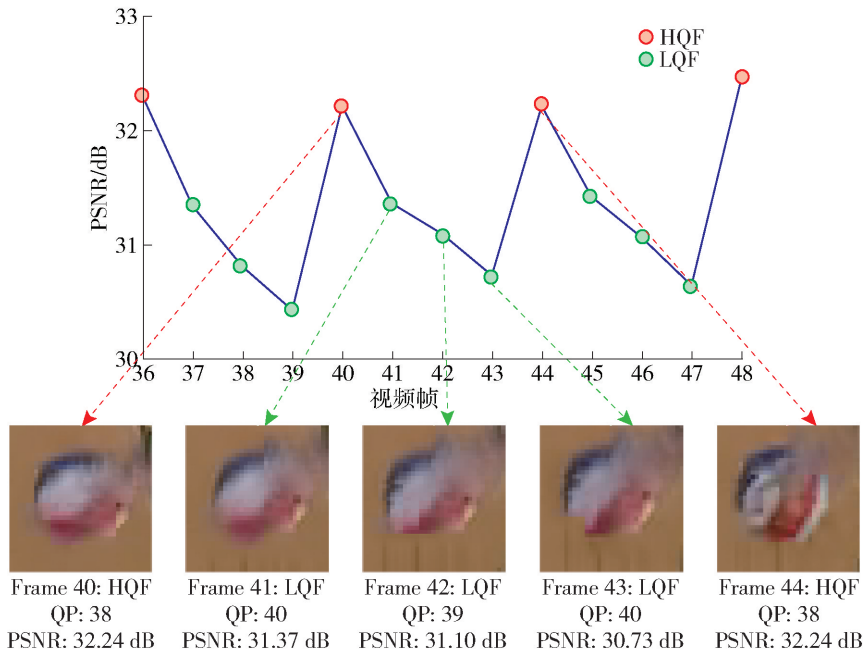


图1 压缩视频的质量波动特点

Fig. 1 Quality fluctuation of compressed video

多样的数个视频进行统计实验,以验证质量波动特点的广泛性。

## 1.2 像素缺失特点

由帧内编码技术和帧间编码技术可知,视频压缩的本质是对视频内容进行预测,即在当前帧和邻近帧中寻找与源像素块相关度最高的最优替代块,而在视频编码强大的压缩效率的要求下,最终选择的替代块可能与源块并不是特别相似。

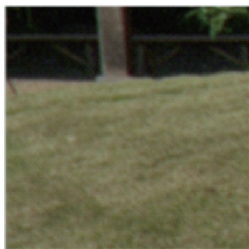
一些压缩后的视频与源视频对应位置的内容存在相当大的差异,损失了相当多的纹理细节,本文称之为像素缺失。压缩视频的像素缺失特点如图2所示。



(a)



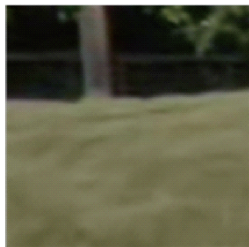
(b)



(c)



(d)



(e)

图2 压缩视频的像素缺失特点

Fig. 2 Pixel missing of compressed video

为了在视频压缩中追求较低数量的传输数据,压缩后的视频不仅导致原始像素信息大量损失,而

且还包含许多模糊区域和伪影。由于有效的像素信息的大量缺乏,使得视频重建困难;同时,大量的模糊区域和伪影也使得网络难以学习特征。因此,本算法在设计网络时,重点关注如何有效地保留和利用有限的原始信息。

## 2 模型设计

根据视频中存在大量的时域信息这一特点,本文提出使用三维卷积神经网络(3D convolutional neural network, 3D-CNN)这种更适合的时空特征提取方法。3D-CNN已成功应用于行为识别等视频高水平视觉任务。本文认为它也可以有效地应用于视频的底层视觉任务,如压缩视频的质量增强。本文以3D-CNN为基础设计了压缩视频质量增强网络(compressed video quality enhancement net, CVQENet),它强大的时空特征表达能力能够有效地解决视频帧的对齐和融合问题。

