

考虑视频内容的 H.265/HEVC 帧层码率分配算法

李维¹, 杨付正², 任鹏²

(1. 西安理工大学 印刷包装与数字媒体学院, 陕西 西安 710048; 2. 西安电子科技大学 通信工程学院, 陕西 西安 710071)

摘要: 为了保证一定视频质量下编码器输出的码率符合给定的目标码率, 提出一种考虑视频内容特性的 H.265/HEVC 帧层码率分配算法。首先从率失真理论分析的角度推导出影响输出码率的主要因素, 然后根据视频编码原理, 预测出帧内容复杂度参数, 从而建立一种更有效的帧层码率分配算法。实验结果表明, 所提算法可以使帧层目标码率与编码码率保持更好的一致性, 且重构视频质量平均提高了 0.103 dB。

关键词: 码率分配; 内容特性; 速率控制; H.265/HEVC

中图分类号: TN911.72

文献标识码: A

H.265/HEVC frame-level bitrate allocation algorithm considering video content

LI Wei¹, YANG Fu-zheng², REN Peng²

(1. Faculty of Printing, Packaging and Digital Media Technology, Xian University of Technology, Xi'an 710048, China;

2. Faculty Communication Engineering, Xidian University, Xi'an 710071, China)

Abstract: In order to guarantee the consistency between the coded bits and target bits, a frame-level bitrate allocation algorithm considering the content feature was proposed for H.265/HEVC codec. The main factors to affect the output bitrate were first discussed by the analysis of rate-distortion theory. According to the video coding principle, the related parameter was then predicted to reflect the content complexity. Finally, an efficient frame-level bitrate allocation algorithm was constructed. Experimental results show that the target bits and coded bits keep a better consistency, and the reconstruction video quality is improved by 0.103 dB in average.

Key words: bitrate allocation; content feature; rate control; H.265/HEVC

1 引言

视频业务的广泛应用和用户对其质量要求的不断提高, 促使视频编码技术不断的改进和更新。目前视频编码标准 H.264/AVC^[1] 凭借其高的编码性能已广泛加入到各种视频产品中。然而随着超高清视频业务的逐步兴起和盛行, 其数据量使存储和传输面临严峻的考验, 为此 JCTVC 组织制定下一代视频编码标准 H.265/HEVC^[2]。

H.265/HEVC 编码的最终目标是在给定的带宽内为用户提供最佳的视频质量。为了使编码码率与

带宽保持更好的一致性, H.265/HEVC 引入了码率控制模块。码率控制模块通常被分解为 2 步: 码率分配和量化参数的确定。由于码率分配是码率控制的首要一步, 其准确程度直接影响着后续编码的性能, 因此本文重点研究码率分配算法。

在高效视频编码标准 H.265/HEVC 的制定过程中, JCTVC 组织也逐次接纳了一系列不断改进的码率分配方案, 具有影响力的主要有 JCTVC-H0213^[3] 和 JCTVC-K0103^[4]。JCTVC-H0213 提案是最早被采纳的码率分配算法。与 H.264/AVC 码率分配类似, JCTVC-H0213 码率分配算法利用已有带宽、帧

收稿日期: 2014-10-13; 修回日期: 2015-04-20

基金项目: 国家自然科学基金资助项目(61371089, 61072102); 高等学校学科创新引智计划基金资助项目(B08038); 校科研启动基金资助项目(108-256211407)

Foundation Items: The National Natural Science Foundation of China (61371089, 61072102); The Project of Introducing Talents of Discipline to Universities (B08038); The School Scientific Research Project (108-256211407)

率、编码结构、GOP 大小等相关参数确定各个 GOP、帧、编码树单元的目标比特数。该算法的最终效果并不理想，编码后生成的码率与给定的目标码率差距通常较大，且在一定程度上恶化了视频的观看质量。在第 11 次 JCTVC 会议上，JCTVC-K0103 提案被提出并被加入到 H.265/HEVC 的参考软件 HM 中。与 JCTVC-H0213 不同，JCTVC-K0103 提出一种基于 R-λ模型的码率控制算法，在性能上有较大提高，但存在一定的缺陷。

