

H.264 视频编码率失真优化 和码率控制技术研究进展

崔子冠^{1,2}, 朱秀昌^{1,2}, 干宗良^{1,2}, 唐贵进^{1,2}, 刘 峰^{1,2}

(1. 南京邮电大学江苏省图像处理与图像通信重点实验室, 江苏南京 210003;
2. 南京邮电大学宽带无线通信与传感网技术教育部重点实验室, 江苏南京 210003)

摘 要: 在 H.264 率失真优化和码率控制(RC)技术研究的基础上,根据控制目的将率失真优化和 RC 的发展分为 3 个阶段:即率失真建模,考虑质量波动对主观质量的影响,以主观失真为指导的 RC. 分析了各阶段有代表性的 RC 技术及其特点,指明了 RC 需要进一步研究的关键和难点问题. 对面向传输的错误弹性编码、可分级编码、多描述编码、以及立体与多视编码中的 RC 技术进行总结,分析进一步发展趋势. 对下一代高性能视频编码(HEVC)中的 RC 应该如何发展进行了探讨.

关键词: 视频编码; 率失真优化; 码率控制; H.264

中图分类号: TN919.8 **文献标识码:** A **文章编号:** 0372-2112 (2013) 12-2443-08

电子学报 URL: <http://www.ejournal.org.cn>

DOI: 10.3969/j.issn.0372-2112.2013.12.019

Advances in Rate Distortion Optimization and Rate Control Techniques for H.264 Video Coding

CUI Zi-guan^{1,2}, ZHU Xiu-chang^{1,2}, GAN Zong-liang^{1,2}, TANG Gui-jin^{1,2}, LIU Feng^{1,2}

(1. *Image Processing and Image Communication Lab, Nanjing University of Posts and Telecommunications, Nanjing, Jiangsu 210003, China;*

2. *Key Lab of Broadband Wireless Communication and Sensor Network Technology, Nanjing University of Posts and Telecommunications, Ministry of Education, Nanjing, Jiangsu 210003, China*)

Abstract: Based on the researches of rate distortion optimization (RDO) and rate control (RC) techniques for H.264, this paper divides the development of RDO and RC into three phases based on RC intention, namely rate distortion modeling, considering the effect of quality fluctuation on subjective quality, subjective distortion guided RC; analyzes the characteristics of representative RC techniques of each phase, and focuses on the key and difficult issues of RC needed to be further researched. Then summarizes the RC schemes for transmission oriented error resilient video coding, scalable video coding (SVC) and multiple description coding (MDC) as well as stereo and multi-view video coding, and analyzes further research trends. Finally, how to develop the RC schemes for high efficiency video coding (HEVC) is discussed.

Key words: video coding; rate distortion optimization; rate control; H.264

1 引言

目前主流的 H.264 编码标准基于预测加变换的混合框架,采用多种帧内与帧间模式和基于率失真优化(Rate Distortion Optimization, RDO)的模式选择等技术,获得了极高的压缩效率. 下一代高性能视频编码(High Ef-

iciency Video Coding, HEVC)即 H.265 在许多编码模块中进一步改进,其压缩效率比 H.264 又提高一倍. 码率控制(Rate Control, RC)是视频编码(VC)中的关键模块,根据各种约束(如目标码率、缓冲区、延迟、计算复杂度、信道状况等)自适应调节编码参数来规范输出码流使之适合信道带宽,并最优化视频质量. RC 分两步:一是在

收稿日期:2012-12-25;修回日期:2013-07-01;责任编辑:孙瑶

基金项目:国家自然科学基金(No. 61071091, No. 61071166, No. 61071167, No. 61101105);江苏省高校自然科学基金(No. 12KJB510019);南邮校科研基金(No. NY212015, No. NY212083);教育部博士点基金(No. 20113223120001);工业和信息化部通信软科学课题(No. 2011-R-70);江苏高校优势学科建设工程—“信息与通信工程”;江苏省自然科学基金(No. BK20130867)

约束条件下为编码单元(Basic Unit, BU)分配最优目标比特;二是根据率失真模型(Rate Distortion Model, RDM)设计准确的控制技术使实际码率尽量接近目标码率.公式表述如下:

$$\begin{aligned} \min & \frac{1}{N} \sum_{i=1}^N D_i(Q_i), \\ \text{s. t.} & \sum_{i=1}^N R_i(Q_i) \leq R_T \end{aligned} \quad (1)$$

其中 Q_i 为量化参数, $D_i(Q_i)$ 与 $R_i(Q_i)$ 为第 i 个 BU 的失真与码率, R_T 为目标码率. 此有约束的最优化问题可由 Lagrange 乘子法转化为无约束问题, 建立 Lagrange 代价函数, 通过调节参数使代价函数最小即可, 即:

$$\min J = D + \lambda R \quad (2)$$

其中 λ 是 Lagrange 乘子, J 是代价函数.

鉴于 RC 对 VC 的重要性, 近年来针对 H.264 在各种环境下的编码与传输问题, 已经提出了很多有效的 RC 技术. 文献[1]对 2007 年之前的 RC 技术进行了综述, 但不断出现的新需求使 RC 的发展日新月异. 本文主要对 2007 年以后 RC 的最新发展进行总结. 第 2 节根据控制目的将单层 VC(即一个信源编码产生一个码流)的 RC 分为 RDM 建模、考虑质量波动对主观质量的影响、以主观失真为指导的 RC 这 3 个阶段, 总结各阶段有代表性的 RC 技术及其特点. 在此基础上, 第 3 节重点指明单层 VC 的 RC 需要进一步研究的关键和难点问题. 第 4 节对面向传输的错误弹性编码(Error Resilient Video Coding, ERVC)、可分级编码(Scalable Video Coding, SVC)、多描述编码(Multiple Description Coding, MDC)、以及立体与多视编码(Multi-View Video Coding, MVC)中的 RC 技术进行总结, 分析进一步发展趋势. 第 5 节对 HEVC 中的 RC 应该如何发展进行探讨. 第 6 节总结全文.

