

基于 SATD 的 H.265/HEVC 拉格朗日因子选择算法

李维¹, 赵凡¹, 张二虎¹, 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: 为了进一步提高帧内粗略模式判决的编码性能, 提出了一种基于 SATD 的 H.265/HEVC 拉格朗日因子选择算法。首先, 分析了哈达玛变换的特性, 从率失真理论的角度入手, 以 SATD 为失真测量尺度, 建立了一种新的拉格朗日因子计算方法。其次, 结合离散余弦变换原理, 推导出所提方法中涉及的参数获取途径。实验结果表明, 所提算法可以更准确地进行帧内预测模式判决, 在同等码率的情况下可以有效地提高重构视频质量。

关键词: 帧内预测模式判决; 拉格朗日因子; 率失真优化; H.265/HEVC

中图分类号: TN911.7

文献标识码: A

Lagrange multiplier selection based on SATD for H.265/HEVC

LI Wei¹, ZHAO Fan¹, ZHANG Er-hu¹, REN Peng²

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

2. School of Telecommunications Engineering, Xidian University, Xi'an 710071, China)

Abstract: In order to improve the coding performance of intra rough mode decision, a SATD-based Lagrange multiplier selection algorithm was proposed for H.265/HEVC. The characteristic of Hadamard transform was first analyzed, and following the new Lagrange multiplier calculation approach based on the distortion measurement of SATD was constructed according to the rate-distortion theory. Then combing the principle of discrete cosine transform, the related parameters were obtained by the formula derivation. Experimental results show that the proposed method supports more accurately intra prediction mode decision, and the reconstruction video quality is improved in average while keeping the same bit-rate.

Key words: intra prediction mode decision, Lagrange multiplier, rate-distortion optimization, H.265/HEVC

1 引言

网络技术和电子技术的不断演进, 促使了多媒体业务的快速发展, 如电视广播、视频会议、移动 TV、医疗检测、远程教育和虚拟现实等。随着越来越多用户的参与, 这些业务逐步活跃在人们生活的各个角落。视频已成为不可或缺的载体, 且用户对视频质量的要求也不断提升。为了在有限信道内进行传输和存储, 庞大的视频信息需先经过视频编码器进行压缩。目前, 视频编码标准 H.264/AVC^[1] 凭借其较高的编码性能已广泛渗透到各种媒体产品中。然而由于具有更好的用户体验, 超高清视频业

务逐步走入人们视野, 但冗余的数据量使存储和传输面临严峻的考验。因此, 视频编码联合小组 (JCTVC) 制定了下一代视频编码标准 H.265/HEVC^[2], 以进一步提高视频压缩效率。

H.265/HEVC 依然沿用 H.264/AVC 中的混合编码框架, 如帧内预测和基于运动补偿的帧间预测、变换量化、熵编码、环路滤波等。但同时也引入了多种新的编码技术, 如基于四叉树的自适应分割技术、多角度帧内预测技术、先进的运动矢量预测技术、模式依赖的正弦离散变换技术、自适应系数扫描技术、像素自适应补偿技术和内部位深扩展等。其中, 率失真优化技术仍起着举足轻重的作用。每

收稿日期: 2016-09-06; 修回日期: 2017-04-12

基金项目: 中国博士后科学基金资助项目 (No.2017M610644); 国家自然科学基金资助项目 (No.61671376, No.61671374)

Foundation Items: China Post Doctoral Science Foundation (No.2017M610644), The National Natural Science Foundation of China (No.61671376, No.61671374)

