

摘要

在资源受限条件下,为获得最优恢复视频质量,需要通过码率控制技术为视频编码器确定合适的编码参数。该编码参数的准确程度直接影响编码器性能。因此,设计合适的码率控制方案和算法成为提高视频编码器性能的关键技术之一。本文在 H.264/AVC 视频编码标准基础上,研究了实时低码率视频码率控制技术,主要研究内容如下:

1. 序列初始量化参数的确定。本文研究了一种确定初始量化参数的方案。在该方案中,考虑了图像复杂度对确定编码参数的影响,研究了一种利用 Sobel 算子衡量图像复杂度的方法,并进一步研究确定了该方案模型参数与图像复杂度之间的关系。由于该方案考虑了图像复杂度的影响,获得的初始量化参数较为适应视频复杂度的差异,从而提高了编码器的编码效率。

2. 研究了一种低码率线性码率控制方案。在分析一些已有码率控制方案的基础上,本文通过理论推导建立了一种基于码率-量化参数的码率控制方案,并对该方案进行了验证。由于该方案可以直接确定量化参数,从而避免了由量化步长转化为量化参数产生的舍入误差。实验结果表明该方案具有良好的性能。

3. 低码率跳帧算法的改进。在分析低码率帧间预测模式与编码效率关系的统计特性基础上,研究了一种降低跳帧发生概率的算法。该算法根据缓冲区的不同状态实时更新量化参数,并利用帧间预测模式与编码效率关系的统计特性为当前宏块确定最终编码模式。实验结果表明,在低码率情况下,本文研究的算法可以有效降低跳帧发生的概率。

关键词: 码率控制 初始 QP 选择 跳帧 H.264/AVC

Abstract

To obtain the optimal reconstructed video quality, bit-rate control technology should be used to select the proper encoding parameters for encoders with the constrained resources. Since coding efficiency highly depends on the accuracy of the selected coding parameters, how to design an appropriate rate control method is a key enabling factor for video coding. This thesis investigates rate control mechanism in low bit-rate and real-time applications based on H.264/AVC video coding standard. Major contributions are summarized as follows:

1. A method for ascertaining initial quantization parameter is designed in this thesis. In the designed method, the influence of image complexity on quantization parameter determination is taken into account, and an image complexity metric based on Sobel operator is presented. Furthermore, the relationship between the model parameter and the image complexity is studied. Due to the consideration of image complexity in the designed method, the obtained initial quantization parameter can adapt to the video complexity differences, thus improving the efficiency for encoders.

2. Based on the analysis of traditional bit-rate control, a linear bit-rate control scheme is established and further demonstrated. With the established scheme, the quantization parameter can be determined directly, thus avoiding the round-off error, which is brought by the transformation from quantization step into quantization parameter. The experimental results show the established scheme can obtain favorable performance.

3. Modification for frame skipping algorithm in low bit-rate. In this thesis, the statistical property for the relationship between coding efficiency and macro-block prediction modes is analyzed in low bit-rate, and then a method for reducing the probability of frame skipping is designed. In the designed method, the quantization parameter is updated in real time according to the buffer state, and the final coding modes are determined using the statistical property for the relationship between coding efficiency and macro-block prediction modes. The experimental result shows the designed method can reduce the probability of the occurrence of frame skipping in low bit-rate.

Keywords: rate control initial QP selection frame skipping H.264/AVC

第一章 绪论

1.1 引言

数字视频因具有直观、形象、简单和易懂等优点,成为目前多媒体通信系统的主要信息载体。据统计,人类通过视觉所获取的信息约占全部信息的 70%。因此,数字视频系统受到人们的广泛关注^[1-2]。然而,其庞大的数据量对存储空间提出了严格的要求,同时通信系统有限的带宽也制约了其应用和发展。为了解决这一问题,国际标准化组织旗下的活动图像专家组(International Organization of Standardization/Moving Picture Experts Group, ISO/MPEG)和国际电信联盟旗下的视频编码专家组(International Telecommunication Union/Video Coding Experts Group, ITU/VCEG)制定了一系列视频编码标准^[3-16]。2003 年 5 月,由 VCEG 和 MPEG 组织组成的视频联合工作组(Joint Video Team, JVT)制定了 H.264/AVC 视频编码标准。与已有的视频编码标准相比,该标准在编码效率上有了较大提高^[17-18]。

为了在信道带宽和传输时延受限的情况下,有效地传输视频数据以满足网络视频业务的服务要求,通常需要对视频编码进行码率控制。即使是在视频存储业务中,由于解码器缓冲区的限制,也需要对视频编码过程进行码率控制。所谓码率控制就是通过设置适当的编码参数,在保证恢复视频质量的前提下,使视频编码产生的码率尽可能接近目标码率。其目的主要有两个,一是根据信道带宽调整编码参数,使视频编码器的输出比特率尽量接近信道带宽;二是在给定码率限制下使编码失真达到最小。因此,码率控制是视频编码器的重要组成部分。本文在 H.264/AVC 视频编码标准基础上,对其推荐使用的码率控制机制进行研究,主要对序列初始量化参数的选择、P 帧量化参数的确定以及跳帧算法进行了改进。

本章内容安排如下:1.2 小节介绍了 H.264/AVC 视频编码基本过程,并对与码率控制相关的关键编码技术,如帧内预测、帧间预测、率失真优化准则以及码率控制机制的研究意义进行了详细阐述;1.3 小节给出了本文的研究内容和组织结构。

1.2 H.264/AVC 视频编码技术简介

为了更好地理解码率控制机制的作用,本小节首先介绍了 H.264/AVC 视频编码器的整个编码过程,然后简要介绍了与码率控制相关的关键编码技术。最后,分析了码率控制的研究意义。

1.2.1 H.264/AVC 的基本编码过程

图 1.1 给出了与 H.264/AVC 视频编码标准兼容的编码器结构示意图^[19]。该编

码器采用基于块的预测编码和变换编码的混合编码方法。在图 1.1 中, 首先将视频帧(或场)划分为 16×16 的宏块(Macro Block, MB), 然后对每一个宏块进行预测编码。之后, 对最优预测模式相应的预测差值进行变换得到相应变换系数, 再对变换系数进行量化操作, 即可得到相应量化系数。最后, 对量化系数进行变长编码或算术编码得到相应的视频压缩码流。由于当前编码帧可能作为后续编码帧的参考帧, 所以, 必须在编码端对当前帧进行恢复, 即实现本地解码。为提高恢复帧质量, 有必要对其进行滤波处理, 并将滤波之后的图像作为后续编码帧的参考图像。

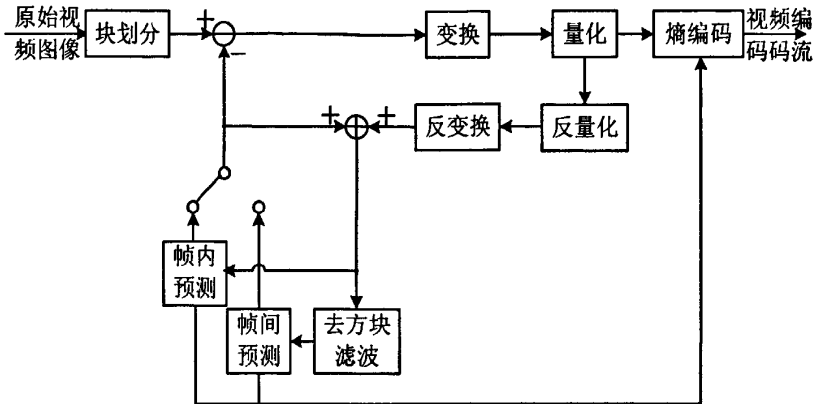


图 1.1 H.264/AVC 的编码框架

在视频压缩过程中, 编码器通过编码控制算法确定当前编码单元的编码参数, 例如, 使用率失真优化准则选择最佳编码模式, 采用码率控制机制确定合适的量化参数等。由于编码器采用这些编码参数直接对序列进行编码, 参数的准确度严重影响了编码器的输出码率和恢复视频质量。所以, 编码控制算法在视频编码器中起着至关重要的作用。作为编码控制算法之一的率失真优化准则, 其模型参数由码率控制得到的量化参数确定。也就是说, 码率控制算法影响率失真优化准则的性能。因此, 码率控制算法对编码器的性能具有重大影响。

1.2.2 与码率控制相关的关键编码技术

在资源受限条件下, 为获得最优恢复视频质量, 需要通过码率控制为编码器确定合适的量化参数。该参数不仅影响变换系数的量化精度, 还影响率失真优化准则为每一个宏块确定的最佳预测模式。为了更清楚地认识率失真整个优化过程, 以更好地理解码率控制机制在率失真优化准则中的作用, 本小节首先介绍了宏块的预测模式, 然后阐述了率失真优化过程, 最后分析了码率控制机制在率失真优化准则中的作用。

1. 去除空间冗余的帧内预测模式^[20-21]

H.264/AVC 视频编码标准的帧内预测技术, 充分挖掘了图像自身的空域相关

性。这主要体现在两个方面，一是考虑到视频内容多样性，即有些纹理丰富，有些纹理平坦，该标准采用自适应预测块大小。其中，最大预测块可设置为 16×16 ，最小可设置为 4×4 ；二是考虑到图像纹理方向对预测结果的影响，该标准根据预测块大小，采用不同预测方向的预测模式。其中， 16×16 块的预测模式有 4 种， 4×4 块的预测模式有 9 种，分别如表 1.1、表 1.2 所示。

由于采用了大小自适应变化的块以及多种不同方向的预测模式，H.264/AVC 的帧内预测能够在尽量保留图像细节信息的前提下，较大程度地去除图像空间冗余信息，从而提高编码器的编码效率。

表 1.1 16×16 亮度分量块的预测模式

模式	模式的具体实现
模式 0(垂直)	由当前宏块上边的像素点预测当前宏块中相应像素点的数值
模式 1(水平)	由当前宏块左边的像素点预测当前宏块中相应像素点的数值
模式 2(DC)	由当前宏块上边和左边像素点的平均数值预测当前宏块中相应的像素点
模式 3(平面)	用线性“平面(plane)”函数对当前编码宏块上边和左边的宏块的样值进行插值