2 单层 VC 中 RDO 和 RC 研究进展

2.1 RDM 建模

H.264 在 2003 年颁布之后, 由于其高效的压缩性能和网络适应性, 基于 H.264 的各种应用开始蓬勃发展, 亟需准确的 RDM 和高效的 RC 技术以精确控制码率. H.264 基于 RDO 选择最优的编码模式, 新技术的引入使得编码残差呈现出与以往标准不同的特性. 因此这一阶段的主要任务在于考虑 H.264 编码特性对残差分布准确建模, 并建立准确的 RDM. 目前精确控制 H.264 的 I 帧及 P 帧输出码率的目标已基本实现. 由于 I 帧及 P 帧编码特性不同, 导致 RDM 也不相同. 下面分别对 I 帧及 P 帧的残差分布及有代表性的 RDM 进行对比和分析.

2.1.1 P 帧残差分布及有代表性的 RDM

残差 DCT 系数的分布对 RDM 建模非常重要, 文献

[2]认为 DCT 系数服从 Laplacian 分布并给出严格数学分析. 目前 H.264RC 中广泛使用的二次 RDM 就是根据 Laplacian 分布推导而来^[3,4]. 文献[5]发现 Laplacian 分布具有指数形式, 导致拖尾衰减过快, 而实际 DCT 系数直方图常有大量的拖尾系数, 因此用 Cauchy 分布对 DCT 系数进行建模更准确. 文献[6]直接建立 R 及 D 与 DCT 系数为零的比率 ρ 之间的关系, 也取得了准确的控制效果. 基于这些统计分布提出的有代表性的 RDM 及其特点总结于表 1. 注意, 这一阶段 RC 的目标主要是为了精确控制 R , 因此 RDM 常表现为 R -量化步长(Q_{step})的函数关系, 而 D 模型常用简单的 $D = kQ_{\text{step}}$ 或 $D = kQ_{\text{step}}^2$ 这种形式, 即认为 D 与 Q_{step} 或 Q_{step}^2 成正比. 如 H.264 最有代表性的 RC 提案 JVT-G012^[3]采用线性 MAD 预测模型解决蛋鸡悖论, 根据流量传输模型为 BU 分配目标比特, 并用二次 RDM 计算 Q_{step} , 以较低的计算复杂度获得了较好的 RC 效果.

基于 RDM 的 RC 算法具有类似的控制框架, 在此框架下很多学者提出了有效的改进机制, 主要集中于 RDM 的进一步改进^[7]、BU 目标比特的优化分配、BU 的灵活划分、头码率的准确估计、缓冲区与延迟的有效管理、编码复杂度的有效测量、模型参数更新时历史数据点的有效选择等方面.

表 1 P 帧残差分布及有代表性的 RDM

残差分布	文献	RDM
Laplacian 分布	[3]	$R-H = (a/Q_{\text{step}} + b/Q_{\text{step}}^2) \times \text{MAD}$
	[4]	$R-H = A \times Q_{\text{step}}^2 + B \times Q_{\text{step}} + C$
	[8]	$R-H = \alpha \cdot \text{SATD}(Q_{\text{step}})/Q_{\text{step}}$
Cauchy 分布	[5]	$R = aQ_{\text{step}}^{-\alpha}, D = bQ_{\text{step}}^2$
ρ 域模型	[6]	$R(\rho) = \theta \times (1 - \rho), D(\rho) = \sigma^2 e^{-\alpha(1-\rho)}$

2.1.2 I 帧复杂度测量及 RC

编码常采用 GOP 结构, 大部分帧都采用 P 帧编码, 因此 P 帧 RC 研究较多而 I 帧 RC 研究较少, 但 I 帧作为每个 GOP 的起始参考帧非常重要, 有效进行 I 帧 RC 将有助于提高整个 GOP 的编码质量. 文献[9]采用梯度信息测量 I 帧复杂度, 计算简单但与 I 帧编码特性不符, 改进效果有限. 文献[10]用统计窗中所有像素与其均值的平均绝对差测量 I 帧复杂度, 且分析 4×4 块是否具有水平、垂直、及左右对角线方向的条纹, 若存在复杂度即设为零. 此方法是想将 H.264 帧内预测特性考虑进去, 但只判断 4 种条纹方向及判断出条纹后复杂度即设为零与 H.264 帧内编码特性不符, 导致 I 帧复杂度测量仍不够准确.

I 帧 RC 要解决好 3 个问题: (a) 准确的 I 帧复杂度

测量; (b) I 帧目标比特最优分配; (c) 准确的 I 帧 $R-Q$ 模型. 针对这 3 个问题, 文献[11]通过边缘检测估计 4×4 块的预测模式和残差, 用残差的平均绝对值测量 I 帧复杂度, 之后提出一种新的 $R-Q_{\text{step}}$ 模型, 同时考虑缓冲区和序列特性为 I 帧分配合适的目标比特, 得到了较好的 I 帧 RC 效果, IPPP 方式编码时序列 PSNR 波动减少了 60% 以上. 文献[11]与 JVT-G012^[3]及 JVT-W042 对 Mobile 序列编码的效果如图 1.