一种候选编码模式都需经过预测、变换、量化和熵编码过程，计算出其对应的码率和失真。通过率失真代价函数确定最优的编码模式，以保证最终的编码码率满足给定的信道带宽。

关于率失真优化问题，通常采用拉格朗日优化技术来求解^[3]。而在拉格朗日优化中，一个关键的问题就是如何选择拉格朗日因子，其直接控制着码率和失真之间的权衡问题。Wiegand 等^[4]提出一种有效的拉格朗日因子计算方法，即在量化参数与拉格朗日因子之间建立一种指数关系，该方法已广泛被业界所采纳且成功地应用到 H.265/HEVC 中。Sun 等^[5]考虑了用户感知特性，提出了一种改进的上下文自适应的拉格朗日因子选择方法。基于 ρ 域的率失真模型，Chen 等^[6]建立了一种内容自适应的拉格朗日因子选择方法。此外，考虑到帧内帧间技术造成编码失真不断向后续待编码单元传播，一种新的自适应拉格朗日因子计算方法^[7]被提出且用于码率控制中。随后 Li 等^[8]定量地研究了当前编码单元对后续编码性能的影响，提出了一种新的拉格朗日优化模型。上述方法都是在编码失真度量尺度为均方误差和 (SSE) 的前提下获取拉格朗日因子的，其不同程度地提高了视频压缩效率，保证所选取的最优编码模式能够有效满足信道带宽的要求。

在 H.265/HEVC 帧内预测模式选择中，首先要采用拉格朗日优化方法进行粗略模式判决 (RMD)。该过程使用的拉格朗日因子是通过基于 SSE 获取的拉格朗日因子进行开方得到的，这种方法简单高效，但仍存在一定的缺陷。在 RMD 中，编码失真的度量尺度是经哈达玛变换后的绝对误差和 (SATD)，即编码失真是在频域中量度的，而编码失真 SSE 是在空域中量度的，因此，拉格朗日因子的选择不应该通过简单的关系式得到。

因此，本文提出了一种新的拉格朗日因子选择算法，用于 H.265/HEVC 的 RMD 中，来进一步提高帧内编码效率。基于哈达玛变换原理，本文分析了预测残差的分布特性，然后依据率失真理论建立了基于 SATD 的率失真模型，继而推导出拉格朗日因子选择算法。同时，结合离散余弦变换 (DCT) 原理，估计出算法中涉及的参数。实验结果表明，本文所提算法可以有效地获取 RMD 中所用的拉格朗日因子，以进一步提高帧内编码效率，在同等码率的条件下，重构视频质量有一定的提高。

2 H.265/HEVC 帧内拉格朗日因子选择算法

拉格朗日优化是视频编码中一种有效的工具，其根据率失真代价 J 来选择最优的编码模式，使编码器达到最优的率失真性能。

$$\min\{J\}, J = D + \lambda R \quad (1)$$

其中， λ 为拉格朗日因子， R 和 D 分别表示编码码率和编码失真。

在帧内编码中，H.265/HEVC 支持 35 种帧内预测模式，包括 33 种角度模式、DC 模式和 Planar 模式，如表 1 所示。同时，图 1 给出了角度模式的具体预测方向。其中，V 表示垂直角度，H 表示水平角度。因此，利用拉格朗日优化技术，从 35 种帧内预测模式中选择出一种最优的预测模式来最大化地降低帧内冗余信息。但如对 35 种帧内预测模式逐一进行变换、量化、熵编码、反量化、反变换、像素重构来获取准确的编码码率和编码失真，编码器将要承受很大的计算负荷及处理时延，这对视频业务的传输是非常不利的。所以，在帧内预测模式判决中，H.265/HEVC 分 2 步实现：RMD 和 RDO (rate-distortion optimization)。

