知识库编码中知识库图像的自适应量化参数决策

# 讨论列表

* 2017/12/8—
* ~~2017/12/5—讨论1）基于率失真优化的知识库图像量化参数选择问题的公式化（4.2.1）是否合理；2）码率估计模型（4.2.2）是否合理；3）失真估计模型（4.2.3）是否合理。~~
* ~~2017/11/23—由于需要修改的内容较多，仅完成了文档结构的修改，还没有完成理论分析部分的修改，大概还需要1天修改。~~
* ~~2017/11/22—主要讨论4.6节OPT-DQP建模中两部分内容：(提纲)~~
  + ~~首先是讨论像素块使用intra模式编码的码率估计方法（p13）是否可行：该方法根据像素块的像素值方差以及像素的方向性强弱估计码率的方法；~~
  + ~~然后是讨论根据训练样本拟合OPT-DQP模型的结果（p16）如何改进：训练模型和样本的一致性R^2=0.5,模型不能很好地描述样本特性。考虑有以下需要改进的方法：~~

# 摘要

**研究目标：**

本文主要研究在知识库编码中，如何自适应地调整每个知识库图像的量化参数，使得序列的编码效率最优。

**研究方法：**

为此，本文分析知识库图像量化参数与序列编码效率之间的关系模型，将量化参数的选择问题转化为基于率失真的最优化问题并求解以选择最优量化参数。

**技术指标：**

本文提出的知识库图像的自适应量化参数选择方法，将最优化测试序列化编码效率，并在最差情况下不降低序列的编码效率，预期在HM测试序列上获得平均10%的编码效率增益。

# 研究背景

需要调研的几个方面：

* 知识库编码现有的方法和问题
  + 目前知识库编码为每个知识库图像分配固定QP，该QP与序列的编码QP的差值为固定值。固定量化参数
* 已有的内容自适应的QP分配方法
* 为了有效估计QP对编码效率的影响，需要调研码率控制或码率分配等问题中对基于率失真的编码效率的估计模型。

## 码率失真估计模型：

* **R-D模型：**

所有的码率失真估计模型都是基于一定的信源假设下的R-D模型，下面列举几个主要的R-D模型。

根据文献【quantizeTIT68gish】【rateCtrlTCSVT97thChiang】【RDOTIP09xli】，在高码率环境以及信源符合laplcacian分布的假设下，rate和distortion之间满足关系：

根据文献【RDModelTSP98Mallat】【adapQpICIP09xli】【adapQpISCAS10xli】，在高斯信源假设下，rate和distortion之间满足关系：

其中是变换系数的方差。

根据文献【RDModelTSP98Mallat】，在低码率（rate<1bit/pixel）情况下，rate-distortion满足关系：

在大多数情况下b=1。

根据文献【RDModelTCSVT05kamaci】，变换系数符合cauchy分布，rate和distortion之间的关系满足：

* **R估计模型**

**R-Q模型：**

在的关系下，研究者提出R-Q模型：

根据文献【rateCtrlTCSVT97thChiang】【rateCtrl99TCSVTvetro】**【**rateCtrlTIP16tszhao**】**【modeSelectJSTSP13tszhao】【rateCtrlTCSVT08hlwang】【rateCtrlICASSP07hlwang】，R-Q模型可以描述为quadric-Q模型：

根据文献【rateCtrlTCSVT05swma】【rateCtrlTCSVT11sdhu】【rateCtrlTIP11sdhu】【rateCtrlTCSVT07yliu】【rateCtrlISCAS07yliu】【rateCtrlTCSVT08yliu】【rateCtrlMMSP07dkkwon】【rateCtrlICIP11jyliu】【rateCtrlTCSVT07dkkwon】【rateCtrlSPIC14yli】【rateCtrlTCSVT09jpdong】，R-Q模型可以表述为linear-Q模型：

除了【rateCtrlMMSP07dkkwon】【rateCtrlICIP11jyliu】【rateCtrlTCSVT07dkkwon】【rateCtrlSPIC14yli】，所有文献都使用MAD（motion-compensated absolute difference）度量预测残差，但是由于MAD需要在编码之后得到，因此文献都使用已编码图像的预测残差估计当前编码图像的预测残差，**这样的问题是**需要进行编码才能获得预测残差，且预测残差并不能直接反映参考图像对当前编码图像的影响，不适用于知识库编码。而【rateCtrlMMSP07dkkwon】【rateCtrlICIP11jyliu】【rateCtrlSPIC14yli】则根据原始像素的帧差度量预测残差，虽然可以避免编码获得预测残差，**但是问题是**：基于原始像素的帧差并不准确（intra模式使用平均值作为预测值，inter则使用参考图像作为预测值），当编码图像和参考图像的差异较大时，帧差并不能反映预测残差，不适用于R-Q模型，导致R的估计不准确。·【rateCtrlTCSVT07dkkwon】则使用变换后的预测残差来度量，尽管更准确，**却导致了**在真正编码之前需要对图像进行预分析的编码，且由于要根据RDcost最优化获得图像的最佳QP，预分析编码需要在多个QP下遍历，大大增加了复杂度，考虑到这个方法用于块级的QP调整，在知识库编码中，帧级QP调整应该不需要这么精确，知识库编码中可以不使用变换预测残差，而使用其他文献中的方法估计预测残差。

根据文献【RDModelTCSVT05kamaci】【rateCtrlICASSP09yjcho】【rateCtrlTCSVT08jyliu】，R-Q模型为幂函数模型：

根据文献【rateCtrlTCSVT13sswang】，R-Q模型的形式取决于预测残差失真的分布，对于H.264，块大小固定为16x16，预测残差失真的分布服从单一laplacian分布，但是对于HEVC，由于块划分的存在，预测残差失真需要由多个laplacian分布综合（每一层块大小对应一个laplacian分布），因此，R-Q和D-Q模型更新为：

其中是预测残差失真的laplacian分布。

**R-λ模型：**

在的前提下，研究者提出R-λ模型：

根据文献【rateCtrlTIP13bli】【rateCtrlTMM15lili】【rateCtrlISCAS15lili】【rateCtrlTCSVT15ycgong】【rateCtrlTIP16mhwang】，编码比特数与λ之间存在指数关系：

**Header-R模型：**

根据文献【rateCtrlTCSVT07dkkwon】，编码图像的header bits并不是固定不变的，为此建立header bits model：

其中是MV中非零分量（水平、垂直）的个数，是MV的个数。

* **D估计模型**

**D-Q模型：**

**在什么R-D的前提下？？？**

根据文献【quantizeTIT68gish】，当使用均匀量化时，D-Q模型为：

根据文献【rateCtrlTCSVT08hlwang】【rateCtrlICASSP07hlwang】【rateCtrlTIP16tszhao】**【**rateCtrlTCSVT11sdhu**】**，D-Q模型为线性模型：

【rateCtrlICASSP07hlwang】声明线性D-Q模型用于P/B帧。I帧的失真可能需要修正

根据文献【RDModelTCSVT05kamaci】【rateCtrlICASSP09yjcho】【rateCtrlTCSVT08jyliu】，D-Q模型为幂函数模型：

**D-λ模型：**

在的前提下，研究者提出R-λ模型：

根据文献【rateCtrlTIP13bli】【rateCtrlTMM15lili】【rateCtrlISCAS15lili】【rateCtrlTCSVT15ycgong】【rateCtrlTIP16mhwang】，编码比特数与λ之间存在指数关系：

由于，该模型可以转换为D-Q模型中的幂函数模型。

## 码率失真传递模型：

**失真传递模型：**

根据文献【rateCtrlTCSVT11sdhu】【rateCtrlTIP16tszhao】，失真传递模型为：

其中。另外，码率之间并没有相关性。而文献【modeSelectJSTSP13tszhao】在失真传递的基础上建立了RDcost传递模型为：

其中表示第l个temporal layer的RDcost变化量。该模型的应用范围是描述H.264的HBP结构中编码图像的RDcost与其最邻近参考图像的RDcost的变化量之间的关系。

根据文献【adapQpISCAS10xli】，参考图像和编码图像之间的失真传递关系可以表示为：

其中表示编码图像中skip的块在图像中的面积占比，表示参考图像的失真，表示编码图像非skip块的平均失真。

根据文献【rateCtrlTCSVT08jyliu】，在SVC编码的spatial scalable应用中，基本层（参考图像）和增强层（编码图像）之间的distortion存在依赖关系：

其中。

根据文献【rateCtrlICASSP09yjcho】【rateCtrlTCSVT13yjcho】，在SVC编码的temporal scalable应用中，基本层（参考图像）和增强层（编码图像）之间的distortion存在依赖关系：

其中。

**码率传递模型：**

另外根据文献【adapQpISCAS10xli】，低层图像和高层图像之间的**平均码率**传递关系表示为：

其中表示hierarchical temporal layer中的第k层平均bits，表示第k层与第0层的QP之差，a是一个常数，文章中使用定值2.5。**但是这里并没有考虑**不同序列中图像之间的差异性，a对于不同序列应该是自适应的，需要引入衡量层间图像之间的差异性因素进行调整；同时，平均码率的依赖关系并不能反映关键帧和知识库图像之间的码率依赖关系。

根据文献【rateCtrlTCSVT08jyliu】，在SVC编码的spatial scalable应用中，基本层（参考图像）和增强层（编码图像）之间的rate存在依赖关系：