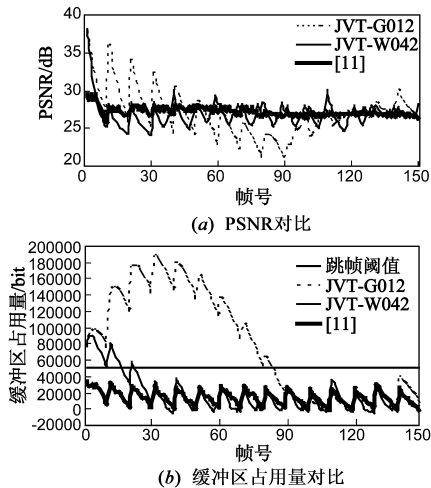


图1 Mobile序列编码结果对比

2.2 考虑质量波动对主观质量的影响

通过建立准确的 $R-Q$ 模型精确控制输出码率的目标基本完成之后, RC 研究的重点开始转向在 R 限制下为用户提供更好的主观感受, 其中最小化失真变化的 RC 和基于感兴趣区域(ROI)的 RC 很有代表性.

2.2.1 最小化失真变化的 RC

之前 RC 的目标都是在 R 限制下最小化平均失真(MINAVE), 由于视频内容的多样性这种优化策略在编码时可能产生明显的质量波动, 从而影响主观感受. 针对 R 限制下对质量平稳性要求严格的应用(如视频流), 就需要将最大失真(MINMAX)或失真方差(MINVAR)最小化, 目的在于得到平滑一致的视频质量. 公式表示如下:

$$\begin{aligned} \min & \left(\max_{i \in \{1, \dots, N\}} D_i(Q_i) \right), \\ \text{s. t.} & \sum_{i=1}^N R_i(Q_i) \leq R_T \end{aligned} \quad (3)$$

如文献[12], 为常质量应用, 基于 MINMAX 提出一种帧层多次编码方法, 在最大失真与平均失真的控制上都好于 JM 算法. 文献[13]在 R 约束下对时域滑动窗中的帧实施 MINVAR 准则, 并用 Lagrange 乘子法进行求解, 得到了一致的图像质量. 文献[14]对质量平滑和码率(缓冲区)平稳这一矛盾问题进行研究, 提出自适应窗口大小机制在保持质量平滑的同时保证缓冲区在安

全水平以内.

2.2.2 基于 ROI 的 RC

针对传统 RC 把图像所有区域同等看待的缺点, 利用人眼对图像不同区域的关注度不同, 首先检测出 ROI 区域如人脸、运动目标等, 之后在 R 限制下对 ROI 细量化, 对非 ROI 粗量化, 以保证 ROI 的空时质量来优化主观感受^[15]. 基于 ROI 的 RC 在低码率应用中(如视频会议)有广泛的应用前景. 难点在于: (a) 编码之前要先检测出 ROI, 准确地分割视频对象需要较高计算量而较小的计算量无法得到准确的检测效果; (b) 根据关注度设定经验加权系数在 ROI 与非 ROI 之间进行 R 分配和 QP 调整, 会产生较明显的方块效应; (c) 最大问题在于 RDM 仍以(加权)MSE 作为失真度量, 其编码结果是使(加权)PSNR 最优, 与主观质量仍不完全一致.

2.3 以主观失真为指导的 RC

RC 第二阶段开始考虑为用户提供更好的主观感受, 但失真准则仍是客观失真. 基于均方误差的峰值信噪比(PSNR)计算简单且物理意义明确, 在视频质量评价中广泛应用, 但 PSNR 与主观质量并不成正比, 这是由于人类视觉系统(HVS)对视频质量的感知是高度非线性的, 以 PSNR 为指导的 RC 不能得到最优的主观质量. 因此如何根据 HVS 特性建立一种主观质量评价方法并用以指导视频编码成为一项重要的研究课题^[16,17]. 即 RC 第三阶段的研究重点由第一阶段 R 模型的精确建模(以精确控制输出码率)转到主观 D 模型的有效建模(以获得最优的主观 $R-D$ 性能). 视频质量专家组(VQEG)和相关学者在这方面也做了初步研究, 其中最具代表性的有基于视觉失真敏感度(VDS)或显著度(Saliency)的 RC、基于恰可察觉失真(JND)的 RC 和基于结构相似性(SSIM)的质量评价与编码方法.

2.3.1 基于 VDS 或 Saliency 的 RC

这类方法与基于 ROI 的 RC 有些类似, 区别在于更充分地利用 HVS 特性建立 VDS 或 Saliency 图, 以此进行优化的 R 分配. 相比 ROI 与非 ROI 两级划分, R 分配更加合理, 主观质量更优. 如文献[18]根据运动、纹理及空时掩模效应建立 VDS 图进行 R 分配以优化视觉质量. 文献[19]融合运动、色彩、方向等多种特征建立 Saliency 图来指导 R 分配.