表 1 帧内预测模式及其对应的模式名

帧内预测模式	相应的模式名
0	Planar
1	DC
2, ..., 34	角度

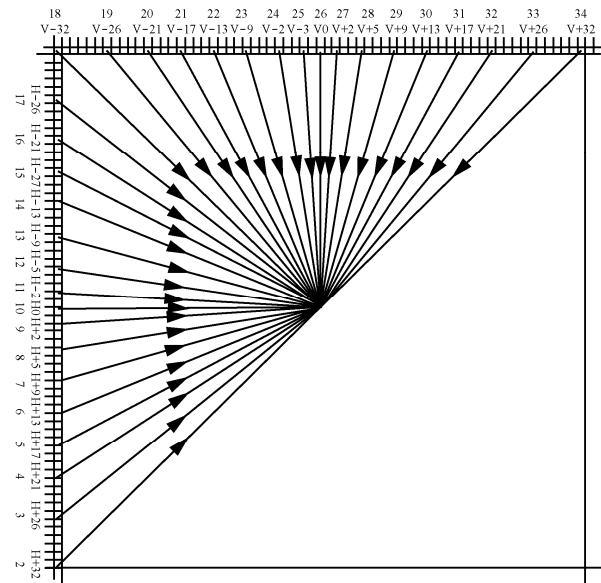


图 1 角度预测模式的方向

首先, 进行粗略模式判决, 即计算这 35 种帧内预测模式的基于 SATD 的率失真代价为

$$J_{SATD} = SATD + \lambda_{pre} R_{pre} \quad (2)$$

其中, λ_{pre} 表示基于 SATD 的拉格朗日因子, R_{pre} 是指编码帧内预测模式所需的比特数。根据式(2)结果从中选择出 N 种率失真代价 J_{SATD} 最小的帧内预测模式作为最可能模式 (MPM)。 N 的取值依赖预测单元(PU)大小, 当 PU 大小分别取 $\{4 \times 4, 8 \times 8, 16 \times 16, 32 \times 32, 64 \times 64\}$ (像素) 时, N 值分别为 $\{8, 8, 3, 3, 3\}$ 。

其次, 对这 N 种 MPM 进行 RDO 评价, 选择基于 SSE 的率失真代价 J_{SSE} 最小的预测模式作为该 PU 的最优帧内预测模式。基于 SSE 的率失真代价为

$$J_{SSE} = SSE + \lambda_{mode} R_{mode} \quad (3)$$

其中, λ_{mode} 为基于 SSE 的拉格朗日因子, R_{mode} 指帧内预测模式和预测残差的总编码比特数。

假设在高码率条件下, 熵受限标量量化的编码器的率失真模型^[9]可表示为

$$R(D) = \frac{1}{2} \text{lb} \left(\frac{\delta^2}{D} \right) \quad (4)$$

其中, δ^2 表示 DCT 域残差的方差。结合式(4), 可推导出拉格朗日因子

$$\lambda = -\frac{dD}{dR} = 2 \ln 10 \delta^2 10^{-2R} \quad (5)$$

在均方误差的条件下, 失真模型可用量化参数 Q 表示, 则最终可得

$$\lambda = cQ^2 \quad (6)$$

其中, c 为常量。可见拉格朗日因子与量化参数有关, 因此, 在 H.265/HEVC 帧内模式判决中, λ_{mode} 和 λ_{pre} 定义^[10]为

$$\begin{cases} \lambda_{mode} = a \cdot 2^{\frac{Q-12}{3}} \\ \lambda_{pre} = \sqrt{\lambda_{mode}} \end{cases} \quad (7)$$

其中, a 为常量。从式(7)可以看出, 基于 SATD 的拉格朗日因子 λ_{pre} 与基于 SSE 的拉格朗日因子 λ_{mode} 之间呈简单的指数关系。由于 SATD 与 SSE 分别在频域和空域得到, 因此, 基于 SATD 的拉格朗日因子 λ_{pre} 的获取方法具有一定的局限性。

3 新的帧内拉格朗日因子选择算法

在 H.265/HEVC 中, 对 35 种帧内预测模式逐

一进行哈达玛变换

$$T = HBH^T \quad (8)$$

其中, H 表示哈达玛变换矩阵, B 和 T 分别指空域的预测残差块和变换域残差块。一般地, 当 PU 大小是 4×4 时, 采用哈达玛变换矩阵 $H_{4 \times 4}$; 当 PU 大小是 8×8 时, 采用哈达玛变换矩阵 $H_{8 \times 8}$; 当 PU 大小是 16×16 、 32×32 和 64×64 时, 先进行 8×8 大小的二叉树分割, 然后采用哈达玛变换矩阵 $H_{8 \times 8}$ 对分割块进行逐一变换。之后, 按照式(9)计算 SATD。