### 2.1 结构设计

像素缺失特点导致压缩视频存在大量的模糊伪影,因此,视频增强领域被广泛使用的光流计算操作在此处并不十分适用。本文提出的CVQENet网络选用3D-CNN将对齐操作隐含在网络设计中。同时,考虑到压缩视频已经损失较多的细节且时、空下采样操作会明显限制增强网络的能力,故只在网络的重建部分统一进行时、空下采样,以保证输出帧与输入帧具有相同的维度。

CVQENet的网络结构如图3所示。考虑到LD配置下需要有规律地引入HQF来缓解误差的累积,同时结合压缩视频的质量波动特点,本文的网络输入为1帧LQF和2帧HQF,而并非3帧连续帧。同样,每个HQF也通过其最近的HQF得到增强,使得HQF的信息得到高效利用。由于3D-CNNs具有参数量巨大的固有缺陷,本文提出了一种3D-Block来解决这一问题,如图4所示。其中3D时间卷积是卷积核大小为 $3 \times 1 \times 1$ 的Conv3D,3D空间卷积是卷积核大小为 $1 \times 3 \times 3$ 的Conv3D。类似于P3D<sup>[20]</sup>,每个3D-Block由一个相当于二维CNN的

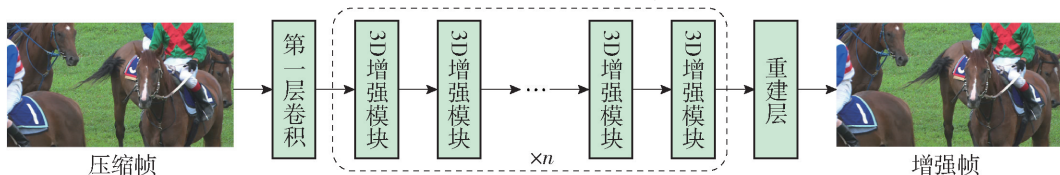


图3 CVQENet模型结构

Fig. 3 Structure of CVQENet

空间卷积滤波器和一个一维 CNN 的时间卷积滤波器组成,并且在每个 3D-Block 中使用残差学习。该 3D-Block 不仅减小了模型的尺寸,而且增加了模型的深度,显著提高了模型的性能。

随着分辨率的提高需要大量的原始信息,许多超分辨率网络为了获得更好的性能,避免丢失更多的信息,并不进行空间下采样。因为本任务中输入的压缩视频丢失了大量的源信息,所以不使用空域的降采样操作,所有卷积步长都设置为 1。

选择 3D-CNN 作为构建块不仅因为它是一个非对齐的网络,而且还因为它具有在时域内表达特征的能力。这也是近年来它被广泛应用于模式识别的原因。由于压缩视频增强的重点在于有效利用帧间信息,所以本文在设计网络时充分利用了 3D-CNN 在时域上的能力。在初始阶段和增强阶段,帧通道数量保持为 5 个,直到构建阶段才对通道进行下采样,以保证在时域内充分融合。

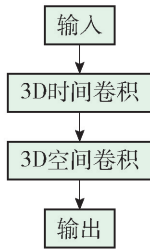


图4 3D-Block 结构  
Fig. 4 3D-Block structure

## 2.2 训练集及网络训练

实验用到的数据集来自 MF-CNN 提供的开源

数据集,共包含 160 段未压缩的视频序列。

CVQENet 模型选用 L1 loss。考虑到直接获取编码视频中运动向量(motion vector, MV)的难度较大,受 MF-CNN 的启发,本文通过最小化当前帧和相邻帧之间的均方误差(mean square error, MSE)训练 CVQENet 模型。

## 2.3 测试条件及环境配置

实验环境对于实验数据的获取有重大影响,本文的实验基于 H. 265 标准参考软件 HM16.9 开展,采用默认低延迟 B 帧(low delay B, LDB)编码配置测定实验数据。此外,为更好地展示本文网络对不同分辨率的序列均有较好的预测性能,本文对 JCT-VC 给出的 5 类(A、B、C、D、E)不同分辨率下的 18 个标准序列在 3 个 QP 值(32、37、42)上进行了验证实验。为了对重建质量有一致的衡量标准,所有实验在 Xeon Gold 6142 CPU、RTX 3090 GPU 的 Linux 操作系统的计算机上进行。