2.3.2 基于 JND 的 RC

JND 指人眼不能感知到的最大失真, 即人眼对 JND 阈值以下的 D 不敏感. JND 可对感知冗余有效建模, 考虑了空时对比敏感度、亮度适应效应、对比度掩模效应等对主观失真建模及 R 调节具有很好的指导作用, 如可对感知重要的区域细量化以保证质量而对感知冗余进行有效删除. 文献[20]基于多种空时视觉特性建立 JND 模型以此指导视频编码. 文献[21]根据视网膜的凹

性(Foveation)增强 JND 提出 FJND 模型,并用于 MB 层 QP 及 λ 调整.

2.3.3 基于 SSIM 的质量评价与编码方法

文献[22]基于 HVS 对结构信息更加敏感的特性提出基于块的 SSIM 测量,克服了 PSNR 基于点而忽略像素之间相关性的缺点,与主观质量非常一致且计算量较小,已被引入 JM 中以评价编码视频的主观质量.基于 SSIM 的质量评价及编码方法近年来得到广泛研究,但 SSIM 形式较为复杂,其与 R 或 Q 之间的闭解形式很难得到,是 SSIM 用于 VC 的难点所在.文献[23]提出 SSIM 失真与 R 存在指数关系,在 BU 层进行基于 SSIM 失真的最优 R 分配;之后将 SSIM 失真引入 RDO 模式选择中,发现基于 SSIM 的 R - D 曲线与基于 MSE 的 R - D 曲线相邻点处切线的坡度非常相似,使用之前关键帧的 RDM 和前一编码帧的 R - D 数据对来自适应的预测当前帧的 λ .文献[24,25]将 SSIM 失真引入 RDO 框架后,先提出一种减少参考的 SSIM 估计模型并结合 R 模型提出了帧层自适应的 λ ,之后在 MB 层进一步根据运动信息和视觉速度感知的不确定性来调节 λ .文献[26,27]首次提出 SSIM- Q_{step} 存在线性关系,并结合 R 模型进行 SSIM 最优的 MB 层 RC;进一步将 SSIM 引入 RDO 框架,提出 MB 层自适应的分析型 λ 来更好地调节 R 和 SSIM 失真.相比之前算法的最大贡献在于得到了 λ 与 Q_{step} 以及图像特性(复杂度)之间的解析式.文献[26,27]与 JVT-G012^[3]对 Deadline 序列的编码效果如图 2.

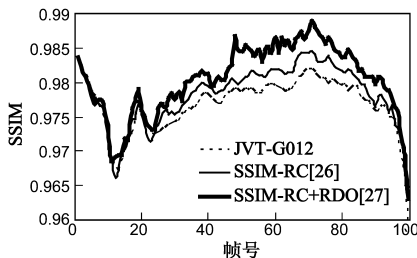


图2 Deadline序列编码结果(96kbps)

3 单层 VC 需要进一步研究的问题

(1)对场景变换(SC)的处理及 GOP 变长编码.目前编码大多采用固定 GOP 结构,即周期性地插入 I 帧而不考虑序列特性.当序列包含 SC 或剧烈运动时编码效率很低且质量波动较大,在低码率应用中尤其严重,如果在编码时考虑 SC 并自适应地插入 I 帧则可更好地对 P 帧进行预测从而提高编码性能.如文献[28]基于 SC 检测提出了 GOP 长度自适应的编码机制.

(2) λ 自适应调整.在 MB 层 QP 确定之后,H.264 引入 λ 进行 RDO 模式选择, λ 仅是 QP 的函数并控制着 R 与 D 之间的权衡.不同图像具有不同的率失真特性,根

据图像内容自适应地调整 λ 有望获得更好的率失真性能.如文献[29]对基于 MSE 的 RDO 中的 λ 在 MB 层根据编码复杂度自适应调整.文献[30]发现 λ 与视频信号的 R - D 特性之间存在线性关系,提出一种启发式的不依赖于 QP 的 λ 计算方法,改进了编码效率.

(3)多维 RC.传统 RC 都是在 R 约束下调节 QP 来优化视频质量,而帧率与图像分辨率常作为预定值独立于 QP.这类机制在高码率应用中性能较好,但在低码率应用中常导致空间质量很低甚至被动跳帧引起时域质量的大幅波动.如 JM 通过监测缓冲区,一旦超过阈值(缓冲区的 80%)就跳过下一帧,这种被动跳帧在以下两种情况经常发生:一是 I 帧编码后由于 R 剧增导致后续 P 帧被连续跳过;二是剧烈运动或 SC 时,由于 I 宏块数增加导致后续带有重要信息的帧被跳过.HVS 对低运动序列更注重空间质量而对高运动序列更注重时域连续性^[31],因此低码率应用中可以根据序列特性和缓冲区联合调节帧率、空间分辨率和 QP(即多维 RC)以优化空时质量.如何准确估计跳过帧与编码帧的失真,并用以确定跳帧数目和下一编码帧的 QP 使局部窗中帧平均失真最小是研究难点.编码端自适应跳帧与解码端插帧相结合^[32],以及编码端空间下采样与解码端上采样相结合^[33]来提高编码效率和解码端空时质量是今后一个重要的研究课题.

(4)便于硬件实现的 RC.传统 RC 都是基于软件实现的,由于算法的复杂性不适合硬件系统采用.但近年来随着视频电话、数字摄像机等广泛应用,亟需适合硬件实现的 RC 机制来提高这些设备的编码性能.便于硬件实现的 RC 算法^[34]需要具备以下特性:(a)低复杂度,考虑到计算能力和功耗;(b)较小的缓冲区需求和延迟,考虑到编码实时性;(c)算法设计上要便于硬件实现,比如指令级优化.