表 1.2 4×4 亮度分量块的预测模式

模式	模式的具体实现
模式 1(垂直)	由当前编码宏块相邻的上边的样值进行垂直方向的外推
模式 2(水平)	由当前编码宏块相邻的左边的样值进行水平方向的外推
模式 3(直流)	利用当前编码宏块相邻的上边和左边的像素点平均值预测当前编码宏块中的像素点
模式 4(左下方向对角线)	在当前编码宏块的右上方和左下方之间沿 45° 方向进行插值
模式 5(右下方向对角线)	对当前编码宏块沿着右下方 45° 方向进行插值
模式 6(垂直-向右)	沿着右下与垂直线呈 30° 的方向进行插值
模式 7(水平-向下)	沿着右下与水平线呈 30° 的方向进行插值
模式 8(垂直-向左)	沿着左下与垂直向呈 30° 的方向进行插值
模式 9(水平-向上)	沿着右上与水平线呈 30° 的方向进行插值

2. 去除时域冗余的帧间预测模式

在以往编码标准中，帧间待编码块最小尺寸为 8×8 。由于视频内容的多样性，较大块可能包含多个具有不同形状或者不同运动状态的对象，特别是复杂纹理区域的块，用 1 个 16×16 或 4 个 8×8 的块都不能准确描述该块的纹理信息，导致其预测差值较大，从而降低编码效率。因此，H.264/AVC 视频编码标准对宏块进行了进一步划分。对于亮度分量，每个宏块(16×16)可划分为 1 个 16×16 、2 个 16×8 、

2 个 8×16 或 4 个 8×8 。而 8×8 子宏块可以进一步划分为：1 个 8×8 、2 个 8×4 、2 个 4×8 或 4 个 4×4 。

帧间预测模式的多样性使 H.264/AVC 编码器可以根据不同视频内容特性自适应地选择不同预测模式。例如，对于纹理变化平缓的区域，可以使用较大块进行预测，而对于纹理较多的区域则使用较小块进行预测，从而提高编码效率。

3. 率失真优化准则

H.264/AVC 视频编码器以宏块为基本单位进行编码。为使每个宏块获得最优编码模式如宏块预测模式，宏块运动矢量等，H.264/AVC 编码器采用了基于拉格朗日乘法的率失真优化准则，如下式所示，

$$R_{dcost} = D_{istortion} + \lambda B_{irate} \quad \text{式(1-1)}$$

其中， R_{dcost} 、 $D_{istortion}$ 和 B_{irate} 分别表示采用某种编码模式进行编码的率失真代价、编码失真和编码比特数。 λ 为拉格朗日乘子，它是量化参数的函数。公式(1-1)给出的率失真代价综合考虑了恢复视频质量($D_{istortion}$)和编码比特数(B_{irate})两个因素，选择率失真代价最小的编码模式是恢复视频质量和编码码比特率之间一个折中的方案^[20]。

对于不同类型帧的宏块，其最佳预测模式的选择过程不同。如果当前帧为 I 帧，为其宏块选择最佳预测模式，需遍历帧内预测所有预测模式。利用式(1-1)计算相应的 R_{dcost} ，最终选择使 R_{dcost} 最小的预测模式作为当前宏块的最佳预测模式。如果当前帧为 P 帧，除了对帧内编码的预测模式进行遍历之外，还要遍历基于时域帧间编码的所有预测模式(不同宏块划分对应不同预测模式)。在所有预测模式(包括帧内和帧间预测模式)中，选择使 R_{dcost} 最小的预测模式作为 P 帧宏块的最佳预测模式。

由上述可知，H.264/AVC 视频编码器根据基于拉格朗日乘法的率失真优化准则确定当前编码宏块的最佳预测模式。由于编码器采用该预测模式直接对当前宏块进行编码，所以该预测模式直接影响当前宏块的编码比特数和恢复宏块的失真。因此，率失真优化准则的性能在很大程度上影响着编码器的编码效率。

通过式(1-1)拉格朗日乘子可知，率失真优化准则的性能受码率控制影响。码率控制根据当前有限的资源为编码器确定量化参数，编码器利用该参数计算率失真优化准则的拉格朗日乘子。由于不同量化参数计算得到的拉格朗日乘子不同，使得特定的预测模式通过式(1-1)计算得到的 R_{dcost} 不同，最终导致当前宏块根据率失真优化准则选择的最佳预测模式不同。由此可见，码率控制所确定的量化参数间接地影响宏块的最佳预测模式选择，进而影响编码器的性能。

1.2.3 研究码率控制机制的意义

受通信信道带宽和存储空间限制, 视频应用对编码效率有很高的要求。而现有的无损编码方法所能达到的编码效率却远远不能满足实际视频应用, 尤其是实时性要求较高、码率较低的应用。如果考虑人类的视觉特性, 在保证恢复视频质量无显著下降的前提下, 可以采用合适的量化参数对视频数据进行量化, 以去除视频信号的生理视觉冗余信息, 获得远远高于无损编码的编码效率。由于量化操作所用的量化参数直接影响编码器的输出比特率和恢复视频质量, 因此, 有必要为编码器确定一个合理的量化参数。为解决这一问题, 引入了码率控制机制。码率控制根据给定的目标比特率、图像格式以及在编码过程中获得的图像复杂度、缓冲区占有度等信息为编码器实时地更新量化参数, 以在资源受限条件下, 用最少的编码比特数获得质量最优的恢复视频。

在视频序列中, 每一帧内容的运动剧烈程度以及图像纹理复杂度的不断变化, 使编码器的输出编码比特数也在不断变化。然而, 大部分信道带宽固定不变, 如 PSTN、ISDN 等。即使信道带宽随着时间推移而不断变化, 如因特网的信道, 其传输速率的变化也不一定与编码比特流速率的变化相一致。由于编码比特率与信道带宽不一致将会导致信道带宽的浪费或编码比特流的丢失, 从而严重影响恢复视频的质量。为了解决这一问题, H.264/AVC 视频编码标准所推荐使用的码率控制方案引入了缓冲区机制。编码器利用该缓冲区对视频压缩码流进行平滑, 使输出的比特率大小与当前信道带宽相匹配。然而, 缓冲区容量是有限的。如果不根据缓冲区状态更新编码参数, 将会导致缓冲区“溢出”。缓冲区“上溢”意味着图像信息的丢失, 缓冲区“下溢”意味着信道带宽的浪费。为了防止缓冲区发生“上溢”和“下溢”现象, 必须通过码率控制实时调整编码过程使用的量化参数。

码率控制必须根据当前资源为编码器确定合适的量化参数。如果量化参数过大, 用于当前序列的编码比特数就会减少。这可能引起缓冲区“下溢”, 导致信道带宽浪费。与此同时, 过大的量化参数使图像严重失真, 从而降低恢复视频质量; 如果量化参数过小, 虽然可以尽可能地保留图像细节信息, 但却使得量化系数较大, 导致编码比特数大幅度增多。这可能引起缓冲区“上溢”而丢失编码信息, 从而严重影响恢复视频质量。除此之外, 量化参数还影响宏块最佳编码模式的选择。因此, 为在资源受限条件下, 获得尽可能高的恢复视频质量, 必须通过码率控制为编码器确定合适的量化参数。

现有的各种视频编码标准虽然没有对码率控制做出明确的规定, 但都为自己推荐了相应的码率控制方案, 如 MPEG-2 的 TM^[22], H.263 的 TMN^[23], MPEG-4 的 VM8^[24]等。然而, 以往的码率控制对 H.264/AVC 视频编码标准并不适用, 并且研究 H.264/AVC 视频编码标准的码率控制比研究其它视频编码标准的码率控制更

加困难。这是因为该视频编码标准的码率控制存在“鸡蛋悖论”，即对当前编码单元进行编码时，需要根据率失真优化准则为当前编码单元选择最佳编码模式，而最佳编码模式的选择需要已知当前编码单元的量化参数。量化参数的确定需要已知当前编码单元与其预测值之间的平均绝对差(Mean Absolute Difference, MAD)。然而，该 MAD 只有在获得了最佳编码模式之后才能确定。

目前，H.264/AVC 视频编码标准推荐使用的码率控制算法主要有两种。一种是基于 TM5^[22]的改进算法^[25-27]。该算法首先利用假想参考解码器(Hypothetical Reference Decoder, HRD)为当前编码单元分配编码比特数。然后，利用率失真模型为编码器确定编码参数。该算法可以实现帧层和宏块级的码率控制，其计算复杂度较高；另一种是基于基本编码单元的自适应码率控制算法^[28-29]，它引入了流量流体模型、基本编码单元等新概念。该算法首先利用给定的帧率、目标比特率以及在编码过程中获得的缓冲区占有度、目标缓冲区级别，通过流量流体模型和线性自回归理论为基本编码单元分配编码比特数。然后，利用已编码帧基本编码单元的实际 MAD，通过提出的线性模型预测当前基本编码单元的 MAD。从而，解决了 H.264/AVC 码率控制的“鸡蛋悖论”问题。最后，该算法利用二次率失真模型为当前基本编码单元确定量化参数。由于该算法充分利用了视频序列之间的时间相关性以及较为准确的二次率失真模型，其码率控制效果较前者算法有所提高。

由上述可见，码率控制影响着缓冲区状态、宏块的最佳预测模式等，进而影响着编码器的性能。因此，对其研究具有重要意义。

1.3 本文研究内容及组织结构

1.3.1 本文研究内容

码率控制是视频编码器的重要组成部分，其目的是在资源受限条件下，为编码器确定合适的编码参数，以获得最优的恢复视频质量。它贯穿于整个视频编码过程，不仅仅影响着编码比特率的稳定，还影响着整个视频序列的编码失真。码率控制与具体应用环境密切相关，不同的应用环境对码率控制的准确性、复杂度、实时性等要求不同。目前，H.264/AVC 视频编码标准推荐使用的码率控制算法，在给定目标码率较高时，确定的编码参数较为准确，编码器能够获得较高质量的恢复视频。然而，当给定目标码率较低时，码率控制的性能较差，从而降低了编码器的编码效率。对此，本文结合国家自然科学基金“提高多视点视频编码效率的技术研究(项目编号: 60772134)”和西安电子科技大学-华为多媒体通信联合实验室专项研究课题，探索性地研究了低码率情况下的码率控制。