影响码率分配性能的主要因素有给定的目标码率和视频自身的内容特性。目前码率分配算法都已考虑目标码率这一因素的影响。因此本文通过考虑视频内容特性，建立一种新的 H.265/HEVC 帧层码率分配方案，来提高码率控制的性能。

2 H.265/HEVC 码率分配

已知当前待编码的 GOP 中各帧的平均目标比特数 T_A 的计算方法为

$$T_A = R_p + \frac{R_p N_c - R_c}{S} \quad (1)$$

其中， R_p 为一帧的平均目标比特数， N_c 为已编码帧的总个数， R_c 为已编码帧的总比特数， S 为滑动窗口大小，取值为 40。则当前待编码的 GOP 的目标比特数 T_G 为

$$T_G = T_A N_G \quad (2)$$

其中， N_G 为 GOP 中所含帧个数。已知当前待编码 GOP 的目标比特数后，对其所含的图像进行逐帧编码。当前待编码帧的目标比特数 T_C 为

$$T_C = \frac{T_G - R_c}{\sum_{N_n} \omega_i} \quad (3)$$

其中， R_c 表示当前正在编码的 GOP 生成的编码比特数； N_n 为当前正在编码的 GOP 中剩余未编码图像的个数， ω 表示图像的码率在整个 GOP 码率中所占的比重，其获取方法如下。

假设当前待编码的 GOP 中各帧所使用的拉格朗日因子 λ 有

$$\lambda_0 : \lambda_1 : \dots : \lambda_{N_G} = \eta_0 : \eta_1 : \dots : \eta_{N_G} \quad (4)$$

各个帧的 η 值在随机接入 (RA) 配置下为

$$\begin{aligned} \eta_0 &= 1.0 \\ \eta_1 &= 0.725 \log \lambda' + 0.7963 \\ \eta_i &= m_i \eta_1, \quad i = 2, \dots, 7 \end{aligned} \quad (5)$$

其中， m 为恒定值，参数 λ' 为

$$\lambda' = 0.5 \lambda' + 0.5 \lambda_i \quad (6)$$

其中， λ_i 表示前一个已编码 GOP 中的关键帧所采用的拉格朗日因子。

在得到各帧采用的拉格朗日因子比值后，利用 R-λ模型计算出当前待编码的 GOP 中各帧的目标比特数 R_i

$$\left. \begin{aligned} \lambda_i &= \alpha_i b_i^{\beta_i} = \alpha_i \left(\frac{R_i}{dh} \right)^{\beta_i} \\ \sum_{N_G} R_i &= T_G \end{aligned} \right\} \Rightarrow R_i \quad (7)$$

其中， b 为一个像素的目标比特数， h 和 d 分别为图像的长和宽， α_i 和 β_i 为参数。最终 w_i 为

$$w_i = \frac{R_i}{\sum_{N_G} R_j} \quad (8)$$

权值 w 用于反映待编码 GOP 中各帧的码率分配关系。

H.265/HEVC 帧层码率分配算法虽然对编码性能有所贡献，但其仍存在缺陷。从上述描述可以看出，帧层码率分配没有考虑当前待编码图像的相关信息，其涉及的所有变量都是通过利用已编码图像信息预测得到的，忽略了当前待编码图像的内容特性。图 1 给出了 Racehorse 序列下各个帧对应的实际编码码率与目标码率之间的关系。测试使用 H.265/HEVC 的 RA Main10 配置，参考软件选用 HM11.0。可以看出，帧与帧之间内容的变化，使码率分配算法的参数无法及时更新，造成帧层输出的比特数很难与目标比特数保持一致，两者之间比特误差较大。因此在 H.265/HEVC 帧层比特分配上不应忽视当前待编码帧的内容特性。

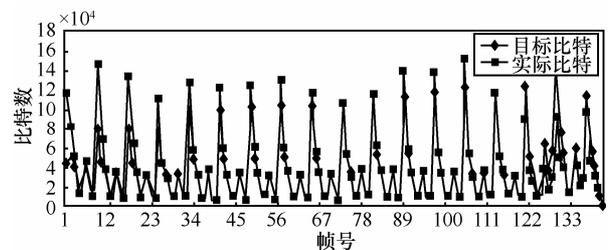


图 1 帧层目标比特与实际比特之间的关系

3 考虑内容特性的帧层码率分配算法

按照当前 GOP 的目标比特数 T_G ，各个帧分配的目标比特数可被描述为

$$\begin{aligned} \min_{T_1, T_2, \dots, T_N} & D_1 + D_2 \dots + D_N \\ \text{s.t.} & R_1 + R_2 \dots + R_N = T_G \end{aligned} \quad (9)$$