(5)联合功率或复杂度 RDO(PRDO、CRDO)的 RC.随着便携设备和网络技术的发展,带宽不再是限制视频通信的唯一条件,便携设备的功耗和续航能力也是设计编码算法时要考虑的关键因素之一.功耗主要来自于编码与传输两部分,编码功耗随计算复杂度单调增加,而传输功耗与压缩码流的大小成正比.因此如何在功耗或复杂度约束下最优化视频的 R - D 性能成为一项重要的研究课题.目前大多数研究是将功耗约束转化为计算复杂度约束,并在传统 R - D 分析的基础上加上功率或复杂度约束,建立统一的功率^[35]或复杂度 R - D 分析模型^[36],通过调整复杂度控制参数达到功耗和 R - D 性能的最佳.关键问题在于:(a)建立合适的功耗或计算复杂度测量方法;(b)建立简单有效的 PRD 或 CRD 分析模型;(c)对分析模型的有效求解.

(6)视频流中的 RC 及多路复用中的 RC.随着

H.264 在视频流和多路复用中的广泛应用,需要有效的 RC 算法在延迟受限的情况下进行常质量控制.文献[37]为视频流提出一种延迟失真优化(DDO)分析框架,以寻求延迟与失真之间的最优平衡来保证接收端连续高质量的播放.文献[38]为多路视频在共享信道中的高效传输使用双帧技术进行码流间及帧间联合最优比特分配,以最小化总体 MSE 失真.另外把编码端的 R - D 分析与网络参数联合优化以提供端到端的服务质量保证(QoS)也是将来的一个研究热点.

(7) 视频转码的 RC. 转码是指码流从一个码率转换到另一个码率,或在不同编码格式之间相互转化.其中 RC 是视频转码中的关键问题^[39],对码率自适应、码流串连影响重大,RC 的好坏直接影响转码过程中引入的质量下降程度.

4 面向传输及多视中的 RC 进展

4.1 ERVC 与联合信源信道 RC

无线网络中如何在编码端根据丢包特性通过调节编码参数以提高码流的抗误码性能是一项重要的研究课题.H.264 的抗误码机制主要有:帧内 MB 刷新,自适应参考帧,自适应冗余图像,数据分区等.帧内 MB 刷新主要有随机或定期刷新、基于运动信息或视觉关注度的刷新,但这类机制没有充分考虑信源与信道的状况无法达到编码效率与抗误码性能的最优平衡.由于信源和信道的时变特性,帧内 MB 刷新数目和位置的选择至关重要,因为帧内 MB 会降低压缩效率而位置不佳将不能阻止误码传播,因此有必要建立一种针对丢包环境下的 RDM 和 RC 机制在 R - D 最优意义下选择帧内刷新率与刷新位置^[40].自适应参考帧^[41]是为了解决由于信道误码造成的编码端与解码端参考帧不匹配的问题.自适应冗余图像^[42]通过编码适当的冗余信息以减轻信道误码对解码质量的影响.数据分区可以将编码信息按照重要性分开传输,以便对重要数据实施更强的保护.

联合信源信道 RC 与 ERVC 类似,也是基于端到端 RDM 联合最优地选择编码模式和 QP.文献[43]针对丢包网络提出一种 RDM,克服了 RC 与误码恢复互不相关的缺点,在 R 限制下能充分利用信道带宽并使得信源编码与信道传输的总失真最小.文献[44]把端到端总失真视为信源量化失真、错误传播失真和错误隐藏失真三项之和,很好地揭示了丢包环境下真正的 R - D 关系.

无线网络鲁棒视频传输最核心的问题在于:(a)端到端失真的准确建模;(b)在编码端准确估计解码端失真;(c)用总失真指导编码参数的调整.如文献[45]的帧层失真估计尤其考虑了 H.264 的新特性(如分像素的运动补偿、帧内预测和去块效应滤波等内插操作)对

失真建模的影响.文献[46]使用有限状态 Markov 链对连续丢包信道中的传输失真进行建模.最近文献[47, 48]对传输失真中的截断与误码扩散两个重要特性进行建模,并根据信源与信道特性得到了传输失真的递归计算公式,相比之前的失真模型更加准确而且计算复杂度更低.

4.2 SVC 的 RC

SVC 将视频编解码成一个基本层和多个增强层,解码端根据运算能力及网络状况选择解码部分码流以获得所需的视频质量.SVC 中 RC 的关键问题在于根据各层之间相关性进行最优的 R 分配与控制^[49,50],难点在于:(a)根据不同层的特性提出更加准确的 RDM;(b)不同层之间的最优 R 分配;(c)如何保证 SVC 编码质量的平稳性.

4.3 MDC 的 RC

MDC 相比 SVC 的优点是每个描述均可独立解码,而 SVC 中基本层一旦丢失,则增强层无用.MDC 中 RC 的关键问题在于根据不同的多描述方案^[51]在多个描述之间最优分配 R 以最小化 D .目前 MDC 的 RC 算法较少且控制效果一般,这是因为不同的 MDC 方案会产生不同的编码特性,导致不同描述及其之间的 R - D 特性难以建模.因此 MDC 的 RC 需要根据具体的多描述方案来分析 R 和 D 特性,如文献[52]为基于分层次 B 帧的 MDC 进行最优 R 分配以平衡编码效率和错误弹性.