$$SATD = \sum_{i=1}^L \sum_{j=1}^L |t_{ij}| \quad (9)$$

其中, 变量 t_{ij} 表示残差块经过二维哈达玛变换后的块内系数, 参数 L 表示块内水平或垂直方向上所含系数个数。

哈达玛变换具有能量集中的特性, 主要将能量集中在二维矩阵的左上角。哈达玛变换核只包括 1 和 -1 这 2 个元素, 各行或各列之间彼此正交。图 2 分别给出了 4×4 和 8×8 大小的哈达玛变换矩阵核。由于哈达玛变换的运算速度快, 对存储空间要求较少, 且其特性非常接近 DCT, 因此, 被应用于 H.265/HEVC 的残差变换中。

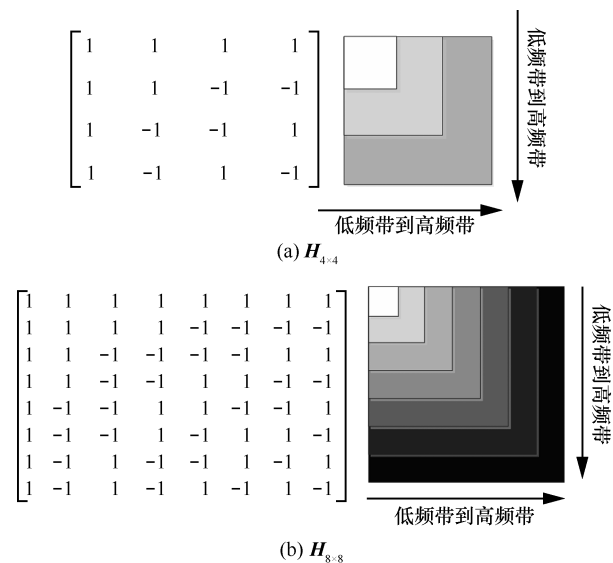


图 2 哈达玛变换矩阵及对应的频带

3.1 拉格朗日因子获取

在高码率假设下, DCT 残差分布通常被描述为零均值的高斯分布。由于哈达玛变换与 DCT 具有类似的特性, 因此, 哈达玛变换残差也可以被描述为高斯分布。

$$f(x) = (2\pi\delta_h^2)^{\frac{1}{2}} e^{-\frac{x^2}{2\delta_h^2}} \quad (10)$$

其中, x 表示哈达玛变换后的残差, δ_h 为哈达玛变换后残差的标准差。在失真测度为 SATD 的情况下, 率失真模型为

$$R(D) = \frac{1}{2} \text{lb} \left(\frac{\pi\delta_h^2}{2eD} \right) \quad (11)$$

则可推导出拉格朗日因子为

$$\lambda' = -\frac{dD}{dR} = \sqrt{\frac{\pi}{2e}} \ln 10 \delta_h 10^{-R} \quad (12)$$

将式(5)代入式(12), 可得

$$\begin{aligned} \lambda' &= -\frac{dD}{dR} = \sqrt{\frac{\pi}{2e}} \ln 10 \delta_h 10^{-R} \\ &= \sqrt{\frac{\pi}{4e \ln 10}} \ln 10 \frac{\delta_h}{\delta} \sqrt{\lambda} \\ &= c' \frac{\delta_h}{\delta} \sqrt{\lambda} \end{aligned} \quad (13)$$

其中, $c' = \sqrt{\frac{\pi}{4e \ln 10}} \ln 10$, 根据式(13), 本文最终提出基于 SATD 的拉格朗日因子 λ'_{pre} 。

$$\lambda'_{\text{pre}} = \frac{\delta_h}{\delta} \sqrt{\lambda_{\text{mode}}} \quad (14)$$