其中。

根据文献【rateCtrlICASSP09yjcho】【rateCtrlTCSVT13yjcho】，在SVC编码的temporal scalable应用中，基本层（参考图像）和增强层（编码图像）之间的rate不存在依赖关系：

其中。

**预测残差传递模型：**

根据文献【adapQpICIP09xli】，在hierarchical编码结构中，低层图像和高层图像的预测残差能量之间有线性关系：

定义为低层图像和高层图像之间的差错传递因子，通过统计之前编码图像的预测残差能量和QP差的比值来计算。但是该因子是建立在低层图像和高层图像之间本身就很高的相似性上，并不普适，因此需要对于不同序列都需要实时计算。同时，该公式的推导包含极大的简化和不合理的假设，关系描述并不准确。

根据文献【rateCtrlTIP11sdhu】，当前编码图像和编码顺序上前一图像的预测残差之间呈比例关系：

其中是编码图像的预测残差。同时文章经过实验发现服从laplacian分布，且对于不同序列，分布的具体参数不同。但是**这个模型的问题是**：在hierarchical参考结构中，编码图像并不一定参考前一帧，关联性并不一定存在，同时，预测残差的比值作为相似性，似乎不是那么合理。

而文献【rateCtrlJVCIR05zgli】【rateCtrlTCSVT07yliu】【rateCtrlISCAS07yliu】【rateCtrlTCSVT08yliu】则认为和之间服从一个更普适的线性关系。其中，【rateCtrlTCSVT07yliu】【rateCtrlISCAS07yliu】【rateCtrlTCSVT08yliu】认为时域依赖关系并不能准确估计预测残差，引入一个空域依赖关系，通过使用较少的inter和intra模式编码得到当前编码图像的粗略的预测残差，并以此作为空域信息预测当前编码图像的实际预测残差。文章通过对比前一帧图像使用时域或空域的判决，决定当前编码图像使用哪种依赖关系。**这种解决方法的问题**是需要对当前图像进行two-pass编码，而且并不能确定时域依赖或空域依赖的有效性，其实之所以要引入上述空域和时域依赖关系，是因为当前图像和前一个编码图像之间的内容相似性并没有被考虑进来，导致不能有效综合空域和时域的依赖关系，因此考虑将图像之间的相关性引入以综合空域和时域的依赖性，而不是分别度量。

根据文献【rateCtrlSPIC14yli】，编码图像和参考图像的预测残差与编码图像和原始参考图像的原始预测残差之间存在线性关系：

类似的，文献【ErrorPropagateRDOICME12twyang】【dependentRDOICME15sli】【dependentRDOTCSVT16sli】【adapQpICIP17ymzhou】认为预测残差能量之间存在线性关系：

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ref | R-model |  |  |  | D-model |  |  |  | RD-dependency |
|  | Q-domain | |  | λ-domain | Q-domain | |  | λ-domain |
|  | quadric | first-order\linear | power |  | squared | linear | power |  |  |
| quantizeTIT68gish |  |  |  |  | √ |  |  |  |  |
| rateCtrlTCSVT97thChiang | basicTheory |  |  |  |  |  |  |  |  |
| rateCtrl99TCSVTvetro | √ |  |  |  |  |  |  |  |  |
| rateCtrlTCSVT08hlwang | √ |  |  |  |  | √ |  |  |  |
| rateCtrlICASSP07hlwang | √ |  |  |  |  | √ |  |  |  |
| rateCtrlTIP16tszhao | √ |  |  |  |  | √ |  |  | √ |
| modeSelectJSTSP13tszhao | √ |  |  |  | × | × | × |  |  |
| rateCtrlJVCIR05zgli | √ |  |  |  | × |  |  |  |  |
| RDModelTCSVT05kamaci |  |  | √ |  |  |  | √ |  |  |
| rateCtrlTCSVT08jyliu | |  | √ |  |  |  | √ |  |  |
| 【rateCtrlICASSP09yjcho】 | | | √ |  |  |  | √ |  |  |
| rateCtrlTCSVT13yjcho |  |  | √ |  |  |  | √ |  |  |
| rateCtrlTCSVT05swma |  | √ |  |  |  | √ |  |  |  |
| rateCtrlTCSVT11sdhu |  | √ |  |  |  | √ |  |  | √ |
| rateCtrlTIP11sdhu |  | √ |  |  | × | × | × |  |  |
| rateCtrlTCSVT07yliu |  | √ |  |  | √ |  |  |  |  |
| rateCtrlISCAS07yliu |  | √ |  |  | √ |  |  |  |  |
| rateCtrlTCSVT08yliu |  | √ |  |  | × |  |  |  |  |
| rateCtrlMMSP07dkkwon |  | frame-diff |  |  | × |  |  |  |  |
| rateCtrlICIP11jyliu |  | frame-diff |  |  | × |  |  |  |  |
| rateCtrlTCSVT07dkkwon | | SATD |  |  |  | √ |  |  |  |
| rateCtrlTCSVT09jpdong |  | √ |  |  |  |  |  |  |  |
| rateCtrlSPIC14yli | | frame-diff |  |  |  |  |  |  |  |
| rateCtrlJVCIR05zgli |  | √ |  |  |  |  |  |  | √ |
| rateCtrlTIP13bli |  |  |  | √ |  |  |  | √ |  |
| rateCtrlTMM15lili |  |  |  | √ |  |  |  | √ |  |
| rateCtrlISCAS15lili |  |  |  | √ |  |  |  | √ |  |
| rateCtrlTCSVT15ycgong |  |  |  | √ |  |  |  | √ |  |
| rateCtrlTIP16mhwang |  |  |  | √ |  |  |  | √ |  |
|  |  |  |  |  |  |  |  |  |  |
| rateCtrlTIP94ramchandran |  |  |  |  |  |  |  |  | √ |
| adapQpICIP09xli |  |  |  |  |  |  |  |  | √ |
| adapQpISCAS10xli |  |  |  |  |  |  |  |  | √ |
| rateCtrlTCSVT98ljlin |  |  |  |  |  |  |  |  | √ |
| rateCtrlTCSVT08jyliu |  |  |  |  |  |  |  |  | √ |
| rateModelISCAS15jhhou |  |  |  |  |  |  |  |  | √ |
| qualityCtrlTBC13jhhou |  |  |  |  |  |  |  |  | √ |
| rateCtrlICASSP09yjcho |  |  |  |  |  |  |  |  | √ |
| rateCtrlTCSVT13yjcho |  |  |  |  |  |  |  |  | √ |
| rateCtrlTCSVT13sswang |  |  |  |  |  |  |  |  | √ |
| rateCtrlSPIC14yli |  |  |  |  |  |  |  |  | √ |
| adapQpICIP17ymzhou |  |  |  |  |  |  |  |  | √ |
| rateCtrlICIP02sliu |  |  |  |  |  |  |  |  | √ |

# 知识库图像的自适应量化参数决策模型

知识库图像量化参数选择的目标是使序列的编码效率最优，所以量化参数选择的问题可以转化为最优化问题并求解，表示为：

其中编码效率受到知识库图像量化参数以及其他编码因素和内容因素的影响。为了求解上述问题，首先需要确定影响的和，然后确定与各影响因素之间的关系式，最后再解决上述最优化问题。

## 影响知识库编码效率的编码因素和内容因素分析

下面列举了可能会影响知识库编码效率的因素并定性地分析这些因素对编码效率的影响程度：

* 编码因素：在知识库编码中，非知识库图像的编码因素并没有进行适应的修改，仍然与传统编码保持一致，因此影响知识库编码效率的主要编码因素包括，
  + 知识库图像的量化参数
    - 知识库图像量化参数决定知识库图像的重建质量。量化参数越高，知识库图像重建质量越低，能够为关键帧提供的参考信息就越少，编码效率越低。
  + 关键帧的量化参数
    - 关键帧的量化参数决定序列的基准编码码率和失真，虽然不会直接影响序列的编码效率，但是会影响知识库图像码率在序列码率中的比重。量化参数越高，给定量化参数编码的知识库图像码率的比重越大，知识库图像提供的信息的使用比率就越小，编码效率越低。
* 内容因素
  + 知识库图像带来的因素
    - 知识库图像的个数
      * 序列可以按照参考的知识库图像被分为多个序列片段，一个片段中的关键帧参考同一个知识库图像。由于关键帧仅参考一个知识库图像且知识库图像相互独立，因此片段之间相互独立，知识库图像个数的变化并不会影响编码效率。
    - 知识库图像的内容复杂程度
      * 内容复杂程度反映知识库图像携带参考信息的代价，内容越复杂，知识库图像编码的码率越高，编码效率越低。
    - 知识库图像的重建质量
      * 重建质量反映知识库图像在编码重建之后保存的有效参考信息的多少，重建质量越高，关键帧能获取的参考信息也越多，编码效率越高。
      * 重建质量由量化参数决定，量化参数越大，重建质量越小。
  + 关键帧带来的因素
    - 参考每一个知识库图像的关键帧个数
      * 关键帧个数反映知识库图像信息被重用的程度，参考知识库图像的关键帧个数越多，关键帧从知识库图像中获取的总信息量越多，编码效率越高。
    - 关键帧与其参考的知识库图像之间的相似程度
      * 相似程度反映关键帧从知识库图像中获取的信息量多少，相似性越高，关键帧从知识库图像中获取的信息量越多，编码效率越高。
  + 普通图像（非关键帧非知识库图像）带来的因素
    - 普通图像与关键帧的相似程度
      * 普通图像并不能直接从知识库图像中获取参考信息，但可以从关键帧中间接获取知识库图像传递的信息，普通图像与关键帧的相似程度越高，知识库图像能够传递的信息也越多，编码效率越高。