4.4 立体与 MVC 的 RC

立体与 MVC 除了提供 2D 视频信息,还提供表现 3D 场景的深度或多视信息,丰富了用户体验,必将成为今后 VC 领域的研究热点.立体与 MVC 中 RC 的关键问题在于 2D 视频与深度之间以及多个视点之间有效分配 R 以最优视频感知质量,其中深度序列编码的 R - D 特性和不同视点之间相关性分析与建模是 RC 成功的必要条件,这也是立体与 MVC 中 RC 今后研究的热点,比如文献[53]为分别从视点层、视频/深度层、及帧层分配 R ,满足了信道与缓冲区的需求;文献[54]利用二次 R 模型和线性 D 模型对 3DTV 广播的左/右视点的 R 进行联合控制.

另外,2DVC 中基于主观质量或 ERVC 思想都可以推广到 3DVC 的 RC 中以改进主观 R - D 性能与错误鲁棒性.如文献[55]为 MVC 提出基于立体视觉关注度的 R 分配算法.文献[56]对 MVC 在丢包网络中传输由于随机丢帧产生的信道失真进行建模.

5 HEVC 的 RC

针对高清视频的 HEVC 在 2013 年颁布之后,必然会在视频编码领域引发新一轮的研究热潮.但截止目前,极少有对 HEVC 中 RC 进行研究的文献出现.HEVC

中 RC 研究的关键问题在于:(a)考虑新的编码技术对纹理信息和头信息码率所产生的影响,建立准确的纹理与非纹理 RDM 来指导 R 的分配与控制,在准确控制 R 的同时提高编码 $R-D$ 性能;(b)引入主观失真,在一定 R 下通过调节 QP 以最优优化主观质量.文献[57]认为 HEVC 编码单元或变换单元预测与变换的深度不同,四叉树预测残差的统计特性也不同,用 Laplacian 混合分布来描述残差,得到 HEVC 编码的 R 和 D 模型.文献[58]同样基于 Laplacian 分布推导出 $R-Q$ 与 $D-Q$ 模型,计算 QP 时同时考虑了帧间质量波动和缓冲区约束,得到了一致的视频质量.

6 结论

在对 RC 最新研究成果的分析可见,针对 H.264 的基于客观质量的 RDM 和 RC 研究得已较为成熟,而主观失真建模及在 RC 中的应用、面向无线网络的 ERVC、面向异构网络的 SVC 及 MDC 中的 RC、丰富用户体验的立体与 MVC 中的 RC、以及针对高清视频的 HEVC 中的 RC 研究成果较少,控制精度不够且复杂度高,将成为今后 RC 技术研究的重点和难点.

参考文献

- [1] Chen Z Z, Ngan K N. Recent advances in rate control for video coding[J]. Signal Processing: Image Communication, 2007, 22(1): 19 – 38.
- [2] Lam E, Goodman J W. A mathematical analysis of the DCT coefficient distributions for images[J]. IEEE Transactions on Image Processing, 2000, 9(10): 1661 – 1666.
- [3] Li Z G, Pan F, Lim K P. Adaptive basic unit layer rate control for JVT[A]. Proceedings of the 7th JVT-G012 Meeting[C]. Pattaya II, Thailand: JVT, 2003. 1 – 33.
- [4] 袁武, 林守勋, 牛振东, 等. H.264/AVC 码率控制优化算法[J]. 计算机学报, 2008, 31(2): 329 – 339.
Yuan Wu, Lin Shou-xun, Niu Zhen-dong, et al. Efficient rate control schemes for H.264/AVC[J]. Chinese Journal of Computers, 2008, 31(2): 329 – 339. (in Chinese)
- [5] Kamaci N, Altunbasak Y, Mersereau R M. Frame bit allocation for the H.264/AVC video coder via Cauchy-density-based rate and distortion models[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2005, 15(8): 994 – 1006.
- [6] He Z, Mitra S K. Optimum bit allocation and accurate rate control for video coding via ρ -domain source modeling[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2002, 12(10): 840 – 849.
- [7] 霍炎, 李生红, 荆涛. 基于广义拉普拉斯分布的低时延速率控制算法[J]. 电子学报, 2010, 38(5): 1078 – 1083.
Huo Yan, Li Sheng-hong, Jing Tao. Low-delay rate control based on generalized Laplacian distribution[J]. Acta Electronica Sinica, 2010, 38(5): 1078 – 1083. (in Chinese)
- [8] Kwon D, Chen M, Kuo C. Rate control for H.264 video with enhanced rate and distortion models[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2007, 17(5): 517 – 529.
- [9] Yan B, Sun K. Joint complexity estimation of I-frame and P-frame for H.264/AVC rate control[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2012, 22(5): 790 – 798.
- [10] Lee G, Lin H, Wang M. Rate control algorithm based on intra-picture complexity for H.264/AVC[J]. IET Image Processing, 2009, 3(1): 26 – 39.
- [11] 崔子冠, 朱秀昌. H.264 图像复杂度自适应的 I 帧码率控制算法[J]. 电子与信息学报, 2010, 32(11): 2547 – 2553.
Cui Zi-guan, Zhu Xiu-chang. Image complexity adaptive intra-frame rate control algorithm for H.264[J]. Journal of Electronics and Information Technology, 2010, 32(11): 2547 – 2553. (in Chinese)
- [12] Cherniavsky N, Shavit G, Ringenburt M F, et al. MultiStage: A MINMAX bit allocation algorithm for video coders[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2007, 17(1): 59 – 67.
- [13] Lee H, Sull S. A VBR video encoding for locally consistent picture quality with small buffering delay under limited bandwidth[J]. IEEE Transactions on Broadcasting, 2012, 58(1): 47 – 56.
- [14] Xu L, Li S, Ngan K N, et al. Consistent visual quality control in video coding[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2013, 23(6): 975 – 989.
- [15] Wu G, Fu Y, Huang S, et al. Perceptual quality-regulable video coding system with region-based rate control scheme[J]. IEEE Transactions on Image Processing, 2013, 22(6): 2247 – 2258.
- [16] Naccari M, Pereira F. Advanced H.264/AVC-based perceptual video coding: architecture, tool, and assessment[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2011, 21(6): 766 – 782.
- [17] Wu G, Wu T, Chien S. Algorithm and architecture design of perception engine for video coding applications[J]. IEEE Transactions on Multimedia, 2011, 13(6): 1181 – 1194.
- [18] Tang C W. Spatiotemporal visual considerations for video coding[J]. IEEE Transactions on Multimedia, 2007, 9(2): 231 – 238.
- [19] Li Z C, Qin S Y, Itti L. Visual attention guided bit allocation in video compression[J]. Image and Vision Computing, 2011, 29(1): 1 – 14.
- [20] Wei Z, Ngan K N. Spatio-temporal just noticeable distortion profile for gray scale image/video in DCT domain[J]. IEEE