若已知 DCT 域标准差 δ 以及哈达玛变换域标准差 δ_h , 就可以利用式(14)计算出 RMD 中采用的拉格朗日因子。

3.2 参数估计

对于大小为 $M \times M$ (像素) 的残差块, 系数 $r(x, y)$ 经过 DCT 变换后的变换系数为

$$Y(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{M-1} r(x, y) A(x, u) A(y, v) \quad (15)$$

其中, $A(m, n) = C(k) \cos \left[\frac{(2m+1)n\pi}{8} \right]$, 参数 $C(k)$

为已知量, 则

$$\begin{aligned} \delta^2 &\approx \frac{1}{MM} \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \delta^2(u, v) \\ &= \frac{1}{MM} \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} \delta_p^2 [AZA^T]_{u,u} [AZA^T]_{v,v} \\ &= \frac{\delta_p^2}{MM} \sum_{i=0}^{M-1} \sum_{j=0}^{M-1} [AZA^T]_{u,u} [AZA^T]_{v,v} \end{aligned} \quad (16)$$

其中, $[\cdot]_{u,u}$ 表示矩阵中位于 (u, u) 位置处的系数值。符号 Z 定义为

$$Z = \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^M \\ \rho & 1 & \rho & \dots & \rho^{M-1} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ \rho^M & \dots & \dots & \dots & 1 \end{bmatrix} \quad (17)$$

其中, 参数 ρ 用于度量水平方向和垂直方向上像素值之间的相关性, 其值设置为 0.6^[11]。参数 δ_p 用于表示残差块内像素的标准差, 该值可通过平均绝对差值 (MAD) 近似^[11]

$$\delta_p \approx \sqrt{2} MAD \quad (18)$$

因此, 根据式(17)和式(19)就可以获取到 DCT 域标准差 δ 。关于哈达玛变换域的标准差 δ_h 的计算如下所示。

由 3.1 节可知, 哈达玛变换域的预测残差服从均值为 0, 方差为 δ_h^2 的高斯分布, 则根据期望值 E 的定义, 可得

$$E[|x|] = \int_{-\infty}^{+\infty} |x| (2\pi\delta_h^2)^{\frac{1}{2}} e^{-\frac{x^2}{2\delta_h^2}} dx = \sqrt{\frac{2}{\pi}} \delta_h \quad (19)$$

结合式(9), 同时可以得到

$$E[|x|] = \frac{SATD}{L \cdot L} \quad (20)$$

因此, 根据式(19)和式(20), 可得哈达玛变换域的标准差

$$\delta_h = \sqrt{\frac{\pi}{2}} \frac{SATD}{L \cdot L} \quad (21)$$

综上所述为参数 δ 以及 δ_h 的预测方法, 代入式(14)即可计算出基于 SATD 的拉格朗日因子, 总流程如图 3 所示。

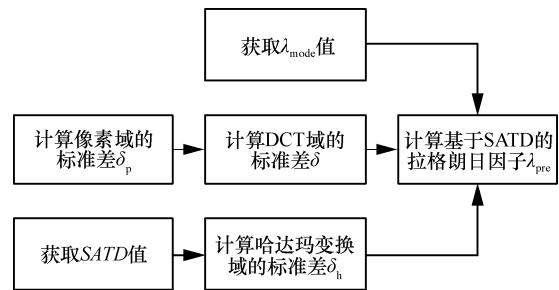


图 3 所提算法流程

4 实验结果

为了验证算法的性能, 本文将所提的基于 SATD 的拉格朗日因子选择算法分别加入到 H.265/

HEVC 原有算法^[2]和 Li 等^[8]所提算法中。本文实验的硬件平台为 Windows7，处理器为 Intel Core i5-4590，CPU 为 3.30 GHz，内存为 4.00 GB。测试软件平台为 HM16.0^[12]，在 Intra_only Main 配置下对标准视频序列进行测试。标准序列选取在不同分辨率和不同应用场景下，JCTVC 提供的 Class B/C/D/E 类序列，其各自对应的分辨率分别为 1 920×1 080、832×480、416×240、1 280×720。所选用的量化参数分别为 22、27、32、37。本文从编码效率和编码复杂度 2 个方面来考察新建算法的总体性能。