## 构建编码效率与知识库图像量化参数和其他影响因素之间的关系

### 以率失真代价度量编码效率，并将知识库图像量化参数选择的问题转化为率失真代价的最优化问题

由于量化参数QP是指向量化步长Q的索引值，Q才是直接影响量化结果的参数，为了便于表达和理解，下文中将QP转化为Q。知识库图像量化步长的约束最优化问题表示为：

其中为给定总码率下使序列图像失真总和最小的知识库图像量化步长集合，每个知识库图像的量化步长取值范围为，且。分别表示序列中非知识库图像和知识库图像的数目，表示第i个非知识库图像在受到量化参数为的知识库图像影响下的失真和码率，是第j个知识库图像的码率。之所以没有考虑知识库图像的失真，是因为在知识库编码中，用于显示的图像仅包括非知识库图像而不包括知识库图像。

其中为给定总码率下使序列图像失真总和最小的知识库图像量化步长集合，每个知识库图像的量化步长取值范围为，且。分别表示序列中非知识库图像和知识库图像的数目，表示第i个非知识库图像在受到量化参数为的知识库图像影响下的失真和码率，是第j个知识库图像的码率。之所以没有考虑知识库图像的失真，是因为在知识库编码中，用于显示的图像仅包括非知识库图像而不包括知识库图像。

上述限定条件的最优化问题可以通过使用lagrange multiplier将限定条件与优化主体结合，得到：

在现有的知识库编码中，每个关键帧仅参考一个知识库图像，并且知识库图像之间相互独立，因此可以将序列分割为子序列，每个子序列仅依赖一个知识库图像，那么序列的RDcost可以转化为子序列的RDcost之和：

其中表示依赖第j个知识库图像的子序列中非知识库图像的个数。由于子序列之间相互独立，序列RDcost最优的问题可以转化为每个子序列的RDcost最优的问题：

为了求解子序列的基于率失真代价最优的量化步长选择问题，需要确定失真和码率受量化步长影响的估计模型。在以往的研究中，研究者在不同的适用条件和应用前提下提出了多种失真和码率的估计模型，本文根据知识库编码的特性，继承并调整这些模型，并验证模型的可用性，从而对知识库编码中的失真和码率进行较准确的估计。

### 建立失真受量化步长影响的估计模型

根据知识库编码的RDcost，序列的失真仅包含非知识库图像的失真。以往的文献给出了很多的失真模型，虽然这些模型拥有不同的形式，但是失真模型的推导都源自于相似的假设前提和推导方法，为了准确描述失真模型，并建立适合于知识库编码的失真模型，本文分析失真的推导过程。

#### 预测残差变换量化失真的理论计算

失真来自于预测残差的变换系数的量化误差，根据文献，预测残差经DCT变换后的系数符合zero-mean Laplacian distribution，其概率密度函数（probability density function）表示为：，

其中是变换系数的幅值，是变换系数的标准差，由于变换系数的均值为0，也是变换系数的能量总和，同时，由于变换域信号的能量和像素域信号的能量一致，因此方差可以表示为：

根据变换系数的分布，假设使用量化步长量化变换系数，那么变换系数量化后的失真可以表示为：

其中是均匀量化器的倍率因子（？？？，可能是死区的扩大倍数），是量化重建值相较于量化起始区间的偏差，且。令，则可以剥离因子使得积分项与内容无关得到：

其中的值域与变化趋势与无关，因此对于不同内容的序列均适用，通过积分可以得到：

综上所述，变换系数的失真（即编码图像的失真）可以表示为：

那么估计失真的关键问题就转变为如何估计预测残差的能量。

#### 非知识库图像的预测残差能量的计算

* 关键帧

对于参考知识库图像的关键帧，由于关键帧与知识库图像之间的相似性并不是很高，关键帧中有较高比率的像素块使用帧内预测模式，因此，关键帧的预测像素有两种来源，一种是帧内预测得到的预测像素，一种是帧间预测得到的预测像素，两种预测像素得到的预测残差的分布并不相同，因此预测残差能量可以拆分为：

其中，分别表示关键帧中使用帧内预测和帧间预测的像素个数。

**（1）帧内预测模式的预测残差能量估计**

对于使用帧内预测的像素，预测像素值由像素块的方向模式和边缘像素块组合得到，但是方向模式需要经过实际编码才能得到，为了不在知识库图像的预分析中引入过多复杂度，我们需要使用合理的方法估计使用帧内预测模式的预测残差能量。

文献【rateCtrlMMSP07dkkwon】【rateCtrlICIP11jyliu】使用像素块的均值作为预测像素，这种方法没有考虑像素块内容的可预测性，当一个像素块中像素之间的相对关系一致时，该像素块的像素值更容易被intra的方向模式预测。文献【intraRateModelISCAS06xjing】认为像素块的梯度与预测残差之间存在线性关系，但是却没有考虑图像内容的复杂度。因此，我们结合像素块的内容复杂度和基于梯度的像素块可预测性，将帧内预测的预测残差能量表示为：

其中表示图像中使用帧内预测模式的像素块个数，为其中第个像素块，为原始像素的平均值，表示的方向可预测性强弱（）。

反映像素块内像素之间的关联性强弱，当像素之间的变化一致时，像素更容易被某种intra方向模式预测，则预测残差越小，反之，像素间变化多样，方向模式很难预测，则预测残差越大。因此需要能反映像素块的方向一致性，我们借鉴SIFT梯度直方图的方法，以梯度直方图的聚集性度量梯度方向的一致性。首先，为像素块中的每个像素值计算其梯度方向角：

为了反映方向角和intra方向模式的关联性，我们将intra方向模式粗分为8个区间，并将梯度值量化到8个方向，考虑到的取值范围为(-90°,90°)，将该曲线划分为8份，每份22.5°，梯度直方图表示为：

当梯度方向杂乱，即的方差小时，像素块方向性不明显，intra模式不能有效预测，残差大；当梯度方向集中，即的方差大时，像素块方向性明显，intra模式能有效预测，残差小。这里使用的标准差对像素块的方差加权，标准差表示为：

将标准差归一化：

其中标准差的最大值在方向梯度集中在某一个方向时获得，表示为：

所以，可预测性因子表示为：

当像素块方向杂乱（即各方向梯度值相近），，像素块自身差异不能有效预测，预测残差大；当像素块方向一致（即方向集中），，预测残差小。

**（2）帧内预测模式的预测残差能量估计**

对于使用帧间预测模式的像素块，预测像素由运动矢量和重建参考图像组合得到，因此预测残差的能量可以表示为：

其中是关键帧像素点由运动矢量指向的参考图像像素点。在大部分情况下接近1，预测残差能量转化为：

其中关键帧和知识库图像的原始像素的预测残差可以通过预分析在原始像素上获得，而参考图像的重建失真则可以基于上一部分对帧内预测模式的预测残差能量的估计计算得到，即：

综合上述所有公式，关键帧的预测残差能量可以表示为：

其中，只有第三项与参考图像的失真有关。那么关键帧的失真表示为：

其中均可以不经过实际编码而通过预分析得到。

* 普通图像（暂不考虑）

每个关键帧主要影响其之后一个intra period中的普通图像，由于这些普通图像通常使用固定的参考结构且量化参数相对固定，因此在给定hierarchical编码结构下，普通图像之间的失真依赖关系相对固定，因此普通图像的失真主要受到关键帧失真的影响，根据文献，我们认为一个intra period中普通图像的平均失真与关键帧的失真呈正比：

其中与intra period的长度和图像内容运动快慢相关，对于不同的序列应该取不同的值，但是对于同一序列中依赖同一知识库图像的关键帧和普通图像，应该是个定值。

**key-distortion expansion实验：需要验证一下两者之间是否有一个线性关系。**

### 建立码率受量化步长影响的估计模型

根据子序列的率失真代价，知识库编码的码率包含两部分：非知识库图像的码率，知识库图像的码率。主流的码率估计模型分为两种：R-Q模型和R-λ模型。R-Q模型描述预测残差变换量化之后所需要的比特数，并使用常数项描述非残差信息所需要的比特数。R-λ模型相较于R-Q模型的优势在于λ的取值是连续的而Q的取值是离散的，连续取值的λ可以更精确地估计码率，但是在为图像或图像块分配码率时，仍然需要根据R-λ模型获取的最佳λ就近选择离散的Q和QP以为图像或图像块分配量化参数，因此R-λ模型与R-Q模型在解决量化参数的选择问题中并没有差异。

R-Q模型包含两种，一种是quadric-R-Q模型，一种是linear-R-Q模型。两种模型都是根据R-D的模型得到的。根据文献，在变换系数服从zero mean Laplacian distribution的高码率情况下，如果失真定义为，码率可以表示为：