即当前 GOP 的总编码比特数满足 T_G 的前提下, 找出一组最优帧层目标码率集 T_1, T_2, \dots, T_N , 使 GOP 编码后的总失真最小。其中, R_i 和 D_i 分别为第 i 帧的编码比特数和编码失真。式(9)的解可以通过拉格朗日优化方法^[5]来获取

$$\min J \quad J(T_1, T_2 \dots T_N) = \sum_{i=1}^N D_i + \lambda \sum_{i=1}^N R_i \quad (10)$$

由于帧间预测技术的使用, 造成帧与帧之间具有一定的依赖特性, 因此要得到帧层最优的目标码率需考虑其率失真性能的相关性。

3.1 码率分配算法理论研究

H.265/HEVC 采用先进的帧间预测技术, 使参考帧的失真会沿着编码方向向后续帧传播, 从而导致各个帧之间的失真具有一定的依赖关系。因此, 第 i 帧的编码失真 D_i 不仅与其生成的比特数 R_i 相关, 还与已编码帧的失真有关, 即

$$D_i = F_i(D_{i-1}, D_{i-2}, \dots, D_1, R_i) \quad (11)$$

其中, 符号 $F_i()$ 表示 D_i 、 $(D_{i-1}, D_{i-2}, \dots, D_1)$ 、 R_i 三者之间的函数关系。从而可以逐次得出

$$D_1 = F_1(R_1) \quad (12)$$

$$D_2 = F_2(D_1, R_2) = F_2(F_1(R_1), R_2) \quad (13)$$

$$D_i = F_i(F_1(R_1), F_2(F_1(R_1), R_2), \dots, R_i) \quad (14)$$

已有研究^[6]表明当前帧失真与其参考帧失真之间存在一定的线性关系。为了验证该结论是否适于 H.265/HEVC, 本节测试 H.265/HEVC 下当前帧失真与它的参考帧失真之间的关系, 如图 2 所示。实验所选用的配置为 RA Main10, 失真则采用均方误差 (MSE) 进行度量。从图 2 可以看出帧与帧之间的失真是以线性的方式传播的。因此有

$$\begin{aligned} D_i &= u_i D_{i-1} + c_i \\ &= u_i u_{i-1} D_{i-2} + u_i c_{i-1} + c_i \\ &\vdots \\ &= u_i u_{i-1} \dots u_2 D_1 + u_i c_{i-1} + u_i u_{i-1} c_{i-2} + \dots + u_i u_{i-1} \dots u_3 c_2 + c_i \end{aligned} \quad (15)$$

参数 u_i 反映了第 $i-1$ 帧的编码失真对第 i 帧编码失真的影响程度, u_i 和 c_i 的取值与 R_i 有关。

因此, 对式(10)中代价函数进行求得

$$\begin{aligned} \frac{\partial J}{\partial R_i} &= \frac{\partial D_i}{\partial R_i} + \frac{\partial D_{i+1}}{\partial R_i} \dots + \frac{\partial D_N}{\partial R_i} + \lambda \\ &= \frac{\partial D_i}{\partial R_i} + \frac{\partial(u_{i+1} D_i + c_{i+1})}{\partial R_i} \dots + \lambda \\ &= (1 + u_{i+1} \dots + u_N u_{N-1} \dots u_{i+1}) \frac{\partial D_i}{\partial R_i} + \lambda \\ &= k_i \frac{\partial D_i}{\partial R_i} + \lambda \end{aligned} \quad (16)$$

则对于第 i 帧, 最优的目标码率就等价于最小化 J_i

$$J_i = D_i + \frac{\lambda}{k_i} R_i = D_i + \lambda_i R_i \quad (17)$$

一般地, 视频编码器的 RD 性能可被描述为

$$D_i = \delta_i^2 e^{-\theta_i R_i} \quad (18)$$

其中, δ_i^2 表示预测残差的方差, β_i 为模型参数。结合式(17), 可以推导出

$$R_i = \frac{-\log\left(\frac{\lambda_i}{\delta_i^2 \theta_i}\right)}{\theta_i} \quad (19)$$