4.1 编码效率

本文采用 $\Delta bit-rate$ 和 $\Delta PSNR$ 作为评价指标^[13]来验证所提算法的有效性。表 2 给出了各类序列下本文所提的基于 SATD 的拉格朗日因子选择算法与 H.265/HEVC 的 RMD 中原有的拉格朗日因子选择算法的对比。可以看出本文算法的 PSNR 平均提高了 0.116 dB，比特率平均降低了 2.62%。图 4 给出了不同序列下的率失真性能曲线。可以看出，本文算法的率失真曲线位于 H.265/HEVC 原有算法的率失真曲线之上。图 5 给出了所提算法与 H.265/HEVC 分别编码同一帧时所得到的帧内预测模式和重构帧。可以看出位于椭圆处的场景，在采用本文算法后得到了更准确的帧内预测模式，其重构效果优于 H.265/HEVC 原有算法所得到的重构效果。综上所述，在给定码率的情况下，本文算法能够达到更好的 PSNR 值。

表 2 本文算法与 H.265/HEVC 原有算法编码性能对比

序列类型	$\Delta bit-rate$	$\Delta PSNR/dB$
Class B	-2.55%	0.115
Class C	-2.38%	0.103
Class D	-2.64%	0.124
Class E	-2.93%	0.125
平均	-2.62%	0.116

本文是在率失真模型的推导下得出的拉格朗日因子，其中的拉格朗日优化是在总的编码码率下研究的。而在实际的 RMD 中，忽略了残差编码比特数，只采用了编码模式比特数进行拉格朗日优化，因此，编码性能只提升了 0.116 dB。为了进一步说明实验对比结果，表 3 给出了各类序列下本文

算法与文献[8]所提算法的对比结果。从表 3 中可以看出，本文算法的 PSNR 平均提高了 0.157 dB，比特率平均降低了 2.96%，其结果进一步说明本文算法具有较高的编码效率。