本论文主要研究内容如下：

1. 序列初始量化参数的选择。序列初始量化参数是指编码视频序列第一帧所

采用的量化参数。本文通过实验分析了初始量化参数对整个视频编码质量的影响。由于不合适的初始量化参数可能引起恢复视频质量的波动，甚至导致跳帧而使图像运动不连续。因此，有必要对初始量化参数进行研究。在目标码率给定的情况下，对于分辨率相同、纹理复杂度不同的视频，其对应的初始量化参数应该不同。但是，H.264/AVC 视频编码标准推荐使用的码率控制在确定初始量化参数时，并未考虑视频纹理复杂度对初始量化参数的影响。本文通过实验研究了一种确定初始量化参数的方法，并对该方法进行分析，发现图像复杂度不同使得序列的最佳初始量化参数存在一定偏差。为确定由图像复杂度不同导致的初始量化参数偏差，本文利用 Sobel 算子研究了一种衡量图像复杂度的方法，并对其准确性进行了验证。在此基础上，本文深入分析了图像复杂度与初始量化参数偏差之间的关系，并给出了求解该偏差的方法。利用本文确定初始量化参数的方法，可以在一定程度上改善实时低码率情况下码率控制的性能，进而提高编码器的性能。

2. 低码率线性码率控制方案。在 IPPP 编码结构中，P 帧出现的概率高达 90%，因此，为 P 帧确定合适的量化参数具有重要意义。目前，H.264/AVC 编码器利用码率与量化步长之间的关系确定量化步长，并将其转换为相应的量化参数之后，才能对当前编码单元进行编码。由于量化步长转换为量化参数的过程中存在不可避免的舍入误差，从而影响码率控制的性能。本文介绍了一些经典的码率控制方案，并分析了低码率情况下其性能下降的原因。在此基础上，通过理论推导研究了一种基于码率-量化参数的线性码率控制方案，并利用 H.264/AVC 的参考软件 JM15.1 对该方案进行了验证。实验结果表明，该方案可显著提高编码器的效率。

3. 低码率跳帧策略的改进算法。码率控制中，缓冲区大小与时延有关。最大时延越小，缓冲区越小；而缓冲区越小，发生跳帧的概率越大。跳帧策略虽然可以有效地避免缓冲区溢出，但却引起恢复视频质量的抖动，从而严重影响视频的主观质量。目前，H.264/AVC 编码器判决当前帧是否采用跳帧策略只是通过比较当前缓冲区占有度与预定义阈值二者之间的大小实现。本文统计了低码率情况下，H.264/AVC 提供的帧间预测模式对编码效率的贡献。在此基础上，设计了一种降低跳帧发生概率的算法。该算法根据缓冲区不同状态，实时地为当前编码单元更新编码参数，并根据帧间预测模式和编码效率关系的统计特性为当前宏块确定最终的编码模式。实验结果表明，与 H.264/AVC 编码器的方法相比，本文所设计的方法在极低码率下能够有效降低跳帧发生的概率，进而提高编码器的性能。

1.3.2 本文的组织结构

本文一共分为五章，各章内容安排如下：

第一章介绍了视频编码的必要性、H.264/AVC 视频编码标准中与码率控制相

关的关键编码技术以及研究码率控制的意义，并给出本论文的研究内容以及文章的组织结构。

第二章介绍了序列初始量化参数的重要性以及一些已有的确定初始量化参数的方法。通过实验推导了本文确定初始量化参数的方案。由于该方案的模型参数受图像复杂度的影响，所以本文研究了一种基于 Sobel 算子衡量图像复杂度的方法，并对该方法进行了实验验证。在此基础上，深入分析了由图像复杂度引入的初始量化参数偏差，并给出了求解初始量化参数的具体步骤。最后给出了本文方法的实验结果与分析。

第三章介绍了一些经典的码率控制方案并给予相应地分析。在此基础上，通过理论推导了一种基于码率-量化参数的码率控制方案，并对该方案进行了实验验证。然后介绍了本文方案确定 P 帧量化参数的具体过程，最后给出了相应的实验结果及分析。

第四章介绍了一些已有的跳帧策略，分析了帧间预测模式与编码效率的统计特性。在此基础上，设计了一种降低跳帧发生概率的算法。最后给出了相应的实验结果与分析。

第五章总结全文并展望了今后的研究工作。

第二章 H.264/AVC 初始 QP 的研究

2.1 引言

码率控制作为视频编码器的重要组成部分^[30], 为不同类型帧确定合适的量化参数, 以在资源受限情况下, 使解码端获得质量最优的恢复视频。因此, 码率控制是视频可靠传输以及恢复视频质量稳定的重要保证。目前, H.264/AVC^[31]的参考软件^[32]将 JVT-G012^[28]和 JVT-W042^[33]作为其推荐使用的码率控制方案。本章基于 JVT-G012 所提的码率控制机制, 对序列的初始量化参数进行了相关研究。

2.1.1 初始 QP 的重要性

序列初始 QP 是指编码视频序列第 1 帧使用的 QP, 其值影响着整个序列的编码质量。在给定目标比特率条件下, 如果编码序列第 1 帧所用的 QP 较小, 则用于第 1 帧的编码比特数就会增多, 这可能导致缓冲区溢出而发生跳帧。与此同时, 由于编码第 1 帧所用比特数较多, 导致当前 GOP 中剩余比特数减少, 从而引起后续帧质量下降; 如果编码序列第 1 帧所用的 QP 较大, 其引起的量化失真较大, 从而使得恢复帧质量降低。由于后续帧直接或者间接地以第 1 帧的恢复帧作为参考帧。因此, 第一帧的失真会传播到当前整个 GOP 中, 进而降低恢复视频质量。

为客观说明序列初始 QP 对整个视频序列编码质量的影响, 本文在给定目标比特率条件下, 采用 JVT-G012 码率控制方案测试了给定序列的一个 GOP(长度为 30 帧)在不同初始 QP 下的编码性能。表 2.1 给出了编码得到的跳帧帧数(Skip Frame Number, SFN)、编码质量和相对误差。其中, 编码质量用峰值信噪比(Peak Signal to Noise, PSNR)进行衡量。相对误差是指实际编码比特率与目标比特率的相对误差, 其表征码率控制的准确度, 计算见(2-1)。

表 2.1 不同初始 QP 对码率控制性能的影响(container.qcif, 目标比特率: 48kbps)

QP	SFN	PSNR (dB)	实际编码比特率 (kbps)	相对误差 (%)
38	0	33.91	48.94	2.0
36	0	34.63	52.04	8.4
34	0	34.78	49.82	3.8
32	0	35.75	51.27	6.8
30	0	35.68	48.27	0.6
28	3	36.00	50.30	4.8
26	7	35.47	50.62	5.5
24	11	34.52	51.10	6.5

$$\text{相对误差} = \frac{\text{实际码率} - \text{目标码率}}{\text{目标码率}} \times 100\% \quad \text{式(2-1)}$$

从表 2.1 可以看出, 对于给定的目标比特率, 采用不同初始 QP 进行编码得到的恢复视频质量、实际编码比特率以及跳帧帧数相差较大, 即不同初始 QP 使编码器的性能不同。采用合适的初始 QP 能够有效地提高编码器的编码效率。因此, 如何确定合适的初始 QP 是码率控制的关键技术之一。

2.1.2 已有的初始 QP 确定方法

理论上, 码率控制应根据视频序列的统计特性(图像复杂度、空间相关性、时间相关性等)和信道带宽, 为编码器确定能够获得最高编码效率的初始 QP, 即最佳初始 QP。但是, 这需要提取序列每一帧的信息并进行大量分析计算, 大大增加了编码器的复杂度。

在 H.264/AVC 视频编码标准推荐使用的码率控制方案中, 序列初始 QP 可以根据经验直接在编码器相应的配置文件中预先设定。在没有预先设定时, 码率控制首先根据目标比特率、帧率以及图像大小计算每个像素点的平均编码比特数(bits per pixel, bpp)。然后, 根据图像大小确定 bpp 与 QP 的对应表。最后, 由该对应表计算序列初始 QP。由此可见, 对于目标比特率、帧率、图像大小相同而图像复杂度不同的视频序列, 采用 H.264/AVC 推荐使用的码率控制为其确定的初始 QP 相同。

事实上, 不同图像复杂度的序列, 选择的初始 QP 应该不同。当测试条件完全相同时, 与低图像复杂度序列相比, 高图像复杂度序列的最佳初始 QP 应相对较大。否则, 将导致缓冲区溢出、恢复视频质量波动等。因此, 在确定初始 QP 时, 若不考虑图像复杂度的影响, 势必会导致码率控制性能的下降, 进而影响编码器的编码效率。文献[34]指出基于方差、梯度、DCT 和边缘算子这 4 种计算图像复杂度方法中, 采用基于梯度算法衡量图像复杂度的方法较为准确。为此, 文献[35-39]通过大量实验, 建立了梯度确定的图像复杂度、编码码率以及 QP 之间的关系, 以此确定序列初始 QP。由于当前帧像素点的梯度值之和并不能准确反映当前帧的图像复杂度, 并且无法建立该图像复杂度与初始 QP 之间的直接关系^[38], 从而导致确定的初始 QP 不是最佳初始 QP, 进而影响编码器的性能。

2.2 自适应的初始 QP 确定方法

2.2.1 确定初始 QP 的方案

为了研究 IPPP 预测结构中, 最佳初始 QP、每个像素点的平均编码比特数 bpp

和图像复杂度之间的关系, 本文使用 H.264/AVC 视频编码标准推荐的码率控制方案, 在不同目标比特率下, 采用不同 QP 作为序列初始 QP 进行编码, 并记录使码率控制性能达到最优的初始 QP 值。判断码率控制性能优劣的规则如下^[40]:

- 1) 实际编码比特率与目标编码比特率之间的相对误差越小越好, 即码率控制的准确度越高越好;
- 2) 缓冲区不发生溢出, 即既不发生上溢也不发生下溢;
- 3) 恢复视频的 PSNR 越高越好;
- 4) 恢复视频质量的波动越小越好。

以表 2.1 为例, 具体说明最佳初始 QP 的确定方法。由表 2.1 可知, 当初始 QP 设置为 38 时, 虽然码率控制的准确度较高, 但恢复视频质量较低; 当初始 QP 分别设置为 36、34 和 24 时, 编码器能够获得较高的恢复视频质量, 但其码率控制准确度较低。当初始 QP 分别设置为 32 和 26 时, 获得的恢复视频质量高于初始 QP 为 36、34 和 24 的情况, 但其码率控制准确度较低。当初始 QP 分别设置为 30 和 28 时, 码率控制得到的编码准确度以及恢复视频的质量相对较高。但是, 当初始 QP 为 28 时, 编码过程中发生了跳帧, 导致恢复视频质量的波动较大, 如图 2.1 所示。所以, 在测试给定的目标比特率条件下, 该测试序列最佳初始 QP 为 30。

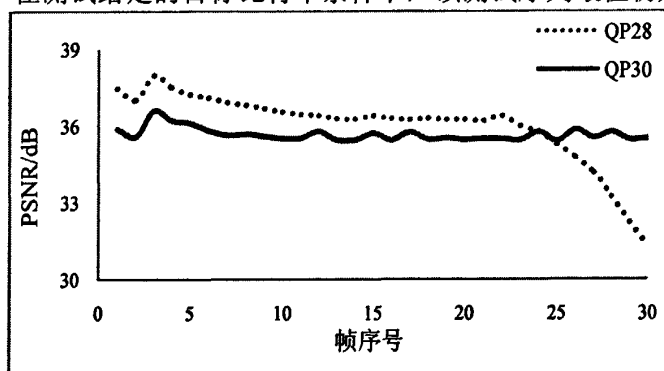


图 2.1 序列中各帧编码质量的波动(container 序列, 目标码率: 48kbps)

在给定目标码率条件下, 根据上述确定初始 QP 的方法, 为不同测试序列确定最佳初始 QP。以每个像素点平均编码比特数 bpp 的对数为横轴, 以所确定的最佳初始 QP 为纵轴, 将不同序列不同目标码率的数据点绘制在同一个坐标系中, 并使用线性回归方法对这些数据点进行分析, 便可以得到图 2.2 所示的最佳初始 QP、bpp 和图像复杂度之间的关系曲线。由该图可以看出, 对于特定序列的最佳初始 QP 和 bpp 之间的关系, 可用如下线性关系式进行描述,

$$Q_p = A_0 + A_1 \cdot \lg(b) \quad \text{式(2-2)}$$

其中, Q_p 和 b 分别为 QP 和 bpp 的数值, A_0 和 A_1 为模型参数。

图 2.2 表明在相同测试条件下, 不同序列最佳初始 QP 之间存在着偏差。从而进一步说明不同序列采用相同的初始 QP 必然会影响某些序列的编码质量。由该图还可以看出, 不同序列 $\lg(b)$ 和 Q_p 的直线经过水平方向的平移可近似重合, 这表明不同序列 A_1 的数值近似相等, 该值可根据拟合曲线进行确定。在测试条件完全相同的情况下, 不同序列 A_0 值不同, 这表明 A_0 是与序列特性相关的一个量, 即与图像复杂度有关。不同序列 A_0 之间的偏差表征了由图像复杂度之差引入的初始 QP 偏差。因此, 如果能够提取衡量图像复杂度的因子, 并且利用该因子推导出 A_0 的数值, 就可以利用式(2-2)所提的方法确定当前序列最佳初始 QP, 从而使码率控制的性能达到最优。

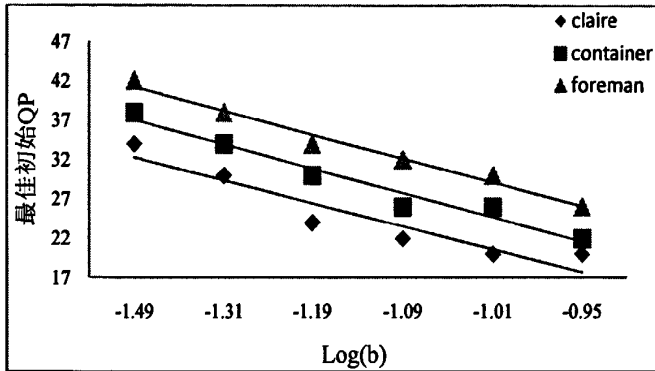


图 2.2 最佳初始 QP、bpp 与图像复杂度之间的关系

2.2.2 基于纹理块的图像复杂度

文献[35-39]利用梯度算子计算当前图像每一个像素点的梯度值, 最终用归一化梯度值(Unitary Grads Value, UGV)衡量图像复杂度。归一化梯度值的计算公式为,

$$U_{gv} = \frac{1}{M \times N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [|I_{i,j} - I_{i,j+1}| + |I_{i,j} - I_{i+1,j}|] \quad \text{式(2-3)}$$

其中, U_{gv} 为 UGV 的值, M 和 N 分别表示图像的宽和高, $I_{i,j}$ 表示坐标为 (i, j) 的像素点的亮度分量灰度值。 U_{gv} 值越大, 图像复杂度越高。

平坦区域占整个图像面积较大的视频序列存在较多空间冗余信息, 对其进行编码, 能够有效去除这些冗余信息, 从而可以用较少编码比特数获得较高恢复视频质量, 这表明这类视频序列的图像复杂度较低。反之, 图像复杂区域占整个图像面积较大的视频序列空间冗余信息较少, 对其进行编码, 如果要获得与低复杂度视频序列相同的恢复视频质量, 则需要花费更多编码比特数, 这表明这类视频序列的图像复杂度较高。在图 2.3 给出的图像中, 可以看出图 2.3(a)图像复杂度最低, 图 2.3(b)次之, 图 2.3(c)图像复杂度最高。为了客观说明这一点, 在相同测试

条件下, 对图 2.3 给出的三幅图像分别进行编码, 表 2.2 给出了相应的测试结果。从表中可以看出, *claire* 图像用最少的编码比特数获得了最高的 PSNR, 所以其图像复杂度最低。而 *silent* 编码消耗的比特数最多, 但是获得的恢复视频质量最差。这说明在给出的三幅图像中, *silent* 的图像复杂度最高。测试结果验证了平坦区域面积较大的图像复杂度较低, 反之, 则其图像复杂度较高。也就是说, 纹理区域面积较大的图像, 其图像复杂度较高, 反之, 则图像复杂度较低。由此可见, 纹理区域面积能够反映图像复杂度, 即可以利用纹理区域面积占整个图像面积的比例来衡量图像复杂度。

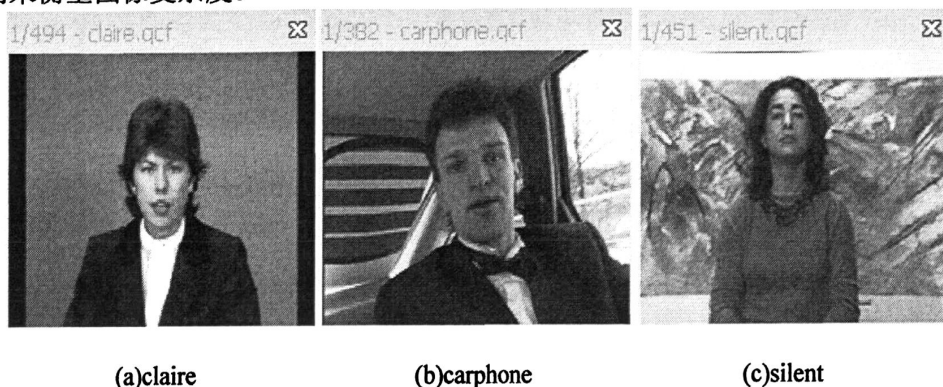


图 2.3 图像复杂度示意图

表 2.2 图 2.3 中图像编码结果

测试序列	PSNR (dB)	实际编码比特数 (bits)
<i>claire</i>	39.449	11064
<i>carphone</i>	36.462	18576
<i>silent</i>	34.991	21744

基于上述分析, 本文利用纹理区域面积(Texture Area, TA)占有度因子(Texture Area Factor, TAF)衡量图像复杂度。纹理区域面积占有度因子的定义为,

$$T_{af} = \frac{T_a}{S_a + T_a} \quad \text{式(2-4)}$$

其中, T_{af} 、 T_a 分别为 TAF、TA 的数值。 S_a 为图像平坦区域面积(Smooth Area, SA)的值。 T_{af} 值越大, 意味着纹理区域越大, 图像复杂度越高; 反之, 则意味着平坦区域越大, 图像复杂度越低。然而, 获得 T_a 和 S_a 的精确值比较困难。为简单有效地计算纹理区域面积占有度因子, 本文采用 H.264/AVC 视频编码标准所支持的最大编码块大小对图像进行分割, 并将位于平坦区域内的块称作平坦块(Smooth Block, SB), 其它块则称之为纹理块(Texture Block, TB)。用纹理块和平坦块的个数分别近似表示 T_a 和 S_a 的值。这样就可以用纹理块的个数(Number of Texture Block,

NTB)与图像总块数(Number of Blocks in Frame, NBF)之间的比值近似表示当前图像纹理区域面积占有度因子, 即:

$$S_{af} \approx \frac{N_{tb}}{N_{bf}} \quad \text{式(2-5)}$$

其中, N_{tb} 、 N_{bf} 分别为 NTB、NBF 的值, S_{af} 为 T_{af} 的近似值。

在平坦块中, 图像亮度分量的灰度值近似相同, 边缘较为平缓, 高频分量较少。而纹理丰富的块, 图像亮度分量的灰度值变化较大, 有较为锐化的边缘, 高频分量较多。基于此, 我们首先通过 Sobel 算子提取当前块每一个像素点的水平和垂直方向的边缘分量。然后, 通过比较该边缘能量与预定义阈值的大小判断该像素点是在纹理变化平缓区域还是在纹理变化锐化区域。最后, 统计当前块中位于纹理区域的像素点个数, 并与预定义的阈值进行比较, 以判定当前块是纹理块还是平坦块。判断当前块是纹理块还是平坦块的具体过程如下:

- 1) 用 Sobel 算子确定当前块每一个像素点的边缘能量。像素点坐标为 (m, n) 的边缘能量 $F(m, n)$ 的计算公式为,

$$S_{ver} = \begin{bmatrix} -1 & -2 & -1 \\ 0 & 0 & 0 \\ 1 & 2 & 1 \end{bmatrix} S_{hor} = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix} \quad \text{式(2-6)}$$

$$H = G(m, n) * S_{hor} \quad V = G(m, n) * S_{ver} \quad \text{式(2-7)}$$

$$F(m, n) = H^2 + V^2 \quad \text{式(2-8)}$$

其中, S_{ver} 、 S_{hor} 分别表示垂直和水平方向的边缘检测算子, H 、 V 分别表示像素点坐标为 (m, n) 的水平和垂直方向的边缘能量, $G(m, n)$ 为像素点亮度值, 式(2-7)中 “*” 表示卷积运算符号^[41]。

- 2) 判断当前像素点所在的区域。当前像素点边缘能量越大, 表明当前像素点与相邻像素点的亮度分量灰度值之差越大; 反之, 则表明当前像素点和相邻像素点的亮度分量灰度值之差越小。所以, 可以通过比较当前像素点的边缘能量与预定义边缘能量阈值来判断当前像素点是位于亮度值变化较大的锐化区域还是亮度值变化较小的平缓区域。
- 3) 判断当前块是否为纹理块。统计当前块位于锐化区域的像素点个数 (Number of Large Energy Pixels, NLEP), 即边缘能量超过预定义边缘能量阈值的像素点个数, 并以该值作为判断当前块是否是纹理块的依据。如果 NLEP 较大, 说明当前块很可能位于亮度分量灰度值变化频繁的纹理区域; 如果 NLEP 较小, 说明当前块可能位于平坦区域与纹理区域的边缘; 如果 NLEP 极小, 则说明当前块完全位于平坦区域。这里, 我们设定当 NLEP

小于设定的阈值时,认为当前块位于平坦区域,为平坦块。否则,即为纹理块。

利用上述方法,确定当前图像所有分块的类型(纹理块和平坦块)之后,就可以得到纹理块和平坦块个数,再利用式(2-5)就能很方便地计算图像复杂度因子 S_{af} 。

为验证本文算法的有效性,在相同测试条件下,采用不同公共测试序列进行编码,并按照实际编码比特数(Encoded Bits, EB)以及恢复帧的 PSNR 对测试序列的复杂度进行排序,得到实际图像复杂度的排序结果。按照纹理区域占有度因子 S_{af} 以及归一化梯度 U_{gv} 大小,分别对测试序列进行排序,得到了本文算法和文献[35-39]采用算法的图像复杂度排序结果,如表 2.3 所示。可以看出,本文算法与实际编码图像复杂度的排序结果相同,而文献[35-39]所采用的算法得到的复杂度排序结果与实际编码图像复杂度的排序结果略有差异。从而,证明了本文设计的衡量图像复杂度算法的有效性。

表 2.3 所提衡量复杂度的算法与文献[35-39]的比较

	missa	claire	highway	carphone	news	foreman	silent
S_{af}	0.52	0.61	0.85	0.90	0.91	0.95	0.98
U_{gv}	4.99	7.39	7.63	13.49	17.46	14.84	14.87
EB(bits)	9088	11064	10432	18576	23680	20368	21744
PSNR(dB)	39.55	39.45	37.74	36.46	36.15	35.45	34.99
实际编码复杂度排序	1	2	3	4	5	6	7
本文方法复杂度排序	1	2	3	4	5	6	7
文献[35-39]复杂度排序	1	2	3	4	7	5	6

2.2.3 确定初始 QP 偏差

由 2.2.1 小节可知,在给定目标比特率条件下,只要求得 A_0 ,就可以利用式(2-2)确定序列最佳初始 QP。而 A_0 是与图像复杂度相关的一个量,不同序列图像复杂度不同, A_0 的值也不同,从而导致序列最佳初始 QP 不同,如表 2.4 所示。表中采用上一小节介绍的纹理区域面积占有度因子 S_{af} 衡量图像复杂度。可以看出,最佳初始 QP 随着图像复杂度的变化而变化,并且随着图像复杂度的增加(S_{af} 数值越大,图像复杂度越高)呈现增长的趋势。

表 2.4 不同序列图像复杂度与最佳初始 QP 之间的关系

	miss	claire	container	highway	carphone	news	foreman	silent
S_{af}	0.52	0.61	0.79	0.85	0.90	0.91	0.95	0.98
最佳初始 QP	32	34	38	38	38	38	42	44

在给定条件下,为了确定由图像复杂度 S_{af} 引入的初始量化参数偏差,将 *claire* 序列作为基准序列,并记录其图像复杂度(以 S_{ref} 表示)以及 A_0 (以 A_{ref} 表示)。对于任何一个序列,式(2-2)所示方案中的 A_0 可表示为,

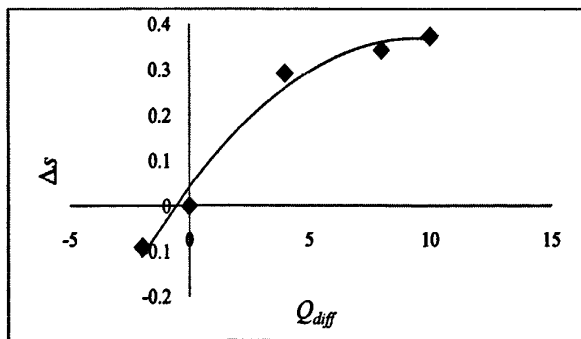
$$A_0 = A_{ref} + Q_{diff} \quad \text{式(2-9)}$$

其中, Q_{diff} 表示由当前序列与基准序列图像复杂度的不同而导致的当前序列与基准序列的初始量化参数偏差。为求得 Q_{diff} , 以测试得到的最佳初始 QP 的差值 Q_{diff} (可通过各个序列的最佳初始 QP 与基准序列的最佳初始 QP 计算得到)为横轴,以相应的图像复杂度因子之差 Δs (可通过各个序列的图像复杂度因子与基准序列的图像复杂度因子计算得到)为纵轴,将具有不同图像复杂度的数据点描绘在同一个坐标系中,如图 2.4 所示。可以看出 Q_{diff} 与 Δs 之间的关系近似抛物线的一部分,即可用下式表示为,

$$\Delta s = aQ_{diff}^2 + bQ_{diff} + c \quad \text{式(2-10)}$$

其中, a 、 b 和 c 为模型参数,可以根据拟合曲线求得。通过式(2-10)求得的 Q_{diff} 数值可能不唯一。由于图 2.4 只是开口向下抛物线的左半部分,所以只能取式(2-10)两个解中较小的一个解,即,

$$Q_{diff} = \frac{-b - \sqrt{b^2 - 4a(c - \Delta s)}}{2a} \quad \text{式(2-11)}$$

图 2.4 最佳初始 QP 之差 Q_{diff} 与图像复杂度之差 Δs 的关系

通过式(2-11)即可求得给定序列与基准序列的最佳初始 QP 之间的偏差 Q_{diff} , 再根据式(2-9)就可以确定式(2-2)中的模型参数 A_0 。

2.2.4 自适应初始 QP 的确定方法

本文将图像复杂度作为考虑因素之一，探索性地研究了一种确定序列初始量化参数的方法。在给定目标比特率条件下，为当前序列确定最佳初始 QP 的具体步骤如下：

- 1) 根据 2.2.2 小节设计的方法计算当前序列的图像复杂度。
- 2) 计算当前序列与基准序列的图像复杂度之差 Δs 。
- 3) 根据式(2-11)确定当前序列与基准序列最佳初始 QP 之间的偏差 Q_{diff} ，再利用式(2-9)确定式(2-2)中的模型参数 A_0 。
- 4) 根据 2.2.1 小节给出的确定序列初始 QP 的方案，为当前序列确定初始 QP。
- 5) 按照 H.264/AVC 视频编码标准推荐使用的码率控制方案，检验所确定的初始 QP 是否在有效范围内，并为序列确定最终的初始 QP。

2.3 实验与分析

将本文确定序列初始 QP 的方法应用于 JVT-G012 的码率控制方案(标记为“JVT-G012+Q_{initial}”)，并与原 JVT-G012 确定初始 QP 的码率控制方案(标记为“JVT-G012”)进行性能比较。采用 Bjøntegaard 方法^[42]比较改进方案与参考方案之间的性能差异，并利用 VCEG-AE07^[43]给出的工具计算 Δ PSNR 的值。 Δ PSNR 表示在相同编码比特率条件下，与参考方案相比，改进方案平均编码质量(以 PSNR 进行衡量)的提高(Δ PSNR 的值为正)或者下降(Δ PSNR 的值为负)的程度。实验使用的预测编码结构为 IPPP，编码帧率为 30 帧/秒。

采用 qcif 格式的标准测试序列对“JVT-G012+Q_{initial}”与“JVT-G012”两种方案的性能进行测试。表 2.5 给出了比较结果。从表中可以看出，“JVT-G012+Q_{initial}”的相对误差较“JVT-G012”平均提高了 1.22%。但是，在相同的编码比特率条件下，“JVT-G012+Q_{initial}”大多数情况下均能够获得较高的恢复视频质量。与参考方案相比，本文研究方案的恢复视频质量平均可提高 0.69 dB，最高可达 1.39dB。

表 2.5 与“JVT-G012”的性能比较

测试 序列 (qcif)	码率 (kbps)	JVT-G012			JVT-G012+Q _{initial}			Δ PSNR (dB)
		Rate (kbps)	PSNR (dB)	相对误 差(%)	Rate (kbps)	PSNR (dB)	相对误 差(%)	
claire	48	39.86	48.07	-16.96	41.05	49.74	-14.48	1.32
	60	61.78	40.72	2.97	60.66	42.03	1.10	
	72	72.94	41.81	1.31	72.35	42.95	0.49	