该公式进行泰勒展开在配合即可得到R-Q模型。（R-λ模型是基于低码率R-D模型R=aD^b推导的）。

根据上一节中对失真的估计方法，我们可以很容易的得到码率的估计方法。

**（1）知识库图像的码率估计模型**

对于知识库图像来说，知识库图像的失真可以表示为：

那么知识库图像的码率可以表示为：

由于，上式可以转化为：

**（2）关键帧的码率估计模型**

对于关键帧图像来说，其失真表示为：

那么关键帧的码率可以表示为：

同样的，由于，上式可以转化为：

**（3）普通图像的码率估计模型**

每个关键帧主要影响其之后一个intra period中的普通图像，由于这些普通图像通常使用固定的参考结构且量化参数相对固定，因此在给定hierarchical编码结构下，普通图像之间的码率依赖关系相对固定，因此普通图像的码率主要受到关键帧码率的影响，根据文献，我们认为一个intra period中普通图像的平均码率与关键帧的码率呈正比：

其中与intra period的长度和图像内容运动快慢相关，对于不同的序列应该取不同的值，但是对于同一序列中依赖同一知识库图像的关键帧和普通图像，应该是个定值。

### header rate model（暂不考虑）

上述模型主要考虑预测残差的码率，并没有考虑头信息（模式信息、MV信息、块划分信息等）的码率。大部分文献认为不同图像类型的头信息码率不同但是个常量，这种假设是基于h.264的块大小固定、模式信息少的情况。在HEVC的复杂块划分和多种模式下，头信息的码率可能不是常量，而是受到图像块划分的细碎程度影响，同时，对于关键帧参考知识库图像的情况，MV的长度可能更大，会消耗更多的比特数，因此头信息的码率也需要自适应调整。

**header Rate model验证实验：暂不考虑。**

### RDcost估计模型

**待得码率和失真模型确定后，综合得到RDcost模型。**

综合得到第m个知识库图像及其关键帧和普通图像的RD-cost的关系式：

为了便于表达，省略了其中m和rec的标记，即表示第m个知识库图像的量化步长，表示图像的重建失真，表示依赖第m个知识库图像的非知识库图像个数，表示参考第m个知识库图像的关键帧图像个数。

其中根据知识库图像和关键帧原始像素确定，是常数。解上述方程的方法有两个：

* 牛顿法：，其中
* 穷举法：DQP为离散值且取值范围有限，根据逐个代入公式以找到最接近0的值。可以采用这个方式。
  + 问题：怎么确定？
    - 使用训练样本训练参数，具体的，对训练数据进行编码，获取最佳（使导数）、以及，从而训练得到。
* DQP的定义，由于知识库图像是从关键帧中提取的，所以知识库图像的QP在关键帧的QP基础上进行调整。（关键帧的QP按照什么方法确定，知识库图像的QP是否对于所有方法都适用？）
  + 定义DQP为关键帧量化参数和知识库图像量化参数之间的差值：
  + 定义DQP为关键帧量化步长和知识库图像量化步长之间的倍率：

根据量化步长和量化参数之间的关系：

得到：

DQP增大，增大。

### 基于内容因素的系数获取（待完成）

**内容因素：**

**在上述R模型中，对于不同的序列、图像类型，常数项可能是不一样的，几乎所有的方法都是通过预编码或根据序列中已有编码数据预测的方法获取常数项，这说明现有模型仍然没有充分考虑不同序列、不同图像类型的内容因素。**

### 模型训练

* 训练样本采集：为了得到系数的统计值，从41个训练序列（包括70个知识库图像）的结果中获取训练样本。样本包括以下变量：
  + 自变量
  + 因变量
    - OPT-DQP
      * 此处optimization以RDcost为目标，因此从训练序列中获取每一个知识库图像及其所属关键帧的rate和psnr（转换为SSE）以计算RDcost。
        + lambda采用0.57\*2^((QP\_key-12)/3)。
        + R的单位是bits
        + NOTE：

~~需要确认RDcost的计算方法，rate的单位是kbps、bps、bits？？？~~

~~需要确认lambda的值~~

~~需要转化各自变量之间的数量级~~

* 训练结果：

## END

## 测试结果

### 测试序列

测试序列使用HM的B、C、D类序列，见下表。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **类别** | **序列** | **帧数** | **场景切换次数** | **I帧个数** | **lib个数** |
| B | BasketballDrive\_1920x1080\_50 | 500 | 1 | 11 | 3 |
| BQTerrace\_1920x1080\_60 | 600 | 1 | 10 | 2 |
| Cactus\_1920x1080\_50 | 500 | 1 | 11 | 1 |
| Kimono\_1920x1080\_24 | 240 | 2 | 11 | 2 |
| ParkScene\_1920x1080\_24 | 240 | 1 | 11 | 1 |
| C | BasketballDrill\_832x480\_50 | 500 | 1 | 11 | 1 |
| BQMall\_832x480\_60 | 600 | 1 | 10 | 1 |
| Flowervase\_832x480\_30 | 300 | 1 | 10 | 3 |
| PartyScene\_832x480\_50 | 500 | 2 | 12 | 2 |
| RaceHorses\_832x480\_30 | 300 | 1 | 10 | 2 |
| D | BasketballPass\_416x240\_50 | 500 | 1 | 11 | 2 |
| BlowingBubbles\_416x240\_50 | 500 | 1 | 11 | 2 |
| BQSquare\_416x240\_60 | 600 | 1 | 10 | 2 |
| RaceHorses\_416x240\_30 | 300 | 1 | 10 | 1 |

### 测试条件

LibVC测试软件基于HM16.15搭建，简称为LibVC-HM16.15。测试使用通用测试条件（Common Test Condition，CTC），见下表。

|  |  |
| --- | --- |
| 参数 | 值 |
| IntraPeriod | 1s（帧数由帧率决定） |
| I type | 基于场景切换的自适应I帧 |
| QP | 22,27,32,37 |
| GOP | 16 |

### 测试结果

下面展示了LibVC-HM16.15已有的编码效率。

* LibVC-HM16.15在不同DQP下编码序列的BD-rate：

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | 2 | 1 | 0 | -1 | -2 | -3 | -4 | -5 | -6 | -7 | -8 | -9 | -10 | -11 | Best |
| B | BasketballDrive | 0.0% | -0.1% | -0.4% | -0.6% | -0.8% | -0.7% | -0.5% | 0.1% | 1.2% | 2.6% | 4.3% | 5.9% | 8.0% | 10.8% | -0.8% |
| BQTerrace | -7.3% | -8.5% | -10.8% | -12.7% | -14.3% | -15.0% | -15.6% | -15.7% | -15.3% | -14.4% | -13.4% | -12.0% | -10.6% | -9.0% | -15.7% |
| Cactus | -3.8% | -4.7% | -7.2% | -9.8% | -12.2% | -14.0% | -15.7% | -16.9% | -17.7% | -18.1% | -18.2% | -18.1% | -17.5% | -16.6% | -18.2% |
| Kimono | -1.5% | -2.0% | -3.1% | -4.4% | -5.5% | -6.3% | -6.7% | -6.9% | -6.7% | -5.9% | -4.6% | -2.8% | -0.6% | 2.3% | -6.9% |
| ParkScene | -3.5% | -4.4% | -6.1% | -7.9% | -9.5% | -10.7% | -11.8% | -12.5% | -12.9% | -12.9% | -12.6% | -11.9% | -10.9% | -9.7% | -12.9% |
| C | BasketballDrill | -4.2% | -5.3% | -8.1% | -11.0% | -14.2% | -16.8% | -19.4% | -21.5% | -23.4% | -24.9% | -26.0% | -26.8% | -27.3% | -27.5% | -27.5% |
| BQMall | -0.4% | -0.6% | -1.1% | -1.8% | -2.5% | -2.9% | -3.4% | -3.6% | -3.8% | -3.9% | -3.8% | -3.6% | -3.2% | -2.9% | -3.9% |
| Flowervase | -3.5% | -4.5% | -6.0% | -7.5% | -8.8% | -9.3% | -9.9% | -10.1% | -9.9% | -9.4% | -8.3% | -7.3% | -5.5% | -3.6% | -10.1% |
| PartyScene | -4.3% | -5.2% | -6.6% | -8.0% | -9.4% | -10.6% | -11.6% | -12.4% | -13.0% | -13.3% | -13.5% | -13.4% | -13.1% | -12.7% | -13.5% |
| RaceHorses | 0.4% | 0.1% | -0.1% | -0.1% | 0.0% | 0.2% | 0.6% | 1.0% | 1.6% | 2.4% | 3.2% | 4.2% | 5.3% | 6.4% | -0.1% |
| D | BasketballPass | 0.1% | -0.1% | -0.9% | -1.7% | -2.5% | -3.2% | -3.7% | -4.1% | -4.5% | -4.6% | -4.8% | -4.7% | -4.6% | -4.5% | -4.8% |
| BlowingBubbles | -1.9% | -2.5% | -3.9% | -5.2% | -6.5% | -7.5% | -8.4% | -9.0% | -9.2% | -9.4% | -9.3% | -9.0% | -8.6% | -8.1% | -9.4% |
| BQSquare | -4.0% | -4.5% | -5.8% | -7.3% | -8.3% | -9.3% | -10.1% | -10.6% | -10.9% | -10.9% | -10.9% | -10.8% | -10.1% | -9.5% | -10.9% |
| RaceHorses | 0.3% | 0.0% | 0.0% | -0.2% | -0.3% | -0.2% | -0.2% | -0.1% | 0.0% | 0.4% | 0.6% | 0.8% | 1.3% | 1.8% | -0.3% |
| average | | -2.4% | -3.0% | -4.3% | -5.6% | -6.8% | -7.6% | -8.3% | -8.7% | -8.9% | -8.7% | -8.4% | -7.8% | -7.0% | -5.9% | -9.6% |