- Transactions on Circuits and Systems for Video Technology, 2009, 19(3): 337 – 346.
- [21] Chen Z Z, Guillemot C. Perceptually-friendly H. 264/AVC video coding based on foveated just-noticeable-distortion model[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2010, 20(6): 806 – 819.
- [22] Wang Z, Bovik A C, Sheikh H, et al. Image quality assessment: from error visibility to structural similarity [J]. IEEE Transactions on Image Processing, 2004, 13(4): 600 – 612.
- [23] Ou T, Huang Y H, Chen H H. SSIM-based perceptual rate control for video coding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2011, 21(5): 682 – 691.
- [24] Wang S Q, Rehman A, Wang Z, et al. SSIM-motivated rate-distortion optimization for video coding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2012, 22(4): 516 – 529.
- [25] Wang S Q, Rehman A, Wang Z, et al. Perceptual video coding based on SSIM-inspired divisive normalization [J]. IEEE Transactions on Image Processing, 2013, 22(4): 1418 – 1429.
- [26] 崔子冠, 朱秀昌. H. 264 结构相似性最优的宏块层码率控制算法 [J]. 电子与信息学报, 2011, 33(6): 1339 – 1344.
- Cui Zi-guan, Zhu Xiu-chang. Structural similarity optimal MB layer rate control algorithm for H.264 [J]. Journal of Electronics and Information Technology, 2011, 33(6): 1339 – 1344. (in Chinese)
- [27] 崔子冠, 朱秀昌. 基于结构相似的 H. 264 主观率失真性能改进机制 [J]. 电子与信息学报, 2012, 34(2): 433 – 439.
- Cui Zi-guan, Zhu Xiu-chang Subjective rate-distortion performance improvement scheme for H. 264 based on SSIM [J]. Journal of Electronics and Information Technology, 2012, 34(2): 433 – 439. (in Chinese)
- [28] Paul M, Lin W S, Lau C, et al. Explore and model better I-frames for video coding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2011, 21(9): 1242 – 1254.
- [29] Zhang J, Yi X Q, Ling N, et al. Context adaptive Lagrange multiplier (CALM) for rate-distortion optimal motion estimation in video coding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2010, 20(6): 820 – 828.
- [30] 袁武, 牛振东, 林守勋, 等. 不依赖于量化参数的视频编码控制优化技术 [J]. 计算机学报, 2011, 34(8): 1519 – 1527.
- Yuan Wu, Niu Zhen-dong, Lin Shou-xun, et al. An adaptable QP-independent approach to rate-distortion optimization for video coding [J]. Chinese Journal of Computers, 2011, 34(8): 1519 – 1527. (in Chinese)
- [31] Ma Z, Xu M, Ou Y, et al. Modeling of rate and perceptual quality of compressed video as functions of frame rate and quantization stepsize and its applications [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2012, 22(5): 671 – 682.
- [32] Cui Z G, Gan Z L, Zhu X C. Joint spatial-temporal quality improvement scheme for H.264 low bit rate video coding via adaptive frameskip [J]. KSII Transactions on Internet and Information Systems, 2012, 6(1): 425 – 444.
- [33] Rhee C, Kim J, Lee H. Bitrate control using a heuristic spatial resolution adjustment for a real-time H. 264/AVC encoder [J]. EURASIP Journal on Advances in Signal Processing, 2012, 87: 1 – 12.
- [34] Kuo C, Chang L, Fan K, et al. Hardware/software codesign of a low-cost rate control scheme for H. 264/AVC [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2010, 20(2): 250 – 261.
- [35] He Z, Cheng W, Chen X. Energy minimization of portable video communication devices based on power-rate-distortion optimization [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2008, 18(5): 596 – 608.
- [36] Su L, Lu Y, Wu F, et al. Complexity-constrained H.264 video encoding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2009, 19(4): 1 – 15.
- [37] Ozcelebi T, Teklap A M, Civanlar M R. Delay-distortion optimization for content-adaptive video streaming [J]. IEEE Transactions on Multimedia, 2007, 9(4): 826 – 836.
- [38] Tiwari M, Groves T, Cosman P C. Delay constrained multiplexing of video streams using dual-frame video coding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2010, 19(4): 1022 – 1035.
- [39] Xu L, Kwong S, Wang H L, et al. A universal rate control scheme for video transcoding [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2012, 22(4): 489 – 501.
- [40] 宋彬, 蒋小兵, 秦浩, 等. 基于 H. 264 具有抗分组丢失能力的帧内刷新算法 [J]. 电子学报, 2009, 37(1): 221 – 224.
- Song Bin, Jiang Xiao-bing, Qin Hao, et al. A packet loss resilient intra refresh algorithm based on H.264 [J]. Acta Electronica Sinica, 2009, 37(1): 221 – 224. (in Chinese)
- [41] Tu W, Steinbach E. Proxy-based reference picture selection for error resilient conversational video in mobile networks [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2009, 19(2): 151 – 164.
- [42] Zhu C, Wang Y, Hannuksela M, et al. Error resilient video coding using redundant pictures [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2009, 19(1): 3 – 14.
- [43] 陈川, 余松煜. 联合编码模式选择的码率控制算法 [J]. 电子学报, 2004, 32(5): 763 – 768.
- Chen Chuan, Yu Song-yu. Joint coding mode selection and rate control algorithm [J]. Acta Electronica Sinica, 2004, 32