可以看出, 当前帧的目标比特数与视频内容特性有关, 纹理平坦和纹理复杂这 2 类区域所分配的目标比特数是有所差异的。因此本文建立一种考虑内容特性的帧层码率分配方案, 来提高码率控制的精确性。

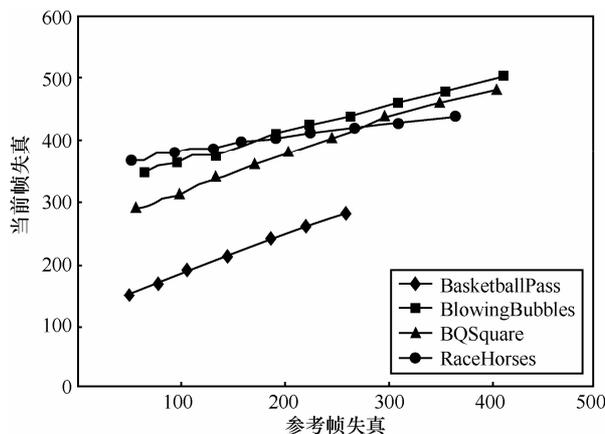


图 2 参考帧编码失真与当前帧编码失真之间的关系

3.2 新的帧层码率分配算法

从 3.1 节可知, 为了有效地建立 H.265/HEVC 帧层码率分配算法, 准确衡量视频帧的内容特性成为其重要的一步。由于在编码之前无法得到当前帧确切的统计特性, 因此对帧内容特性的估量具有一

定的挑战性。考虑到方差度量是视频编码中应用广泛的一种方法, 因此本文首先利用已编码帧和待编码帧的已有信息, 估计出待编码帧的方差, 然后建立 H.265/HEVC 帧层码率分配算法。

假设 $u(i, j)$, $u'(i, j)$, $v(i, j)$ 分别表示位于 (i, j) 处参考帧的原始像素值、参考帧的重构像素值、当前待编码帧的像素值。以原始参考帧 u 作为参考, 经过运动补偿后的残差 $e(i, j)$ 为

$$e(i, j) = v(i, j) - u(d(i, j)) \quad (20)$$

$d(i, j)$ 为进行运动估计时当前编码帧中位于 (i, j) 处像素的位移偏移量。则经运动补偿后的残差方差 δ^2 为

$$\delta^2 = E[e^2(i, j)] = E[\{v(i, j) - u(d(i, j))\}^2] \quad (21)$$

但在实际的编码中, 由于原始参考帧无法在解码端重现, 因此运动估计和补偿都是基于已解码的参考帧 u' 进行的。实际的残差 $e'(i, j)$ 为

$$e'(i, j) = v(i, j) - u'(d(i, j)) \quad (22)$$

则方差 δ'^2 为

$$\delta'^2 = E[e'^2(i, j)] = E[\{v(i, j) - u'(d(i, j))\}^2] \quad (23)$$

式(22)可改写为

$$e'(i, j) = \{v(i, j) - u(d(i, j))\} + \{u(d(i, j)) - u'(d(i, j))\} \quad (24)$$

其中, 前半部分表示以原始参考帧为参考, 当前帧经过运动补偿后得到的残差, 后半部分表示参考帧中经过位移偏移后的原始像素值与已解码像素值之间的误差。假设这 2 部分互不相关, 则

$$\delta'^2 = \delta^2 + E[\{u(d(i, j)) - u'(d(i, j))\}^2] \quad (25)$$

式(25)中的第 2 部分虽然不是参考帧的编码失真, 但却与其具有很强的相关性, 因此假设

$$\begin{aligned} & E[\{u(d(i, j)) - u'(d(i, j))\}^2] \\ & \approx \tau E[\{u(i, j) - u'(i, j)\}^2] \\ & = \tau D_u \end{aligned} \quad (26)$$