* LibVC-HM16.15在使用不同自适应QP决策方法下的编码效率：

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **adap-freq** | | | **adaq-dist** | | | **adaq-freq&dist** | | | **fixdqp-6** | | |
|  | Sequence | Y | U | V | Y | U | V | Y | U | V | Y | U | V |
| B | BasketballDrive | -0.7% | -2.9% | -2.5% | 0.3% | -6.0% | -4.3% | 0.3% | -4.7% | -3.5% | 1.4% | -6.5% | -4.2% |
| BQTerrace | -12.7% | -30.5% | -34.5% | -10.8% | -39.9% | -46.8% | -10.8% | -39.9% | -46.8% | -11.6% | -43.9% | -52.6% |
| Cactus | -16.4% | -32.9% | -34.7% | -15.8% | -26.3% | -26.4% | -15.8% | -26.3% | -26.4% | -16.3% | -29.1% | -29.4% |
| Kimono | -4.8% | -8.0% | -8.9% | -6.2% | -15.0% | -19.9% | -5.2% | -10.5% | -13.8% | -5.8% | -16.8% | -22.9% |
| ParkScene | -11.5% | -22.6% | -28.0% | -11.0% | -18.0% | -21.7% | -11.0% | -18.0% | -21.7% | -11.6% | -23.3% | -29.5% |
| C | BasketballDrill | -24.6% | -32.0% | -31.6% | -25.1% | -32.5% | -32.2% | -25.1% | -32.5% | -32.2% | -23.0% | -29.3% | -29.1% |
| BQMall | -3.3% | -14.2% | -13.4% | -3.0% | -8.9% | -8.5% | -3.0% | -8.9% | -8.5% | -3.2% | -11.7% | -11.6% |
| Flowervase | -7.9% | -12.6% | -11.5% | -7.0% | -38.5% | -33.8% | -7.0% | -38.5% | -33.8% | -10.0% | -41.5% | -37.3% |
| PartyScene | -11.2% | -18.0% | -19.3% | -12.3% | -23.3% | -25.4% | -12.3% | -23.3% | -25.4% | -12.8% | -26.3% | -28.5% |
| RaceHorses | 0.7% | -1.4% | -1.1% | 1.9% | -3.3% | -2.1% | 0.3% | -1.0% | -1.1% | 2.3% | -3.9% | -2.5% |
| D | BasketballPass | -3.2% | -6.8% | -4.0% | -3.4% | -7.8% | -4.6% | -3.4% | -7.8% | -4.6% | -4.9% | -11.0% | -6.6% |
| BlowingBubbles | -7.7% | -10.5% | -13.3% | -8.5% | -15.0% | -18.8% | -8.2% | -13.8% | -17.3% | -9.0% | -17.8% | -22.4% |
| BQSquare | -8.2% | -19.4% | -19.6% | -9.3% | -36.2% | -36.5% | -9.3% | -36.2% | -36.5% | -9.9% | -42.5% | -41.9% |
| RaceHorses | 0.6% | -2.4% | -2.0% | -0.1% | -1.3% | -1.3% | -0.1% | -1.3% | -1.3% | 0.1% | -2.1% | -2.0% |
|  | average | -7.9% | -15.3% | -16.0% | -7.9% | -19.4% | -20.2% | -7.9% | -18.8% | -19.5% | -8.2% | -21.8% | -22.9% |

# 本项研究取得的效果分析

# 附录

## 推导T(θ)

先列出多次使用的积分公式：

函数的形式为：

首先求解第一项：

然后求解第二项，先求解第二项中任意第k项：

那么，的第二项可以表示为：

其中加和项可以按照等比数列加和得到：

综上所述，函数表示为：

## 量化相关知识

* 模式RDO的和量化参数QPq之间的关系：

HM中和量化参数的关系：

* 运动搜索RDO的与之间的关系：
* 量化步长和量化参数之间的关系：
* 失真和量化步长之间的关系：
* 编码图像失真与参考图像失真之间的关系[[1]](#footnote-1)：

其中表示error propagation的intensity，失真以MSE度量。（NOTE：待认证，这种关系可能仅存在于差别不大的图像对之间）

* 基于原始像素计算的MSE与基于重建像素的MSE的关系：

其中是拉格朗日参数：

* 根据文献[[2]](#footnote-2)，D-R关系为：
* 根据文献[[3]](#footnote-3)，

## RD-cost相关知识

* MV的比特数计算方法：MVD使用指数哥伦布码编码，因此，MVD的比特数即为其指数哥伦布码的码长之和。码长求解方法：

## RD估计模型

### Quadratic Rate Distortion Model[1]

**前提假设：**

信源符合Laplacian分布：

给定失真度量方法：

则rate distortion function可以表示为：

**推导：**

进行泰勒展开得到：

展开得到：

舍弃高阶项，可以按照下式估计：

其中，对于不同类型（I/P/B）的帧，不同。

文献阐述失真可以使用1---average quantization scale of a frame with a specific picture prediction type（理解为帧的平均量化步长）、2---Mean squared error（MSE）、3---Mean absolute difference（MAD）度量，且上述估计模型仍然有效（疑问：跟假设中的失真度量不一致，可能也是验证过的）。