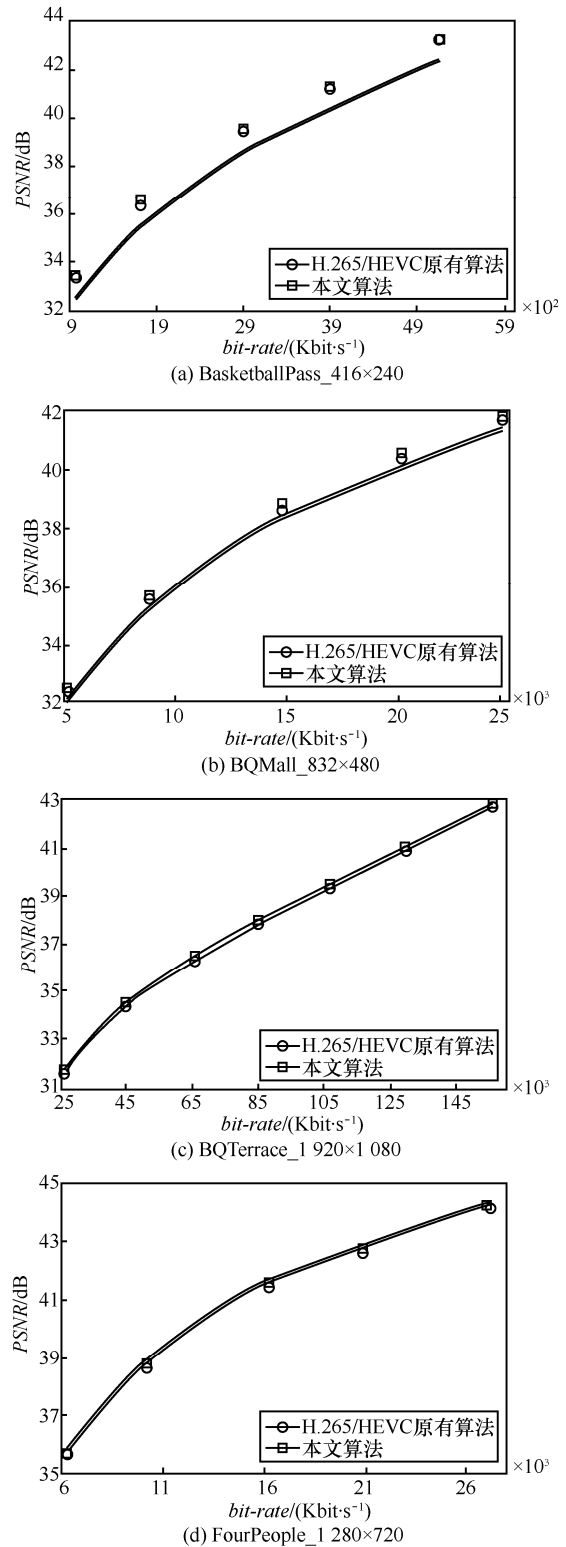


图 4 所提算法与 H.265/HEVC 原有算法的率失真曲线

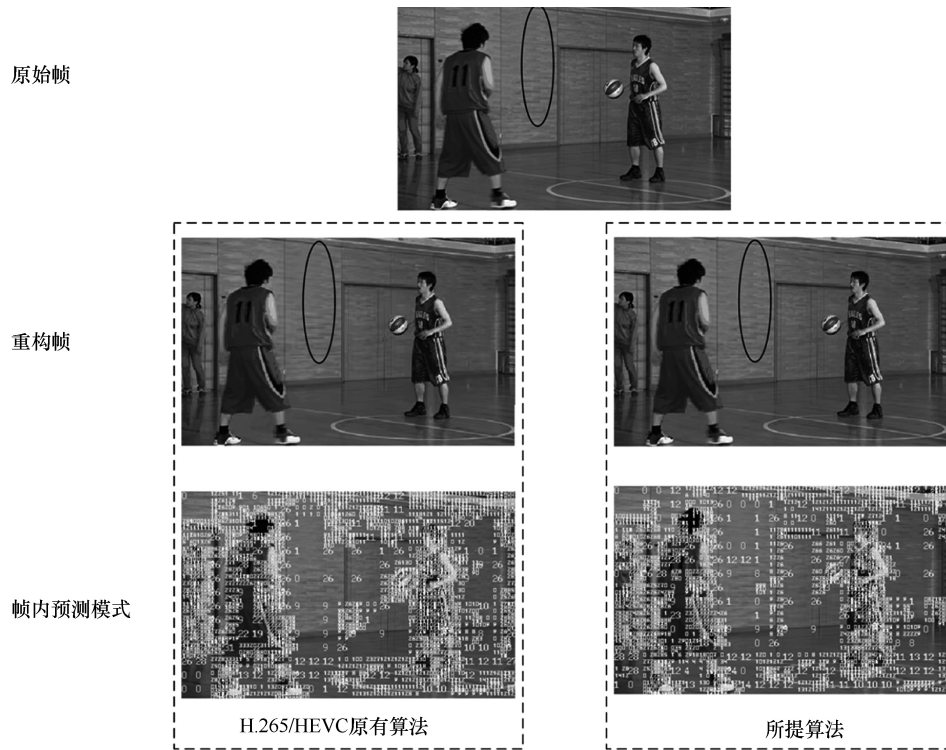


图 5 所提算法与 H.265/HEVC 原有算法效果对比

表 3 本文算法与文献[8]算法的编码性能对比

序列类型	$\Delta bit-rate$	$\Delta PSNR/dB$
Class B	-2.72%	0.158
Class C	-2.94%	0.133
Class D	-2.83%	0.146
Class E	-3.36%	0.192
平均	-2.96%	0.157