D_u 为参考帧的编码失真。经实验测试 τ 取值约为 0.9。结合式(25)和式(26)可得

$$\delta'^2 = \delta^2 + \tau D_u \quad (27)$$

以上分析是针对单向预测时, 当前待编码图像的方差预测方法。当前待编码图像采用双向预测时, 则

$$\delta'^2 = a(\delta_{i1}^2 + \tau D_{u1}) + b(\delta_{i2}^2 + \tau D_{u2}) \quad (28)$$

其中, D_{u1} 为参考帧 1 的编码失真, D_{u2} 为参考帧 2 的编码失真。参数 a 和 b 分别表示两参考帧被当前待编码图像所参考的比例, 该值可以通过前面已编码的同层帧预测得到。为了验证方差估计的准确度, 图 3 给出了实际方差与预测方差之间的关系, 可以看出该方法能够准确地预测待编码帧的方差。实验配置为 RA Main10, 量化参数为 27。

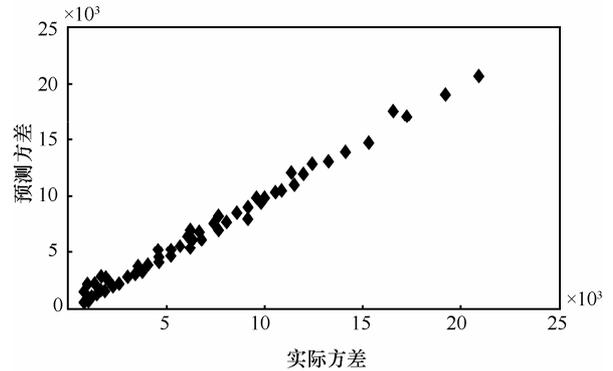


图3 BasketballPass 序列的实际方差与预测方差

在预测出帧内容复杂度后, 本文提出将方差 δ_i^2 加入到原始的 H.265/HEVC 码率分配中, 来建立新的帧层码率分配算法。令

$$\begin{aligned} \eta_0 &= 1.0 \\ \eta_1 &= 0.725 \log\left(\lambda \frac{\delta_0^2}{\delta_{r0}^2}\right) + 0.7963 + \log\left(\rho \frac{\delta_0^2}{\delta_{r1}^2}\right) \\ \eta_i &= m_i \eta_1 + \log\left(\rho \frac{\delta_i^2}{\delta_{ri}^2}\right), \quad i = 2, \dots, 7 \end{aligned} \quad (29)$$

其中, δ_i^2 表示前一个 GOP 中第 i 帧的实际残差方差, 经实验测试 ρ 值取 2.718。将式(29)代入式(4), 结合式(8)计算出 w_i 值, 最终通过式(3)得出当前待编码帧的目标码率。

4 实验结果

将本文所提的帧层码率分配方法加入到 H.265/HEVC 编码器的参考软件 HM11.0 中, 以验证该算法的性能。测试选用不同内容特性的视频序列: BasketballPass、BlowingBubbles、BasketballDrill、BQmall、Fourpeople、Johnny、ParkScene、BQTerrace, 分辨率有 416×240、832×480、1 280×720、1 920×1 080, 每一种分辨率对应 2 个测试序列, 其编码帧数为 200。实验所采用编码配置为 RA Main10^[7]。

为了对码率分配这一模块进行专门测试, 本文对量化参数的确定这一模块进行修改, 以屏蔽其对码率分配最终性能的影响。具体做法是通过遍历所有的量化参数, 最后选择率失真代价最小的量化参数作为码率控制所使用的量化参数。