### first order Rate Distortion Model[3]

**假设：**

对于不同类型（I/P/B）的块，变换系数（即预测残差）编码所需比特数与1/q成正比（根据多次实验结果得到）。

**推导：**

R-Q关系表示为：

### λ-domain Rate Distortion Model[5][6][7]

**假设：**

R-D关系满足：

其中是与序列特性有关的参数。

**推导：**

由于R-D曲线的斜率即为λ，可以推导得：

那么R-λ关系可以表示为：

## 验证试验结果

### R-q/λ验证实验结果

**R-q/λ验证实验：需要实验验证R-q模型或者R-λ模型的正确性。**

**方法：**

**1---不考虑常数项的变化，对于每个序列，分别验证每个知识库图像和每个非知识库图像的R随q或λ的变化趋势是否符合表达式描述的趋势；**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| class | sequence | Lib-R-square | | | Seq-R-square | | |
| quadric-R-Q | linear-R-Q | R-λ | quadric-R-Q | linear-R-Q | R-λ |
| B | BasketballDrive\_1920x1080\_50 | 0.9864 | 0.9821 | 0.9778 | 0.9997 | 0.9758 | 0.9945 |
| BQTerrace\_1920x1080\_60 | 0.9976 | 0.9705 | 0.9862 | 0.9986 | 0.9768 | 0.9958 |
| Cactus\_1920x1080\_50 | 0.9914 | 0.9914 | 0.9914 | 0.9995 | 0.9624 | 0.9939 |
| Kimono\_1920x1080\_24 | 0.9940 | 0.9762 | 0.9949 | 0.9999 | 0.9957 | 0.9990 |
| ParkScene\_1920x1080\_24 | 0.9975 | 0.9975 | 0.9975 | 1.0000 | 0.9990 | 0.9998 |
| C | BasketballDrill\_832x480\_50 | 0.9998 | 0.9975 | 0.9983 | 1.0000 | 0.9990 | 0.9998 |
| BQMall\_832x480\_60 | 0.9980 | 0.9980 | 0.9980 | 1.0000 | 0.9993 | 0.9999 |
| Flowervase\_832x480\_30 | 0.9988 | 0.9971 | 0.9994 | 1.0000 | 0.9992 | 0.9995 |
| PartyScene\_832x480\_50 | 0.9980 | 0.9589 | 0.9900 | 1.0000 | 0.9982 | 0.9988 |
| RaceHorses\_832x480\_30 | 0.9971 | 0.9644 | 0.9893 | 1.0000 | 0.9978 | 0.9985 |
| D | BasketballPass\_416x240\_50 | 0.9994 | 0.9868 | 0.9974 | 1.0000 | 0.9984 | 0.9998 |
| BlowingBubbles\_416x240\_50 | 0.9981 | 0.9595 | 0.9887 | 1.0000 | 0.9980 | 0.9984 |
| BQSquare\_416x240\_60 | 0.9979 | 0.9656 | 0.9940 | 0.9998 | 0.9961 | 0.9981 |
| RaceHorses\_416x240\_30 | 0.9863 | 0.9863 | 0.9863 | 1.0000 | 0.9968 | 0.9975 |
| average | | 0.9957 | 0.9808 | 0.9921 | 0.9998 | 0.9923 | 0.9981 |
| train | beach\_1920x1080\_25 | 0.9989 | 0.9989 | 0.9989 | 1.0000 | 0.9994 | 0.9991 |
| City\_1280x720\_60 | 0.9985 | 0.9780 | 0.9952 | 0.9998 | 0.9915 | 0.9978 |
| Crew\_1280x720\_60 | 0.9989 | 0.9962 | 0.9966 | 1.0000 | 0.9965 | 0.9993 |
| taishan\_1920x1080\_25 | 0.9977 | 0.9706 | 0.9953 | 0.9996 | 0.9952 | 0.9968 |
| vidyo1\_720p\_60 | 0.9959 | 0.9959 | 0.9959 | 0.9999 | 0.9964 | 0.9988 |
| vidyo3\_720p\_60 | 0.9947 | 0.9947 | 0.9947 | 0.9999 | 0.9962 | 0.9982 |
| Traffic\_2560x1600\_30\_crop | 0.9927 | 0.9927 | 0.9927 | 1.0000 | 0.9989 | 0.9997 |
| sunflower\_1920x1080\_25 | 0.9965 | 0.9508 | 0.9898 | 0.9979 | 0.9973 | 0.9958 |
| toys\_and\_calendar\_1920x1080\_25 | 0.9997 | 0.9913 | 0.9963 | 0.9998 | 0.9772 | 0.9974 |
| vintage\_car\_1920x1080\_30 | 0.9996 | 0.9996 | 0.9996 | 1.0000 | 0.9961 | 0.9994 |
| walking\_couple\_1920x1080\_25 | 0.9948 | 0.9948 | 0.9948 | 0.9999 | 0.9943 | 0.9974 |
| harbour\_1280x720\_60 | 0.9895 | 0.9895 | 0.9895 | 0.9999 | 0.9844 | 0.9955 |
| crossroad\_720x576\_30 | 0.9906 | 0.9906 | 0.9906 | 0.9999 | 0.9838 | 0.9987 |
| Intersection\_1600x1200\_30 | 0.9914 | 0.9914 | 0.9914 | 0.9998 | 0.9913 | 0.9981 |
| tractor\_1920x1080\_25\_690 | 0.9963 | 0.9904 | 0.9954 | 1.0000 | 0.9999 | 0.9998 |
| pedestrian\_area\_1920x1080\_25\_250 | 0.9907 | 0.9860 | 0.9921 | 0.9995 | 0.9753 | 0.9936 |
| riverbed\_1920x1080\_25\_250 | 0.9946 | 0.9946 | 0.9946 | 1.0000 | 0.9998 | 0.9999 |
| rush\_hour\_1920x1080\_50\_500 | 0.9843 | 0.9698 | 0.9874 | 0.9999 | 0.9956 | 0.9976 |
| station2\_1920x1080\_30\_300 | 0.9950 | 0.9950 | 0.9950 | 1.0000 | 0.9947 | 0.9996 |
| Jets\_1280x720\_50\_500 | 0.9904 | 0.9541 | 0.9807 | 0.9998 | 0.9809 | 0.9976 |
| Night\_1280x720\_50\_460 | 0.9934 | 0.9934 | 0.9934 | 0.9998 | 0.9982 | 0.9998 |
| Spincalendar\_1280x720\_50\_550 | 0.9993 | 0.9802 | 0.9893 | 0.9998 | 0.9961 | 0.9996 |
| Optis\_1280x720\_60\_600 | 0.9993 | 0.9871 | 0.9934 | 0.9997 | 0.9997 | 0.9971 |
| BigShips\_1280x720\_60\_600 | 0.9998 | 0.9881 | 0.9959 | 0.9999 | 0.9974 | 0.9991 |
| Panslow\_1280x720\_50\_550 | 0.9973 | 0.9783 | 0.9852 | 0.9998 | 0.9699 | 0.9984 |
| ShuttleStart\_1280x720\_60\_600 | 0.9996 | 0.9976 | 0.9997 | 0.9997 | 0.9865 | 0.9974 |
| Sailormen\_1280x720\_60\_600 | 0.9948 | 0.9948 | 0.9948 | 1.0000 | 0.9974 | 0.9997 |
| Cyclists\_1280x720\_60\_600 | 0.9976 | 0.9976 | 0.9976 | 0.9999 | 0.9956 | 0.9995 |
| Raven\_1280x720\_60\_600 | 0.9998 | 0.9993 | 0.9997 | 0.9999 | 0.9991 | 0.9989 |
| Sheriff\_1280x720\_60\_600 | 0.9998 | 0.9907 | 0.9978 | 0.9999 | 0.9990 | 0.9993 |
| Basket\_ccir\_720x576\_25\_250 | 0.9959 | 0.9533 | 0.9917 | 0.9998 | 0.9941 | 0.9989 |
| Boat\_ccir\_720x576\_25\_220 | 0.9955 | 0.9955 | 0.9955 | 0.9996 | 0.9928 | 0.9980 |
| F1\_ccir\_720x576\_25\_220 | 0.9970 | 0.9970 | 0.9970 | 0.9998 | 0.9911 | 0.9963 |
| Football\_ccir\_720x576\_25\_220 | 0.9921 | 0.9921 | 0.9921 | 0.9998 | 0.9923 | 0.9970 |
| Container\_ccir\_720x480\_30\_300 | 0.9978 | 0.9978 | 0.9978 | 0.9999 | 0.9985 | 0.9995 |
| Tempete\_ccir\_720x480\_25\_260 | 0.9998 | 0.9812 | 0.9956 | 1.0000 | 0.9988 | 0.9996 |
| football\_cif\_352x288\_25\_250 | 0.9931 | 0.9931 | 0.9931 | 1.0000 | 0.9940 | 0.9977 |
| new\_cif\_352x288\_30\_300 | 0.9977 | 0.9977 | 0.9977 | 0.9999 | 0.9990 | 0.9997 |
| paris\_cif\_352x288\_30\_1065 | 0.9959 | 0.9690 | 0.9974 | 0.9998 | 0.9969 | 0.9992 |
| foreman\_cif\_352x288\_30\_300 | 0.9963 | 0.9963 | 0.9963 | 0.9999 | 0.9975 | 0.9983 |
| mobile\_cif\_352x288\_30\_300 | 0.9736 | 0.9736 | 0.9736 | 0.9998 | 0.9891 | 0.9963 |
| average | | 0.9953 | 0.9873 | 0.9937 | 0.9998 | 0.9934 | 0.9983 |

## LibVC-HM12.1的编码效率

第一版LibVC软件是基于HM12.1搭建的，简称为LibVC-HM12.1。LibVC-HM12使用长度为8帧的GOP编码结构，其在知识库图像使用不同DQP情况下的编码效率为：

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  | **dqp=0** | | | **dqp=3** | | | **dqp=6** | | | **max-gain** | | |
|  | Sequence | Y | U | V | Y | U | V | Y | U | V | Y | U | V |
| B | BasketballDrive | -1.9% | -5.8% | -4.6% | 0.0% | -0.9% | -0.6% | 1.7% | -1.7% | -0.6% | -1.9% | -5.8% | -4.6% |
| BQTerrace | -6.9% | -5.6% | -5.4% | -9.7% | -16.0% | -19.5% | -9.4% | -31.1% | -39.8% | -9.7% | -16.0% | -19.5% |
| Cactus | -5.7% | -4.3% | -3.0% | -11.4% | -12.5% | -13.1% | -14.5% | -23.1% | -23.5% | -14.5% | -23.1% | -23.5% |
| Kimono | -3.2% | -2.1% | -2.3% | -5.8% | -6.8% | -7.9% | -6.5% | -12.8% | -16.7% | -6.5% | -12.8% | -16.7% |
| ParkScene | -6.0% | -4.9% | -4.5% | -10.2% | -11.0% | -12.0% | -13.7% | -19.5% | -23.4% | -13.7% | -19.5% | -23.4% |
| C | BasketballDrill | -6.2% | -4.8% | -4.3% | -13.3% | -14.2% | -14.0% | -20.2% | -25.7% | -24.5% | -20.2% | -25.7% | -24.5% |
| BQMall | -0.8% | -0.3% | -0.9% | -1.9% | -2.8% | -3.2% | -2.3% | -6.9% | -6.5% | -2.3% | -6.9% | -6.5% |
| Flowervase | -4.3% | -4.4% | -3.7% | -8.0% | -15.1% | -14.1% | -10.5% | -33.7% | -32.1% | -10.5% | -33.7% | -32.1% |
| PartyScene | -5.3% | -4.1% | -3.8% | -8.6% | -10.2% | -11.1% | -11.0% | -19.6% | -21.6% | -11.0% | -19.6% | -21.6% |
| RaceHorses | -0.1% | 0.3% | 0.3% | 0.3% | -0.8% | -0.5% | 1.4% | -2.6% | -1.7% | -0.1% | 0.3% | 0.3% |
| D | BasketballPass | -0.6% | 0.0% | -0.6% | -2.8% | -2.8% | -2.1% | -4.4% | -7.9% | -4.6% | -4.4% | -7.9% | -4.6% |
| BlowingBubbles | -2.9% | -1.8% | -1.7% | -5.8% | -6.1% | -7.9% | -7.5% | -12.8% | -16.9% | -7.5% | -12.8% | -16.9% |
| BQSquare | -3.9% | -2.3% | -1.6% | -6.5% | -14.2% | -14.2% | -7.9% | -31.5% | -31.3% | -7.9% | -31.5% | -31.3% |
| RaceHorses | 0.0% | 0.1% | 0.0% | -0.2% | -0.6% | -0.6% | -0.1% | -1.7% | -0.9% | -0.2% | -0.6% | -0.6% |
|  | average | -3.4% | -2.8% | -2.6% | -6.0% | -8.1% | -8.6% | -7.5% | -16.5% | -17.4% | -7.9% | -15.4% | -16.1% |