- (5): 763 – 768. (in Chinese)
- [44] Zhang Y, Gao W, Lu Y, et al. Joint source-channel rate-distortion optimization for H.264 video coding over error-prone networks[J]. IEEE Transactions on Multimedia, 2007, 9(3): 445 – 454.
- [45] Wang Y, Wu Z, Boyce J M. Modeling of transmission-loss-induced distortion in decoded video[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2006, 16(6): 716 – 732.
- [46] Li Z C, Chakareski J, Niu X, et al. Modeling and analysis of distortion caused by Markov-model burst packet losses in video transmission[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2009, 19(7): 917 – 931.
- [47] Chen Z F, Pahalawatta P V, Tourapis A M, et al. Improved estimation of transmission distortion for error-resilient video coding[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2012, 22(4): 636 – 647.
- [48] Chen Z F, Wu D P. Prediction of transmission distortion for wireless video communication: Analysis[J]. IEEE Transactions on Image Processing, 2012, 21(3): 1123 – 1137.
- [49] 刘家瑛, 郭宗明, Cho Y. 面向 H.264/SVC 空域-质量域可伸缩编码的码率分配算法[J]. 电子学报, 2010, 38(9): 2112 – 2117.
Liu Jia-ying, Guo Zong-ming, Cho Y. Bit allocation algorithm in H.264/SVC spatial-quality with dependent R-D modeling[J]. Acta Electronica Sinica, 2010, 38(9): 2112 – 2117. (in Chinese)
- [50] 李晓峰, 周宁, 刘洪盛, 等. 一种基于缩减栅格算法的 SVC 联合信源/信道编码方法[J]. 电子学报, 2011, 39(4): 859 – 864.
Liu Xiao-feng, Zhou Ning, Liu Hong-sheng, et al. A joint source/channel coding with reduced trellis algorithm for the scalable extension of H.264/AVC[J]. Acta Electronica Sinica, 2011, 39(4): 859 – 864. (in Chinese)
- [51] Kamnoon N, Agrafiotis D, Canagarajah C. Flexible adaptive multiple description coding for video transmission[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2012, 22(1): 1 – 11.
- [52] Tsai W, You H. Multiple description video coding based on hierarchical B pictures using unequal redundancy[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2012, 22(2): 309 – 320.
- [53] Liu Y W, Huang Q M, Ma S W, et al. A novel rate control technique for multiview video plus depth based 3D video coding[J]. IEEE Transactions on Broadcasting, 2011, 57(2): 562 – 571.
- [54] Chang Y, Kim M. A joint rate control scheme in a hybrid stereoscopic video codec system for 3DTV broadcasting[J]. IEEE Transactions on Broadcasting, 2013, 59(2): 265 – 280.
- [55] Zhang Y, Jiang G Y, Yu M, et al. Stereoscopic visual attention-based regional bit allocation optimization for multiview video coding[J]. EURASIP Journal on Advances in Signal Processing, 2010, 60: 1 – 24.
- [56] Zhou Y, Hou C P, Xiang W, et al. Channel distortion modeling for multi-view video transmission over packet-switched networks[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2011, 21(11): 1679 – 1692.
- [57] Lee B, Kim M. Modeling rates and distortions based on a mixture of Laplacian distributions for inter-predicted residues in quadtree coding of HEVC[J]. IEEE Signal Processing Letters, 2011, 18(10): 571 – 574.
- [58] Seo C, Moon J, Han J. Rate control for consistent objective quality in high efficiency video coding[J]. IEEE Transactions on Image Processing, 2013, 22(6): 2442 – 2454.

作者简介



崔子冠 男, 1982 年 1 月出生于河南郑州。2011 年获南京邮电大学工学博士学位, 现为南京邮电大学通信与信息工程学院讲师, 主要研究方向为视频编码与传输。
E-mail: cuizg@njupt.edu.cn



朱秀昌 男, 1947 年 8 月出生于江苏丹徒。教授, 博士生导师, 江苏省图像处理与图像通信重点实验室主任, 主要研究方向为图像处理与多媒体通信。



干宗良 男, 1979 年 9 月出生于新疆石河子。2007 年获南京邮电大学工学博士学位, 现为南京邮电大学通信与信息工程学院讲师, 主要研究方向为图像处理与视频通信。