本文首先从准确度上来分析所提算法的性能。准确度衡量方法为

$$R_e = \frac{|R_T - R_A|}{R_T} \quad (30)$$

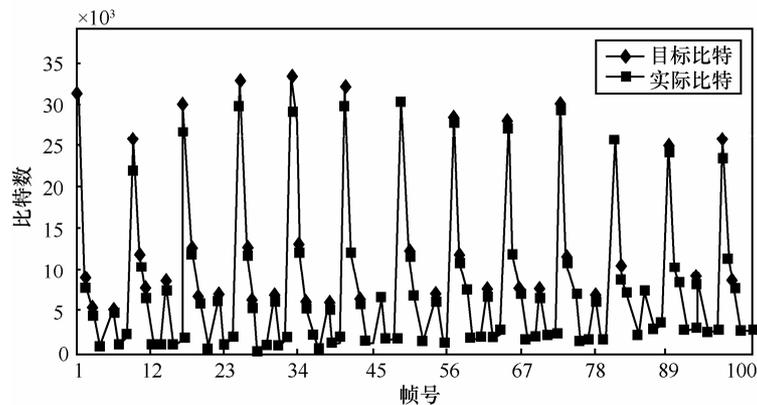
其中, R_e 为相对误差, 反映了实际编码码率 R_A 相对于目标码率 R_T 的偏离程度。表 1 分别给出了 H.265/HEVC 帧层码率分配算法和本文所提算法的相对误差 R_e 。可以看出, 与 H.265/HEVC 帧层码率分配算法相比, 本文所提算法的准确度较高, 帧层的实际编码码率与目标码率具有好的一致性。同时图 4 给出了 H.265/HEVC 帧层码率分配算法和本文所提算法分别计算出来的实际编码码率与目标码率之间的关系。可以看出, 本文所提算法能够根据时域视频内容的变化, 合理地调整帧层的目标

码率, 使实际编码码率满足目标码率。

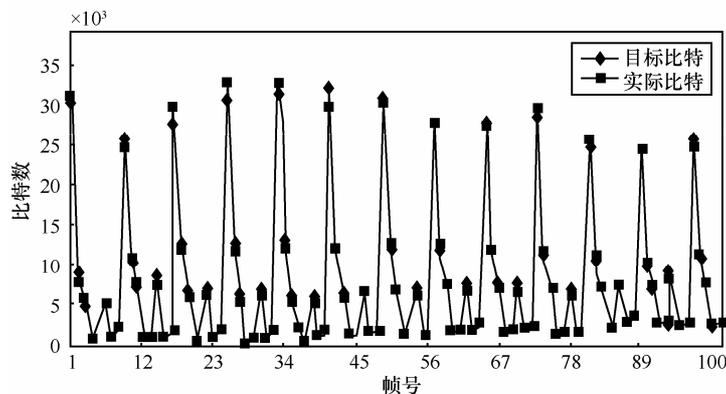
表 1 码率分配算法的准确度比较

序列	H.265/HEVC 帧层码率分配算法的 R_e	本文所提的帧层码率分配算法 R_e
BasketballPass	0.598	0.150
BlowingBubbles	0.185	0.041
BasketballDrill	0.183	0.086
BQmall	0.318	0.193
Fourpeople	0.468	0.139
Johnny	0.106	0.060
ParkScene	0.056	0.026
BQTerrace	0.193	0.049
平均	0.263	0.093

在编码效率方面, 本文使用 $\Delta Bit-rate$ 和 $\Delta PSNR$ 作为评价指标^[8]。 $\Delta Bit-rate$ 用于反映编码输出码率的变化情况, 其值为负, 表示 $Bit-rate$ 减少; $\Delta PSNR$ 用于反映视频质量的变化情况, 其值为正, 表示 $PSNR$ 增加。表 2 给出了各个序列下本文所提算法与 H.265/HEVC 帧层码率分配算法的对比结果。可以得出本文所提算法的 $Bit-rate$ 平均降低了约



(a) H.265/HEVC 帧层码率分配算法



(b) 本文所提的码率分配算法

图 4 BasketballPass 序列下各个帧对应的目标码率与编码码率

3.13%，PSNR 平均提高了约 0.103 dB。此外图 5 给出了一视频序列的 RD 性能曲线，以进一步说明实验对比结果。可以看出本文所提算法的 RD 曲线位于 H.265/HEVC 帧层码率分配算法的 RD 曲线之上，说明在编码码率相同的情况下，本文所提算法能够提供更好的视频质量。图 6 还给出了 H.265/HEVC 帧层码率分配算法和本文所提算法得到的各个帧的 PSNR 值，可以看出本文所提算法得到的各个帧的 PSNR 大多都高于 H.265/HEVC 帧层码率分配算法得到的各个帧的 PSNR。因此本文所提算法的性能优于 H.265/HEVC 帧层码率控制算法。