## 影响编码参数的的权重不应该仅考虑一种类型的权重度量（例如HHI的感知权重），同时也应该考虑其他的权重（例如显著度、编码贡献度等），融合成一种普适（generalized）权重。

### 帧级QP调整

### 块级QP调整

在块级图像分析的情况下，能反映L和之间各区域之间的平均差异，此时，假设L被划分为K个区域，：

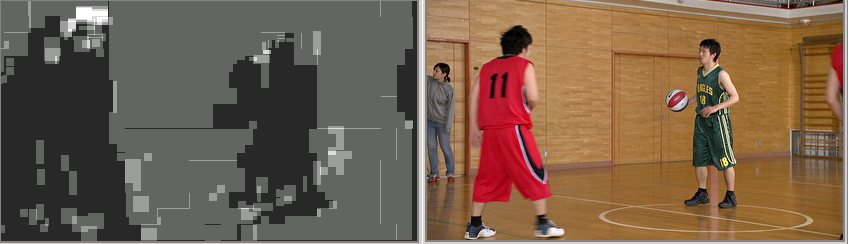
* library picture块级使用次数

library pictures的像素平均使用次数与帧使用次数：

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | seq | pixel-freq-avg | | | frame-freq | | |
| class B | BasketballDrive\_1920x1080\_50.yuv |  |  |  | 4 | 4 | 3 |
|  | BQTerrace\_1920x1080\_60.yuv |  |  |  | 3 | 7 |  |
|  | Cactus\_1920x1080\_50.yuv |  |  |  | 11 |  |  |
|  | Kimono\_1920x1080\_24.yuv |  |  |  | 5 | 4 |  |
|  | ParkScene\_1920x1080\_24.yuv |  |  |  | 8 |  |  |
| class C | BasketballDrill\_832x480\_50.yuv | 1.967 | 7.106 |  | 2 | 10 |  |
|  | BQMall\_832x480\_60.yuv | 2.738 |  |  | 10 |  |  |
|  | Flowervase\_832x480\_30.yuv | 1.688 | 2.149 | 1.903 | 2 | 4 | 4 |
|  | PartyScene\_832x480\_50.yuv | 4.281 | 4.194 |  | 6 | 6 |  |
|  | RaceHorses\_832x480\_30.yuv | 1.458 | 1.057 |  | 8 | 2 |  |
| class D | BasketballPass\_416x240\_50.yuv | 1.705 | 3.061 |  | 2 | 9 |  |
|  | BlowingBubbles\_416x240\_50.yuv | 2.82 | 3.545 |  | 5 | 6 |  |
|  | BQSquare\_416x240\_60.yuv | 3.625 | 2.976 |  | 5 | 5 |  |
|  | RaceHorses\_416x240\_30.yuv | 1.757 |  |  | 10 |  |  |

* 知识库图像的像素被参考次数与编码效率的样例
  + D-basketpass

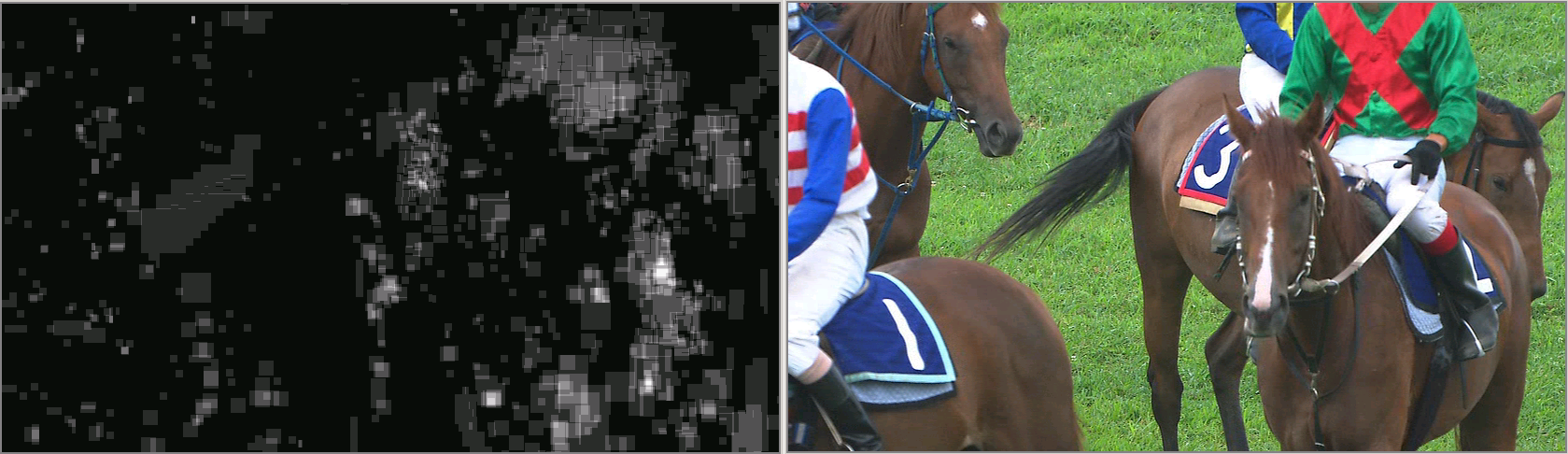
|  |  |  |  |
| --- | --- | --- | --- |
|  | 像素级参考次数 | 帧级参考次数 |  |
| 知识库图像1 | 1.705 | 2 |  |
| 知识库图像2 | 3.061 | 9 |  |
|  | Y | U | V |
| 编码效率 | -4.4% | -7.9% | -4.6% |

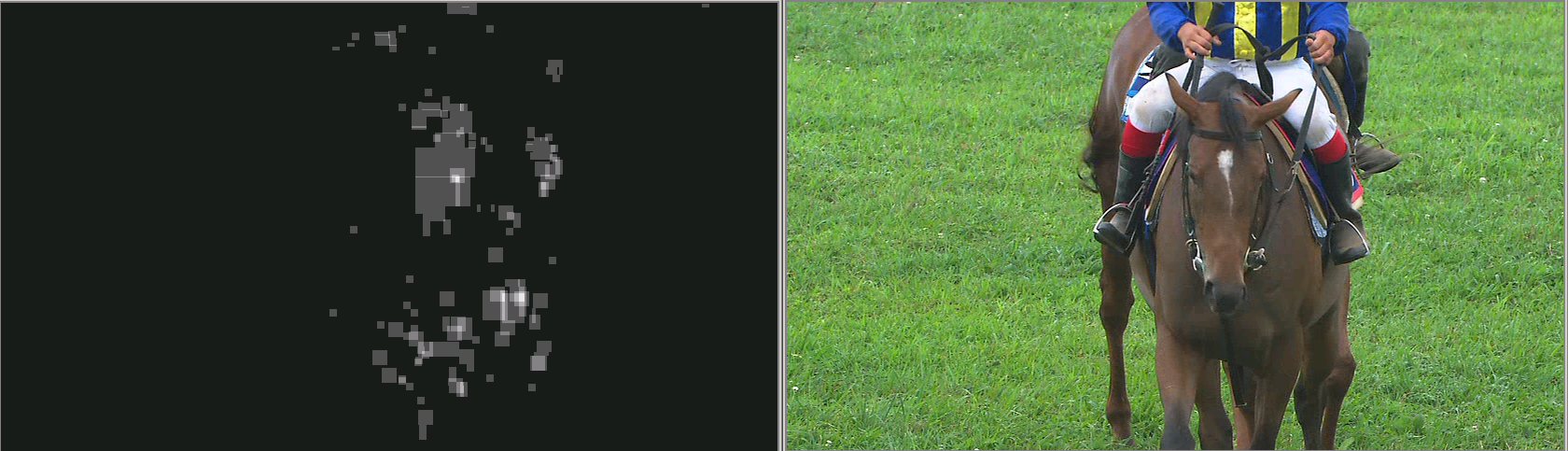




* + C-racehorse

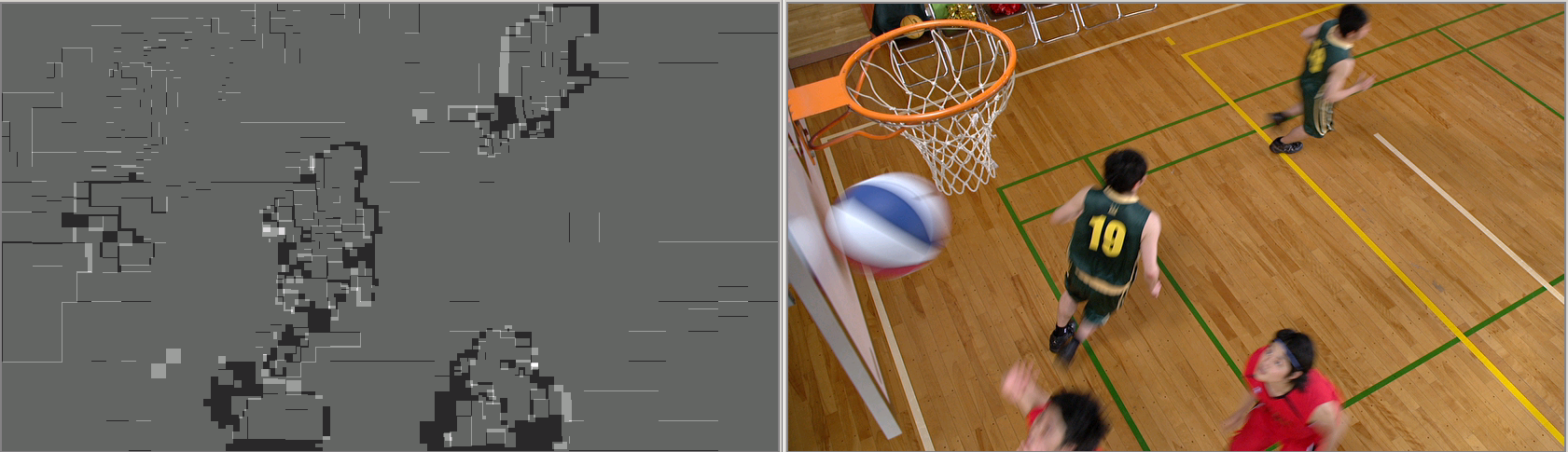
|  |  |  |  |
| --- | --- | --- | --- |
|  | 像素级参考次数 | 帧级参考次数 |  |
| 知识库图像1 | 1.458 | 8 |  |
| 知识库图像2 | 1.057 | 2 |  |
|  | Y | U | V |
| 编码效率 | 1.4% | -2.6% | -1.7% |