	84	84.53	43.61	0.63	85.92	43.72	2.29	
	48	50.27	32.82	4.73	50.69	32.60	5.60	
carphone	60	61.08	33.52	1.80	62.14	33.35	3.57	-0.09
	72	72.85	34.21	1.18	75.34	34.64	4.64	
	84	86.38	35.19	2.83	86.58	35.07	3.07	
	48	48.42	34.85	0.88	50.30	36.00	4.79	
container	60	60.82	35.55	1.37	61.58	37.11	2.63	1.14
	72	72.08	36.2	0.11	73.37	37.87	1.90	
	84	85.13	38.32	1.35	85.62	38.59	1.93	
	48	49.48	32.09	3.08	50.62	32.44	5.46	
foreman	60	61.01	32.69	1.68	62.52	33.38	4.20	0.39
	72	72.56	33.45	0.78	74.90	34.15	4.03	
	84	87.22	34.73	3.83	87.14	34.66	3.74	
	48	48.46	35.84	0.96	49.19	36.61	2.48	
highway	60	59.96	36.52	-0.07	61.79	37.28	2.98	0.59
	72	72.01	37.07	0.01	72.32	37.82	0.44	
	84	84.33	38.21	0.39	84.64	38.28	0.76	
	48	50.07	40.72	4.31	48.38	41.61	0.79	
missa	60	60.93	41.33	1.55	61.62	42.58	2.70	0.91
	72	73.05	41.93	1.46	72.70	43.06	0.97	
	84	82.84	43.27	-1.38	84.59	43.31	0.70	
	48	49.54	32.83	3.21	49.79	34.21	3.73	
news	60	60.44	34.25	0.73	61.87	35.71	3.12	1.39
	72	74.03	34.96	2.82	74.39	36.99	3.32	
	84	85.62	37.93	1.93	85.62	37.93	1.93	
	48	49.93	31.72	4.02	52.78	31.57	9.96	
silent	60	61.86	33.74	3.10	61.93	34.19	3.22	-0.12
	72	72.62	34.08	0.86	75.06	35.4	4.25	
	84	88.26	36.62	5.07	86.96	36.19	3.52	
	average	—	—	—	1.27	—	—	

利用 cif 格式的标准测试序列对“JVT-G012+Q_{initial}”和“JVT-G012”两种方案进行性能测试。表 2.6 给出了测试结果。从表中可以看出，“JVT-G012+Q_{initial}”的相对误差较“JVT-G012”平均提高了 0.52%。但是，在相同的编码比特率条件下，“JVT-G012+Q_{initial}”可以获得更高的恢复视频质量。与参考方案相比，本文研究

方案的恢复视频质量平均可提高 0.29 dB，最高可达 0.81dB。

表 2.6 与“JVT-G012”的性能比较

测试 序列 (cif)	目标 码率 (kbps)	JVT-G012			JVT-G012+Q _{initial}			Δ PSNR (dB)
		RATE (kbps)	PSNR (dB)	相对误 差(%)	RATE (kbps)	PSNR (dB)	相对误 差(%)	
bridge-far	64	70.40	36.85	10.00	72.23	36.48	12.86	0.01
	84	83.82	36.98	-0.21	83.82	36.98	-0.21	
	108	112.77	37.19	4.42	111.18	37.45	2.94	
	132	131.06	37.34	-0.71	134.97	37.73	2.25	
claire	64	67.00	37.76	4.69	65.33	37.11	2.08	0.10
	84	84.98	38.65	1.17	84.98	38.65	1.17	
	108	111.07	39.88	2.84	110.44	40.22	2.26	
	132	135.27	40.82	2.48	135.71	41.39	2.81	
foreman	64	70.14	23.35	9.59	71.13	25.14	11.14	0.58
	84	92.15	25.79	9.70	94.85	26.34	12.92	
	108	114.89	28.30	6.38	114.43	28.38	5.95	
	132	138.21	31.04	4.70	137.89	29.97	4.46	
hall	64	68.68	31.66	7.31	68.97	32.26	7.77	0.23
	84	90.30	34.50	7.50	89.05	34.14	6.01	
	108	113.23	35.33	4.84	116.10	35.81	7.50	
	132	137.02	35.78	3.80	140.01	36.33	6.07	
highway	64	68.57	34.38	7.14	68.01	35.08	6.27	0.13
	84	87.69	35.99	4.39	87.69	35.99	4.39	
	108	108.63	36.42	0.58	112.46	36.53	4.13	
	132	133.19	36.86	0.90	138.29	36.38	4.77	
moth_dotr	64	70.80	36.20	10.63	67.74	36.14	5.84	0.27
	84	89.37	37.42	6.39	93.41	37.58	11.20	
	108	117.62	38.47	8.91	114.00	38.75	5.56	
	132	145.10	39.11	9.92	144.28	39.83	9.30	
news	64	67.95	30.21	6.17	67.00	30.97	4.69	0.56
	84	86.48	32.80	2.95	87.99	32.88	4.75	
	108	111.74	33.88	3.46	111.01	34.56	2.79	
	132	135.59	34.62	2.72	134.75	35.74	2.08	

	64	68.14	25.77	6.47	65.43	27.92	2.23	
waterfall	84	88.27	26.56	5.08	90.98	26.90	8.31	0.81
	108	121.44	27.22	12.44	114.61	28.48	6.12	
	132	135.48	29.07	2.64	138.52	28.25	4.94	
	64	68.67	28.25	7.30	69.37	29.28	8.39	
caphone	84	88.90	30.24	5.83	91.60	30.81	9.05	0.21
	108	114.20	32.22	5.74	114.41	32.02	5.94	
	132	135.15	33.30	2.39	142.72	33.23	8.12	
	64	70.40	36.85	10.00	72.23	36.48	12.86	
bridge -close	84	83.82	36.98	-0.21	83.82	36.98	-0.21	0.01
	108	112.77	37.19	4.42	111.18	37.45	2.94	
	132	131.06	37.34	-0.71	134.97	37.73	2.25	
	average	—	—	—	5.10	—	—	

综上,与“JVT-G012”方案相比,本文所研究的算法可以在一定程度上改善码率控制的性能,进而提高编码器的编码效率。

2.4 小结

确定序列初始量化参数是码率控制的关键技术之一。本章首先通过实验详细分析了初始量化参数对整个序列编码质量的影响,同时介绍了一些确定初始量化参数的方法并给出了相应分析。然后,本章通过实验推导了一种确定初始量化参数的方案。通过对该方案进行分析可知,不同序列图像复杂度的差异导致最佳初始量化参数出现偏差。为了定量求解该偏差,本章设计了一种基于 Sobel 算子的图像复杂度衡量方法,并对其准确性进行了实验验证。在此基础上,深入研究了不同序列图像复杂度与最佳初始量化参数偏差之间的关系,并给出了求解初始量化参数的具体步骤。

将所设计的方案应用于 JVT-G012 码率控制方案中,并与“JVT-G012”方案相比较。实验结果表明,在相同的编码比特率下,对 qcif 和 cif 两种标准测试序列,所提方案获得的恢复视频质量可分别平均提高 0.69dB、0.29dB。可见,所设计的方案能够在一定程度上改善码率控制的性能,进而提高编码器的编码效率。

第三章 低码率线性码率控制方案

3.1 引言

码率控制除了为每一个 GOP 的 I 帧确定合适的 QP 之外, 还需要为 P 帧选择合适的 QP, 以在缓冲区大小、信道带宽等资源受限的情况下, 使解码端获得最优的恢复视频质量。而在 IPPP 编码预测结构中, 大部分帧均为 P 帧。因此, 如何确定 P 帧的 QP 是码率控制的关键问题之一。

目前, 大多数码率控制是通过建立码率(Rate, R)-量化步长(Quantization Step, Q_{step})模型^[44-47](R- Q_{step})为编码器确定合适的编码参数, 如 JVT-G012^[28]和 JVT-W042^[33]所推荐使用的二次模型^[45]。在 H.264/AVC 视频编码标准中, Q_{step} 不能直接作为编码参数, 需要根据量化步长和量化参数之间的对应关系, 将 Q_{step} 转换为相应的 QP 之后才能对视频序列进行编码。然而, Q_{step} 转换为 QP 的过程中存在不可避免的舍入误差, 使得通过 R- Q_{step} 模型确定的 QP 所对应的量化步长与计算得到的量化步长之间存在差异, 从而导致码率控制性能下降。性能下降主要表现为实际编码比特率与给定目标比特率之间存在较大的偏差, 编码器缓冲区产生较大的波动, 恢复视频质量产生较大的波动, 恢复视频 PSNR 下降等。为解决这些问题, 本章通过理论推导以及实验验证研究了一种简单有效的视频编码码率控制方案。通过该方案可以直接为编码器确定量化参数, 从而, 避免了由量化步长转化为量化参数的过程所引入的舍入误差, 提高了码率控制的性能。

本章内容安排如下: 3.2 小节介绍了一些已有的码率控制模型, 并分析了低码率情况下确定 QP 存在的问题; 3.3 小节通过理论推导与实验验证, 研究了一种线性码率控制方案, 并给出了确定 QP 的具体方法; 3.4 小节给出了实验条件与实验结果, 并对实验结果进行了分析; 3.5 小节是本章的小结。

3.2 已有码率控制模型

由于码率控制在视频编码器的重要性, 近几年来, 许多学者对其进行了研究。文献[48]指出量化步长与量化系数为零的个数有关。当量化步长增大时, 为零的量化系数个数就会增多, 并且二者之间存在着特定的线性关系。基于此, 提出了用量化系数零值的比例 ρ 来计算量化步长 Q_{step} , 即,

$$R(\rho) = \theta \times (1 - \rho) \quad \text{式(3-1)}$$

其中, θ 为常数。当目标比特率极低时, 该模型的码率控制性能会有所降低。

文献[44]指出当信源服从拉普拉斯分布时, 对其率失真函数进行泰勒级数展开,

并用量化步长作为衡量失真的标准,提出了二次 R- Q_{step} 模型,即,

$$N_{bc} = \frac{\alpha_1}{Q_{step}} + \frac{\alpha_2}{Q_{step}^2} \quad \text{式(3-2)}$$