## 3 实验结果与分析

为证实所提出的网络 CVQENet 在压缩视频质量增强方面的能力且考虑到对 LD 配置下的压缩视频增强难度较高,本文使用 PSNR 指标在 LD 配置下将 CVQENet 与压缩伪影卷积神经网络(artifacts reduction convolutional neural network, ARCNN)<sup>[4]</sup>、MemNet<sup>[5]</sup>、去噪卷积神经网络(denoising convolutional neural network, DnCNN)<sup>[6]</sup>、多帧质量增强(multi frame quality enhancement, MFQE)<sup>[13]</sup>、MGANet<sup>[19]</sup>等算法进行比较,实验结果如表 1 所示。

表 1 LD 配置下测试序列的 PSNR 提升结果

Table 1 PSNR improvement of test sequences under LD configuration

类别	视频序列名称	ARCNN	MemNet	DnCNN	MFQE	MGANet	CVQENet	dB
A (2 560 × 1 600)	PeopleOnStreet	0.463 7	0.484 1	0.398 0	0.702 6	0.724 2	0.796 8	
	Traffic	0.267 9	0.229 5	0.250 1	0.286 4	0.460 2	0.447 7	
B (1 920 × 1 080)	Kimono	0.246 0	0.255 7	0.231 9	0.492 1	0.472 9	0.240 0	
	ParkScene	0.169 1	0.175 4	0.155 0	0.240 4	0.234 7	0.276 3	
	Cactus	0.105 3	0.148 6	0.186 0	0.267 6	0.371 9	0.426 0	
	BasketballDrive	0.180 3	0.217 3	0.213 5	0.218 9	0.325 1	0.130 1	
C (832 × 480)	BQTerrace	0.110 2	0.171 7	0.100 8	-0.113 2	0.184 1	0.302 6	
	BasketballDrill	0.155 9	0.150 2	0.105 5	0.176 6	0.415 9	0.465 7	
	BQMall	0.166 7	0.227 0	0.142 1	0.072 5	0.398 3	0.524 8	
	PartyScene	0.022 4	0.096 1	0.005 2	-0.158 9	0.251 9	0.425 7	
	RaceHorsesC	0.144 2	0.173 2	0.117 8	0.005 2	0.186 8	0.228 7	

续表 1

类别	视频序列名称	ARCNN	MemNet	DnCNN	MFQE	MGANet	CVQENet
D (416 × 240)	BasketballPass	0.192 7	0.189 2	0.148 7	0.393 6	0.486 5	0.599 5
	BQSquare	-0.110 8	0.050 8	-0.080 6	-0.441 8	0.278 6	0.614 3
	BlowingBubbles	0.093 3	0.134 7	0.170 1	0.142 6	0.280 3	0.475 7
	RaceHorses	0.257 2	0.291 6	0.227 5	0.386 1	0.317 9	0.411 5
E (1 280 × 720)	FourPeople	0.402 0	0.389 3	0.374 7	0.499 7	0.655 4	0.722 8
	Johnny	0.233 2	0.369 2	0.266 5	0.382 3	0.572 7	0.611 2
	KristenAndSara	0.393 4	0.352 5	0.373 3	0.478 4	0.656 1	0.675 0
平均值	QP37	0.194 0	0.228 1	0.188 1	0.254 5	0.404 1	0.465 2
	QP32	0.121 1	0.158 3	0.110 4		0.352 8	0.382 3
	QP42	0.132 7	0.161 1	0.125 9		0.346 3	0.378 6

CVQENet 在 QP 为 32、37、42 时的平均 PSNR 提升分别为 0.382 3、0.465 2、0.378 6 dB。对于 LD 配置 QP 为 37 时, CVQENet 在 PeopleOnStreet 序列上 PSNR 增强了 0.796 8 dB。以 QP 为 37 为例, 相较于计算机视觉顶级会议 CVPR2018 的 MFQE 方法<sup>[13]</sup>, CVQENet 在参数量上升 10.6% 的情况下, PSNR 的增长量提升了 82.8%; 相较于 DCC2020 的 MGANet<sup>[19]</sup>, 其论文内并未公布参数量, PSNR 的增长量提升了 15.1%。

如图 5 所示, 在 LD 配置下, CVQENet 方法在不同 QP 下对测试序列的平均增强效果均优于所有其他比较方法。因为 CVQENet 是在 QP 为 37 的情况下进行随机初始化训练, 并在 QP 为 32 和 42 的情况下进行微调, 所以在 QP 为 37 的情况下明显比其他 2 种 QP 更优。综合参数量和增强效果两方面比较, CVQENet 是比较方法中最先进的。