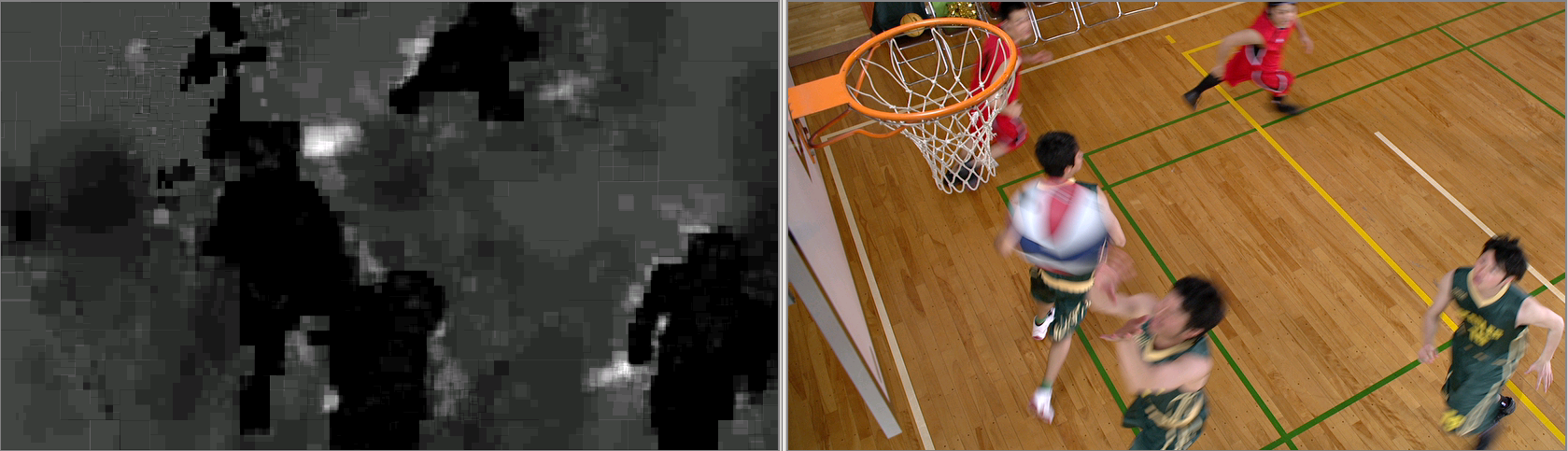




* + C-basketdrill

|  |  |  |  |
| --- | --- | --- | --- |
|  | 像素级参考次数 | 帧级参考次数 |  |
| 知识库图像1 | 1.967 | 2 |  |
| 知识库图像2 | 7.106 | 10 |  |
|  | Y | U | V |
| 编码效率 | -20.2% | -25.7% | -24.5% |





# 参考文献

1. 【rateCtrlTCSVT97thChiang】Tihao Chiang and Ya-Qin Zhang, "A new rate control scheme using quadratic rate distortion model," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 7, no. 1, pp. 246-250, Feb 1997.
2. 【rateCtrlTIP16tszhao】T. Zhao, Z. Wang and C. W. Chen, "Adaptive Quantization Parameter Cascading in HEVC Hierarchical Coding," in IEEE Transactions on Image Processing, vol. 25, no. 7, pp. 2997-3009, July 2016.
3. 【rateCtrlTCSVT05swma】Ma, Siwei, Wen Gao, and Yan Lu. "Rate-distortion analysis for H. 264/AVC video coding and its application to rate control." IEEE Transactions on Circuits and Systems for Video Technology 15.12 (2005): 1533-1544.
4. 【adapQpICIP17ymzhou】MP.PC.4: TEMPORAL CORELATION BASED HIERARCHICAL QUANTIZATION PARAMETER DETERMINATION FOR HEVC VIDEO CODING. Yimin Zhou, Hongyu Wang, Ling Tian, Ce Zhu, University of Electronic Science and Technology of China
5. 【rateCtrlTIP13bli】B. Li, H. Li, L. Li and J. Zhang, " $\lambda $ Domain Rate Control Algorithm for High Efficiency Video Coding," in IEEE Transactions on Image Processing, vol. 23, no. 9, pp. 3841-3854, Sept. 2014.
6. 【rateCtrlTMM15lili】L. Li, B. Li, D. Liu and H. Li, "$\lambda$-Domain Rate Control Algorithm for HEVC Scalable Extension," in IEEE Transactions on Multimedia, vol. 18, no. 10, pp. 2023-2039, Oct. 2016.
7. 【rateCtrlTCSVT15ycgong】Y. Gong, S. Wan, K. Yang, Y. Yang and B. Li, "Rate-Distortion-Optimization-Based Quantization Parameter Cascading Technique for Random-Access Configuration in H.265/HEVC," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 27, no. 6, pp. 1304-1312, June 2017.
8. 【RDModelTSP98Mallat】S. Mallat and F. Falzon, "Analysis of low bit rate image transform coding," in IEEE Transactions on Signal Processing, vol. 46, no. 4, pp. 1027-1042, Apr 1998.
9. 【RDModelICIP04mdai】M. Dai, D. Loguinov and H. Radha, "Rate-distortion modeling of scalable video coders," Image Processing, 2004. ICIP '04. 2004 International Conference on, 2004, pp. 1093-1096 Vol.2.
10. 【rateCtrlTMM12fshao】F. Shao, G. Jiang, W. Lin, M. Yu and Q. Dai, "Joint Bit Allocation and Rate Control for Coding Multi-View Video Plus Depth Based 3D Video," in IEEE Transactions on Multimedia, vol. 15, no. 8, pp. 1843-1854, Dec. 2013.
11. 【rateCtrlTCSVT07dkkwon】D. K. Kwon, M. Y. Shen and C. C. J. Kuo, "Rate Control for H.264 Video With Enhanced Rate and Distortion Models," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 17, no. 5, pp. 517-529, May 2007.
12. 【rateCtrlMMSP07dkkwon】D. K. Kwon, Y. Cho and C. C. J. Kuo, "A Simplified Rate Control Scheme for Non-Conversational H.264 Video," 2007 IEEE 9th Workshop on Multimedia Signal Processing, Crete, 2007, pp. 284-287.
13. 【rateCtrlICIP11jyliu】J. Liu, Y. Cho and Z. Guo, "Practical rate control algorithm for temporal scalability in scalable video coding," 2011 18th IEEE International Conference on Image Processing, Brussels, 2011, pp. 1641-1644.
14. 【rateCtrlTCSVT07yliu】Y. Liu, Z. G. Li and Y. C. Soh, "A Novel Rate Control Scheme for Low Delay Video Communication of H.264/AVC Standard," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 17, no. 1, pp. 68-78, Jan. 2007.
15. 【rateCtrlTCSVT08yliu】Y. Liu, Z. G. Li and Y. C. Soh, "Rate Control of H.264/AVC Scalable Extension," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 18, no. 1, pp. 116-121, Jan. 2008.
16. 【rateCtrlISCAS07yliu】Y. Liu, Y. C. Soh and Z. G. Li, "Rate Control for Spatial/CGS Scalable Extension of H.264/AVC," 2007 IEEE International Symposium on Circuits and Systems, New Orleans, LA, 2007, pp. 1746-1750.
17. 【rateCtrlTIP11sdhu】S. Hu, H. Wang and S. Kwong, "Adaptive Quantization-Parameter Clip Scheme for Smooth Quality in H.264/AVC," in IEEE Transactions on Image Processing, vol. 21, no. 4, pp. 1911-1919, April 2012.
18. 【quantizeTIT68gish】H. Gish and J. Pierce, "Asymptotically efficient quantizing," in IEEE Transactions on Information Theory, vol. 14, no. 5, pp. 676-683, September 1968.
19. 【RDOTIP09xli】X. Li, N. Oertel, A. Hutter and A. Kaup, "Laplace Distribution Based Lagrangian Rate Distortion Optimization for Hybrid Video Coding," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 19, no. 2, pp. 193-205, Feb. 2009.

1. Zhao, Tiesong, Zhou Wang, and Chang Wen Chen. "Adaptive quantization parameter cascading in HEVC hierarchical coding." IEEE Transactions on Image Processing 25.7 (2016): 2997-3009. [↑](#footnote-ref-1)
2. Li, Bin, et al. "$\lambda $ Domain Rate Control Algorithm for High Efficiency Video Coding." IEEE transactions on Image Processing 23.9 (2014): 3841-3854. [↑](#footnote-ref-2)
3. Ma, Siwei, Wen Gao, and Yan Lu. "Rate-distortion analysis for H. 264/AVC video coding and its application to rate control." IEEE Transactions on Circuits and Systems for Video Technology 15.12 (2005): 1533-1544. [↑](#footnote-ref-3)