其中, N_{bc} 表示编码比特数(The Number Of Bits Encoded, N_{bc}), α_1 和 α_2 为模型系数,在编码过程中可通过线性回归方法进行更新。该模型在 MPEG-4 码率控制算法中得到了有效应用^[44]。

Lee 等指出文献[44]所提模型中的编码比特数 N_{bc} 应该区分非纹理信息和纹理信息的编码比特数(The Number Of Texture Bits Encoded, N_{tbe}), 并且该模型没有考虑图像内容复杂度对量化参数的影响。所以,将原始图像与预测图像差值信号的平均绝对值 MAD 作为衡量图像内容复杂度的度量。从而,利用 N_{tbe} 、MAD 以及 Q_{step} 之间的关系提出了如下模型^[49],

$$\frac{N_{tbe}}{MAD} = \frac{\beta_1}{Q_{step}} + \frac{\beta_2}{Q_{step}^2} \quad \text{式(3-3)}$$

其中, β_1 和 β_2 为模型系数,在编码过程中通过线性回归方法进行更新。这种模型由于复杂度相对较低,码率控制性能相对较好,已广泛应用于现有的码率控制方案中,如 MPEG-4 的码率控制方案^[49-50], H.264/AVC 的码率控制方案^[29]等。

通过对率失真函数进行泰勒级数展开并保留常数项,文献[47]提出了修正的二次 R- Q_{step} 模型,如式(3-4)所示。其中, γ_0 、 γ_1 和 γ_2 为模型系数。与文献[44]所提模型相比,修正的二次 R- Q_{step} 模型的码率控制性能有所提高。

$$\frac{N_{tbe}}{MAD} = \frac{\gamma_0}{MAD} + \frac{\gamma_1}{Q_{step}} + \frac{\gamma_2}{Q_{step}^2} \quad \text{式(3-4)}$$

不难看出,上述给出的较为经典的码率控制算法均通过码率-量化步长(R- Q_{step})模型确定量化步长。由于量化步长不能直接作为 H.264/AVC 视频编码器的编码参数,只有根据量化步长和量化参数之间的对应关系,将量化步长转换为相应的量化参数后,才能对视频进行编码。表 3.1 给出了 H.264/AVC 编码器中,量化步长 Q_{step} 与量化参数 QP 之间的对应关系^[20]。由表可以看出, Q_{step} 与 QP 之间的转换过程中存在着不可避免的舍入误差。例如,当通过 R- Q_{step} 模型确定的 Q_{step} 为 14.5 时,为编码器确定的 QP 为 27,而 QP 为 27 所对应的 Q_{step} 为 14;当通过 R- Q_{step} 模型确定的 Q_{step} 为 169 时,编码器确定的 QP 为 49,而该 QP 所对应的 Q_{step} 为 176。这表明采用 R- Q_{step} 模型进行码率控制使得最终确定的 QP 所对应的量化步长与计算得到的量化步长之间存在舍入误差,并且该舍入误差随着 Q_{step} 数值的增大而增大。当给定的目标比特率较低时,通过 R- Q_{step} 模型确定的量化步长较大,导致由量化步长转换为量化参数引入的舍入误差较大,从而影响码率控制性能。

表 3.1 H.264/AVC 编码器的量化步长 Q_{step} 和量化参数 QP 之间的对应关系

QP	Q_{step}	QP	Q_{step}	QP	Q_{step}	QP	Q_{step}
0	0.6250	13	2.75	26	13	39	56
1	0.6875	14	3.25	27	14	40	64
2	0.8125	15	3.50	28	16	41	72
3	0.8750	16	4.00	29	18	42	80
4	1.0000	17	4.50	30	20	43	88
5	1.1250	18	5.00	31	22	44	104
6	1.2500	19	5.50	32	26	45	112
7	1.3750	20	6.50	33	28	46	128
8	1.6250	21	7.00	34	32	47	144
9	1.7500	22	8.00	35	36	48	160
10	2.0000	23	9.00	36	40	49	176
11	2.2500	24	10.00	37	44	50	208
12	2.5000	25	11.00	38	52	51	224

3.3 线性码率控制方案

3.3.1 R-QP 方案

文献[51]指出编码失真 D 与编码码率 R_T 之间存在式(3-5)所示的关系,

$$D = \delta^2 R_T^{1-2r} \quad \text{式(3-5)}$$

其中, 失真 D 以编码视频的均方误差(Mean Square Error, MSE)进行度量, δ 和 r 为模型参数。对式(3-5)左右两边同时取对数可得,

$$\lg D = (1-2r) \cdot \lg R_T + 2 \cdot \lg \delta \quad \text{式(3-6)}$$

恢复视频的峰值信噪比 PSNR 与 D 之间的关系为^[52],

$$P = 10 \times \lg \left(\frac{255^2}{D} \right) \quad \text{式(3-7)}$$

其中, P 为恢复视频的 PSNR 数值。对式(3-7)进行对数运算并整理可得,

$$\lg D = 2 \times \lg 255 - 0.1 \cdot P \quad \text{式(3-8)}$$

由式(3-6)、式(3-8)可得,

$$2 \times \lg 255 - 0.1 \cdot P = (1-2r) \cdot \lg R_T + 2 \cdot \lg \delta \quad \text{式(3-9)}$$

进而, 可以推导出式(3-10)成立,

$$P = a \lg R_T + b \quad \text{式(3-10)}$$

其中, a 和 b 为模型参数。文献[53]中指出, P 与 QP 之间存在着线性关系, 即,

$$P = c \cdot Q_P + d \quad \text{式(3-11)}$$

其中, c 和 d 为模型参数, Q_P 表示量化参数 QP 的大小。由式(3-10)、式(3-11)可得,

$$\lg R_T = \alpha \cdot Q_P + \beta \quad \text{式(3-12)}$$

其中, α 和 β 为本文设计方案的参数。式(3-12)表明 $\lg R_T$ 与 Q_P 之间存在线性关系。

3.3.2 方案验证

为验证式(3-12)所示方案, 本文在 H.264/AVC 的参考软件 JM15.1 平台上, 采用不同的量化参数, 对 foreman、bigships、keiba、soccer、bqsquare 和 city_corr 六个标准测试序列进行编码, 并统计编码比特率的对数与量化参数之间的关系。测试结果如图 3.1 所示。由该图可以看出编码比特率的对数与量化参数之间确实存在线性关系, 符合式(3-12)设计的方案。从而, 验证了本文所设计方案的正确性。

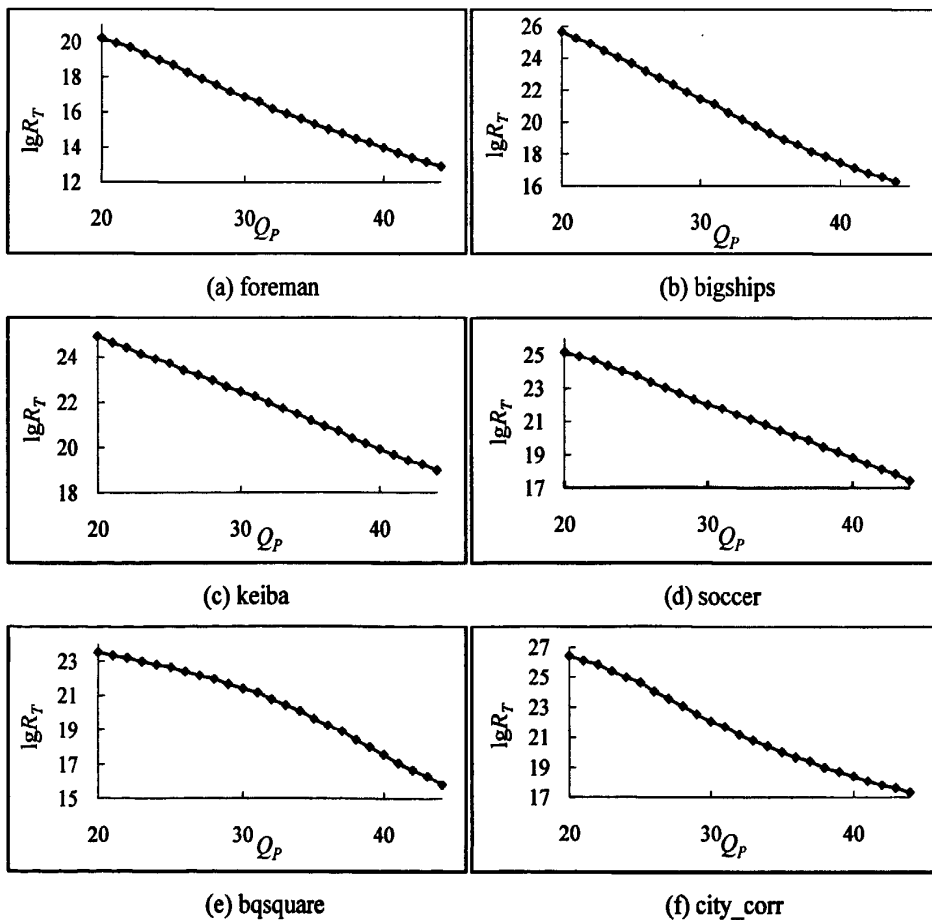


图 3.1 不同序列编码比特率的对数 $\lg R_T$ 与量化参数 Q_P 之间的关系

为了直观地描述 $\lg R_T$ 与 Q_P 之间的线性关系, 表 3.2 统计了上述测试序列 $\lg R_T$ 与 Q_P 之间的相关系数。由表 3.2 可知, $\lg R_T$ 与 Q_P 相关系数的绝对值近似为 1, 这表明 $\lg R_T$ 与 Q_P 之间存在线性关系, 进一步说明了本章设计的 R-QP 方案的正确性。

表 3.2 $\lg R_T$ 与 Q_P 之间的相关系数统计结果

序列	相关系数	序列	相关系数
foreman	-0.9985	Keiba	-0.9998
bqsquare	-0.9882	city_corr	-0.9958
soccer	-0.9998	bigships	-0.9986