表 4 本文算法与 H.265/HEVC 原有算法编码复杂度对比

序列类型	ΔT
Class B	0.82%
Class C	0.65%
Class D	0.33%
Class E	0.48%
平均	0.57%

4.2 编码复杂度

在编码复杂度方面，本文使用编码时间差 ΔT 进行度量，其定义为

$$\Delta T = \frac{T_p - T_o}{T_o} \times 100\% \quad (22)$$

其中， T_p 表示本文计算方法所耗的编码时间， T_o 表示参考算法所耗的编码时间。表 4 给出了本文算法与 H.265/HEVC 原有算法的对比。从实验数据可以看出，与 H.265/HEVC 原有算法相比，本文算法对编码复杂度没有太大影响，其编码时间平均提高了 0.57%。同时，表 5 给出了本文算法与文献[8]算法的对比，本文算法编码时间平均提高了 0.59%。可见，本文所提的基于 SATD 的拉格朗日因子选择算法不仅提高了视频压缩的效率，同时也未给编码器带来大的计算负担。

表 5 本文算法与文献[8]算法编码复杂度对比

序列类型	ΔT
Class B	0.84%
Class C	0.69%
Class D	0.41%
Class E	0.42%
平均	0.59%

5 结束语

在视频编码中，拉格朗日优化技术起着重要的作用，拉格朗日因子直接影响着最优模型选择的结果。本文针对 H.265/HEVC 的帧内粗略模式判决，提出一种基于 SATD 的拉格朗日因子选择算法。通过利用率失真理论，建立出拉格朗日因子计算方法。其次，结合 DCT 变换原理，推导出

相关参数估计方法。实验结果表明, 本文算法具有较高的帧内编码性能, 在相同码率下可以有效地提高视频质量。

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] ORTEGA A, RAMCHANDRAN K. Rate-distortion methods for image and video compression[J]. IEEE Signal Processing Magazine, 1998, 15(6): 23-50.
- [4] WIEGAND T, GIROD B. Lagrange multiplier selection in hybrid video coder control[C]//The IEEE International Conference on Image Processing. Thessaloniki, 2001.
- [5] SUN C, WANG H, KIM T, et al. Perceptually adaptive Lagrange multiplier for rate-distortion optimization in H.264[C]//The Future Generation Communication and Networking. Jeju, 2007.
- [6] CHEN L, GARBACEA I. Adaptive lambda estimation in Lagrangian rate-distortion optimization for video coding[C]//The Visual Communications, Image Processing. San Jose, 2006.
- [7] 李维, 杨付正, 任鹏. 考虑视频内容的 H.265/HEVC 帧层码率分配算法[J]. 通信学报, 2015, 36(9): 76-81.
LI W, YANG F Z, REN P. H.265/HEVC frame-level bitrate allocation algorithm considering video content[J]. Journal on Communications, 2015, 36(9): 76-81.
- [8] LI S, ZHU C, GAO Y. Lagrangian multiplier adaptation for rate-distortion optimization with inter-frame dependency[J]. IEEE Trans on CSVT, 2016, 26(1): 117-129.
- [9] GISH H, PIERCE J. Asymptotically efficient quantizing[J]. IEEE Trans Inf Theory, 1986, 14(5): 676-683.
- [10] KIM I K. High efficiency video coding test mode (HM15) encoder description[S]. JCTVC, 2014.
- [11] PAO I, SUN M. Modeling DCT coefficients for fast video encoding[J]. IEEE Trans Circuits Syst, Video Technol, 1999, 9(4): 608-616.
- [12] BOSSEN F. Common test conditions and software reference configurations[S]. JCTVC, 2013.
- [13] HJONTEGAARD G. Calculation of average PSNR differences between RD-curves[C]//VCEG, 2001.

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。

赵凡 (1969-), 女, 陕西富平人, 西安理工大学副教授, 主要研究方向为多媒体信号分析、压缩与通信。

张二虎 (1965-), 男, 山西永济人, 西安理工大学教授, 主要研究方向为图像分析与机器视觉。

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