LD 配置下 QP 为 37 时, 利用不同方法对

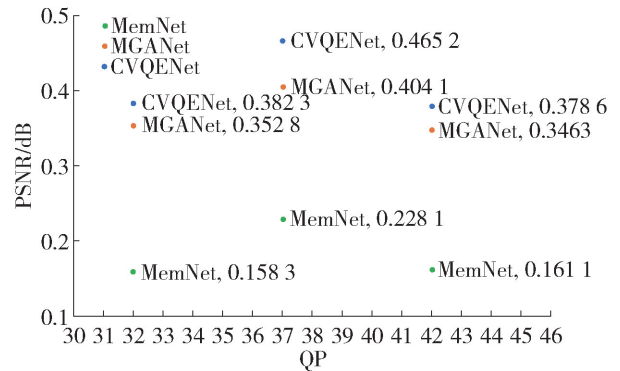


图 5 LD 配置下各算法的增强性能对比

Fig. 5 Enhanced performance comparison of various algorithms under LD configuration

BasketballDrill 序列 51 ~ 70 帧增强后的 PSNR 变化如图 6 所示。可以看出, CVQENet 能够明显增强 LQF, 这使得 PSNR 波动明显小于 H.265/HEVC。综上所述, CVQENet 能够减少视频压缩的质量波

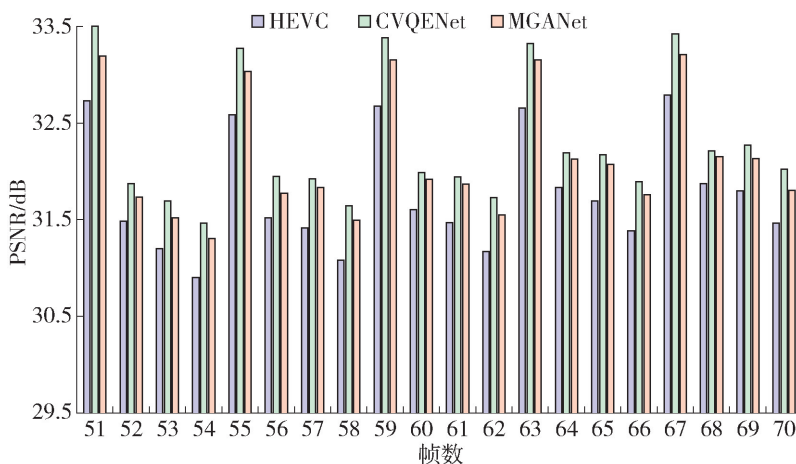


图 6 LD 配置下各算法的 PSNR 波动对比 (QP = 37)

Fig. 6 Comparison of PSNR fluctuations of various algorithms under LD configuration (QP = 37)

动,这对于提升观看压缩视频时的主观感受十分重要。

BasketballDrive 序列、FourPeople 序列、KristenAndSara 序列和 Johnny 序列的原始区域、通

过 H.265 编码后的压缩区域和利用本文提出的 CVQENet 方法进行增强后的增强区域的对比如图 7 所示。可以看出,与压缩视频相比,经过 CVQENet 增强后的视频主观质量有着显著提升。



图7 使用 CVQENet 增强后的画质对比(LD 配置, QP = 37)

Fig. 7 Comparison of image quality enhanced by CVQENet (LD configuration, QP = 37)

## 4 结论

1) 本文通过剖析视频编码技术与压缩视频质量增强任务之间的本质关系,结合大量的主观实验和统计实验,总结出压缩视频存在的质量波动和像素缺失 2 个独有特征。

2) 基于压缩视频的 2 个独有特征,有针对性地提出了一种简单高效的压缩视频质量增强网络 CVQENet。

3) 对提出的 CVQENet 进行对比实验。实验结果表明,CVQENet 在 QP 为 37 时平均 PSNR 提升为 0.465 2 dB,对测试序列的平均增强效果均优于所有其他比较方法,为利用有限的时空信息提高压缩视