表 2 所提算法与 H.265/HEVC 帧层码率分配算法对比结果

序列	所提的帧层码率分配算法	
	$\Delta\text{Bit-rate}/\%$	$\Delta\text{PSNR}/\text{dB}$
BasketballPass	-2.71	+0.133
BlowingBubbles	-2.04	+0.087
BasketballDrill	-2.51	+0.108
BQmall	-2.93	+0.125
Fourpeople	-4.77	+0.089
Johnny	-3.54	+0.114
ParkScene	-2.85	+0.083
BQTerrace	-3.72	+0.086
平均	-3.13	+0.103

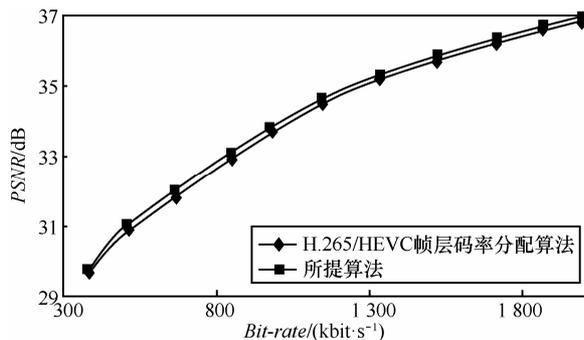


图 5 BQmall 序列的率失真曲线

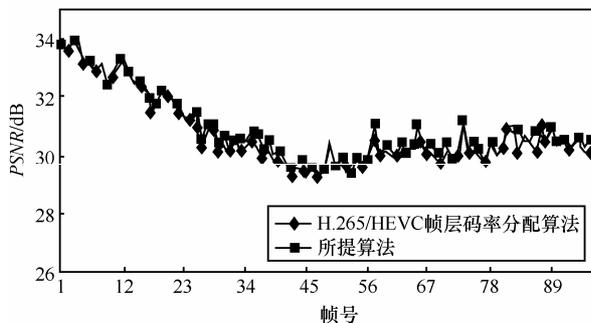


图 6 BQmall 序列各个帧的 PSNR 值

5 结束语

在视频编码中，不同内容特性的视频生成的编

码比特数就有所差异。为了使帧层编码码率满足给定的目标码率，码率分配算法应该根据不同帧内容来合理地分配目标码率，以提高算法的准确性。本文提出了一种新的帧层码率分配算法，通过估计待编码帧的内容复杂度来为各个帧分配合理的目标码率。实验结果表明，本文所提的帧层码率分配算法具有较高的编码性能，能够很好地保持编码码率与目标码率的一致性。

参考文献：

- [1] WIEGAND T, SULLIVAN G J, LUTHRA A. Overview of the H.264/AVC video coding standard [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13(7): 560-576.
- [2] SULLIVAN G J, OHM J R, HAN W J, et al. Overview of the high efficiency video coding[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2012, 22(12): 1649-1668.
- [3] CHOI H, NAM J, YOO J. Rate control based on unified RQ model for HEVC[A]. Proceedings of the 8th JCTVC Meeting[C]. San Jose, 2012.
- [4] LI B, LI H Q, LI L, et al. Rate control by R-lambda model for HEVC[A]. Proceedings of the 11th JCTVC Meeting[C]. 2012.
- [5] ORTEGA A, RAMCHANDRAN K. Rate-distortion methods for image and video compression [J]. IEEE Signal Processing Magazine, 1998, 15(6): 23-50.
- [6] LIU J, CHO Y, GUO Z. Bit allocation for spatial scalability coding of H.264/SVC with dependent rate-distortion analysis[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2010, 20(7): 967-981.
- [7] BOSSEN F. Common test conditions and software reference configurations[A]. Proceedings of 11th JCT-VC Meeting[C]. Shanghai, 2012.
- [8] HJONTEGAARD G. Calculation of average PSNR differences between RD-curves [EB/OL]. http://wftp3.itu.int/av-arch/video-site/0104_Aus/VCEG-M33.doc.

作者简介：



李维（1985-），女，陕西咸阳人，西安理工大学讲师，主要研究方向为视频编码、多媒体通信。

杨付正（1977-），男，山东德州人，西安电子科技大学教授、博士生导师，主要研究方向为视频编码、视频质量评估、多媒体通信。

任鹏（1984-），男，河南新乡人，西安电子科技大学博士后，主要研究方向为数字信号处理、模式识别等。