频的质量提供了一种探索思路。

## 参考文献:

- [1] SULLIVAN G J, OHM J R, HAN W J, et al. Overview of the high efficiency video coding (HEVC) standard [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2012, 22(12): 1649-1668.
- [2] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution [C] // 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops. Piscataway, NJ: IEEE, 2017: 136-144.
- [3] ZHENG B L, CHEN Y W, TIAN X, et al. Implicit dual-domain convolutional network for robust color image

- compression artifact reduction[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2020, 30(11): 3982-3994.
- [4] DONG C, DENG Y B, LOY C C, et al. Compression artifacts reduction by a deep convolutional network[C]//2015 IEEE International Conference on Computer Vision. Piscataway, NJ: IEEE, 2015: 576-584.
- [5] TAI Y, YANG J, LIU X M, et al. MemNet: a persistent memory network for image restoration[C]//2017 IEEE International Conference on Computer Vision. Piscataway, NJ: IEEE, 2017: 4539-4547.
- [6] ZHANG K, ZUO W M, CHEN Y J, et al. Beyond a Gaussian denoiser: residual learning of deep CNN for image denoising[J]. *IEEE Transactions on Image Processing*, 2017, 26(7): 3142-3155.
- [7] WANG T T, CHEN M J, CHAO H Y. A novel deep learning-based method of improving coding efficiency from the decoder-end for HEVC[C]//2017 Data Compression Conference, Piscataway, NJ: IEEE, 2017: 410-419.
- [8] YANG R, XU M, LIU T, et al. Enhancing quality for HEVC compressed videos[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2019, 29(7): 2039-2054.
- [9] YANG R, SUN X Y, XU M, et al. Quality-gated convolutional lstm for enhancing compressed video[C]//2019 IEEE International Conference on Multimedia and Expo, Piscataway, NJ: IEEE, 2019: 532-537.
- [10] YANG R, MENTZER F, VAN GOOL L, et al. Learning for video compression with hierarchical quality and recurrent enhancement[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Piscataway, NJ: IEEE, 2020: 6628-6637.
- [11] FENG L T, ZHANG X F, WANG S S, et al. Coding prior based high efficiency restoration for compressed video[C]//2019 IEEE International Conference on Image Processing. Piscataway, NJ: IEEE, 2019: 769-774.
- [12] GUAN Z Y, XING Q L, XU M, et al. MFQE 2.0: a new approach for multi-frame quality enhancement on compressed video[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021, 43(3): 949-963.
- [13] YANG R, XU M, WANG Z L, et al. Multi-frame quality enhancement for compressed video[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2018: 6664-6673.
- [14] WANG T T, XIAO W H, CHEN M J, et al. The multi-scale deep decoder for the standard HEVC bitstreams[C]//2018 Data Compression Conference. Piscataway, NJ: IEEE, 2018: 197-206.
- [15] HE X Y, HU Q, ZHANG X Y, et al. Enhancing HEVC compressed videos with a partition-masked convolutional neural network[C]//2018 25th IEEE International Conference on Image Processing. Piscataway, NJ: IEEE, 2018: 216-220.
- [16] MA L, TIAN Y, HUANG T. Residual-based video restoration for HEVC intra coding[C]//2018 IEEE Fourth International Conference on Multimedia Big Data. Piscataway, NJ: IEEE, 2018: 1-7.
- [17] KANG J H, KIM S, LEE K M. Multi-modal/multi-scale convolutional neural network based in-loop filter design for next generation video codec[C]//2017 IEEE International Conference on Image Processing. Piscataway, NJ: IEEE, 2017: 26-30.
- [18] SONG X D, YAO J B, ZHOU L L, et al. A practical convolutional neural network as loop filter for intra frame[C]//2018 25th IEEE International Conference on Image Processing. Piscataway, NJ: IEEE, 2018: 1133-1137.
- [19] MENG X D, DENG X, ZHU S Y, et al. MGANET: a robust model for quality enhancement of compressed video[EB/OL]. [2022-07-10]. <https://arxiv.org/abs/1811.09150>.
- [20] QIU Z F, YAO T, MEI T. Learning spatio-temporal representation with pseudo-3D residual networks[C]//2017 IEEE International Conference on Computer Vision. Piscataway, NJ: IEEE, 2017: 5533-5541.

(责任编辑 梁洁)