3.3.3 QP 的计算

本文码率控制方案利用相邻帧之间的时域相关性为当前 P 帧确定 QP。用 $R_{tbe}(i, j)$ 表示第 i 个 GOP 中第 j 帧纹理信息的目标编码比特数(The Number of Target Texture Bits Encoded, R_{tbe}), $R_{atbe}(i, j)$ 表示第 i 个 GOP 中第 j 帧纹理信息的实际编码比特数(The Number Of Actual Texture Bits Encoded, R_{atbe})。用 $Q_P(i, j)$ 表示第 i 个 GOP 中第 j 帧的量化参数。每个 GOP 中前两帧的 QP 按照 H.264/AVC 视频编码标准推荐使用的码率控制方案进行确定。对第 i 个 GOP 中的第 $j(j>2)$ 帧确定 QP 的具体方法如下:

- 1) 判断与当前帧相邻的前两帧的 QP 是否相同。若 $Q_P(i, j-1)$ 不等于 $Q_P(i, j-2)$, 则直接利用这两个已编码帧的数据, 计算本文设计的 R-QP 方案的参数 α 和 β , 如图 3.2 和式(3-13)、式(3-14)所示, 继而可根据式(3-12)计算出当前帧的 QP。

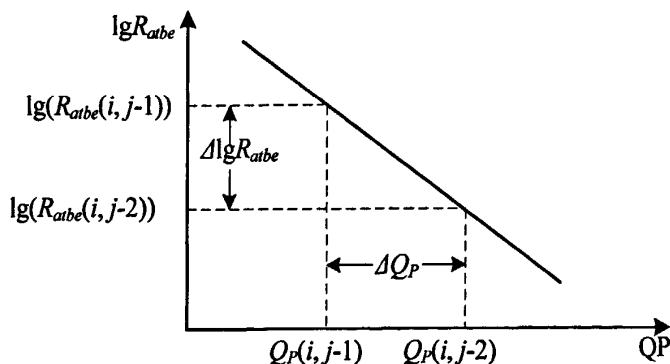


图 3.2 利用已编码帧求解 R-QP 方案的参数

$$\alpha = \frac{\lg(R_{atbe}(i, j-1)) - \lg(R_{atbe}(i, j-2))}{Q_P(i, j-1) - Q_P(i, j-2)} \quad \text{式(3-13)}$$

$$\beta = \frac{Q_p(i, j-1) \times \lg(R_{atbe}(i, j-2)) - Q_p(i, j-2) \times \lg(R_{atbe}(i, j-1))}{Q_p(i, j-1) - Q_p(i, j-2)} \quad \text{式(3-14)}$$

2) 若 $Q_p(i, j-1)$ 等于 $Q_p(i, j-2)$, 首先采用式(3-15)检测已编码帧的 QP 是否准确。若式(3-15)成立, 即已编码帧的目标比特数的均值和实际编码比特数的均值之间的差值小于门限值 T_1 (实验中取值为 300), 则说明已编码帧的 QP 选择较为合理。进而可以采用滑动窗口中的数据计算当前帧的 QP。其中, K 为滑动窗口的长度, 实验中取值为 5。

$$\left| \sum_{k=1}^K R_{tbe}(i, j-k) - \sum_{k=1}^K R_{atbe}(i, j-k) \right| / K < T_1 \quad \text{式(3-15)}$$

其次, 依据式(3-16), 在滑动窗口记录的各帧实际编码比特数中, 寻找与当前帧目标比特数最相近的帧。

$$m = \arg \min_{m \in [1, K]} \{ |R_{tbe}(i, j) - R_{atbe}(i, j-m)| \} \quad \text{式(3-16)}$$

若当前帧的目标比特数和第 $j-m$ 帧的实际编码比特数的差值小于门限值 T_2 (实验中取值为 300), 如式(3-17)所示。根据相邻帧之间的时域相关性, 可以采用第 $j-m$ 帧的 QP 对当前帧进行编码。

$$\left| R_{tbe}(i, j) - R_{atbe}(i, j-m) \right| < T_2 \quad \text{式(3-17)}$$

若式(3-15)或者式(3-17)不成立, 则通过比较前一帧的实际纹理信息编码比特数与当前帧的目标纹理信息编码比特数的大小, 确定当前帧的 QP, 如式(3-18)所示。

$$Q_p(i, j) = \begin{cases} Q_p(i, j-1) - 1, & R_{tbe}(i, j) \geq R_{atbe}(i, j-1) \\ Q_p(i, j-1) + 1, & R_{tbe}(i, j) < R_{atbe}(i, j-1) \end{cases} \quad \text{式(3-18)}$$

3) 最后, 根据下式对计算得到的 QP 进行调整, 以确保 QP 在有效的范围内。

$$Q_p(i, j) = \min\{Q_p(i, j), \max\{Q_p(i, j), 1\}, 51\} \quad \text{式(3-19)}$$

通过上述步骤, 为当前帧选择合适的 QP 之后, 才能对当前帧进行编码。

3.4 实验与分析

将本文设计的 R-QP 方案应用在 H.264/AVC 视频编码标准的参考软件 JM15.1 中, 采用 IPPP 编码预测结构, 在不同目标码率条件下, 对 qcif 和 cif 两种标准测试序列进行测试, 以比较本文设计的方案(标记为“线性方案”)与 JM15.1 中推荐使用的码率控制方案(标记为“二次方案”)的性能。

图 3.3 给出了两种方案的 PSNR 比较曲线。从图中可以看出, 与“二次方案”相比, “线性方案”编码产生的 PSNR 相对稳定, 这表明解码端恢复视频质量的

波动相对较小。

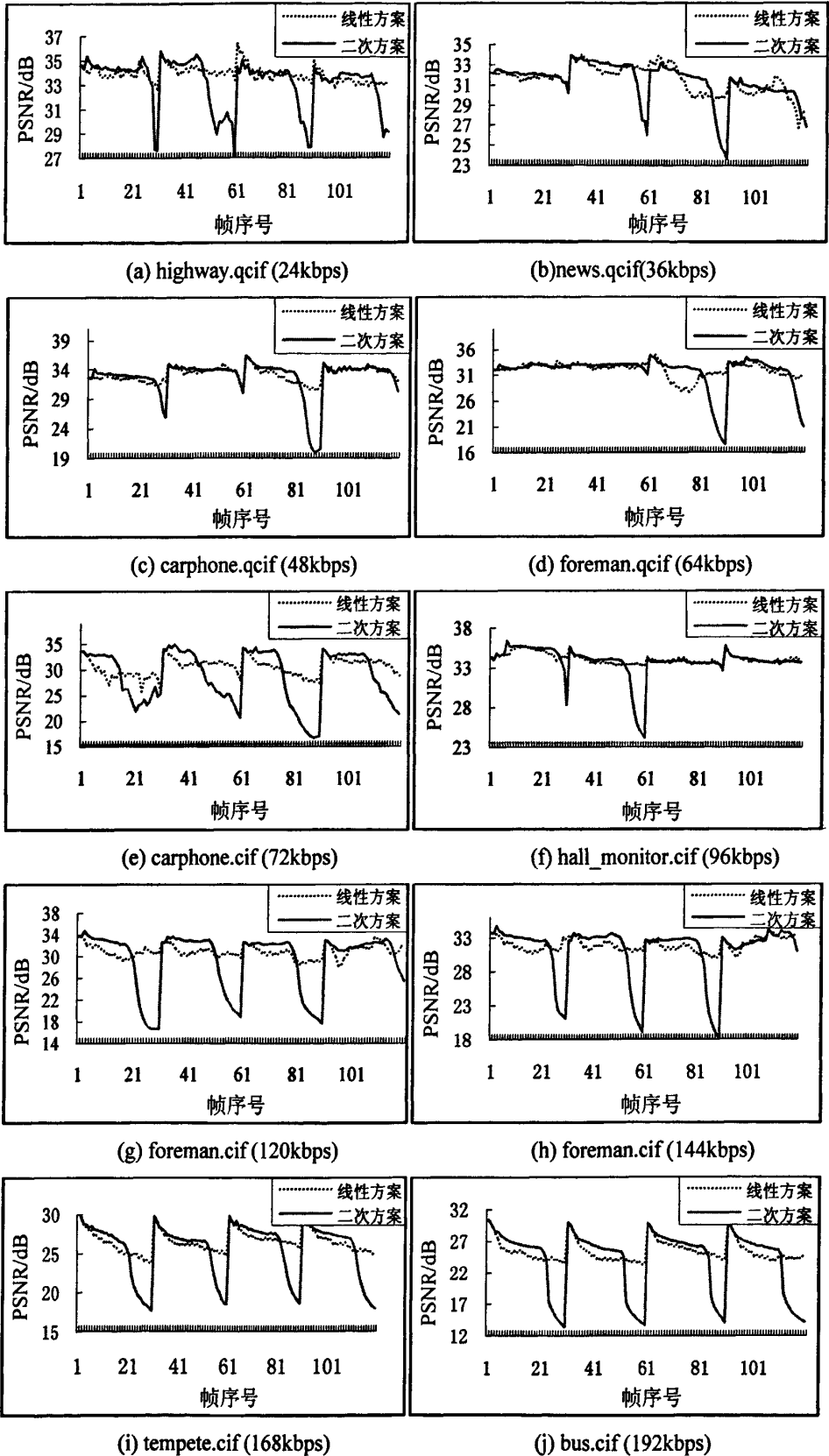


图 3.3 不同序列“线性方案”和“二次方案”的 PSNR 比较曲线

在不同目标比特率条件下,使用“线性方案”和“二次方案”对不同的标准测试序列(qcif和cif)进行编码。表3.3和表3.4给出了相应的测试结果。表中, ΔRATE 表示“线性方案”相对误差与“二次方案”相对误差之间的差值(为负则表示“线性方案”的码率比“二次方案”的码率更接近目标码率;为正则表示“二次方案”的码率比“线性方案”的码率更接近目标码率)。相对误差可根据式(2-1)进行计算。 ΔPSNR 表示在相同目标码率条件下,与“二次方案”相比,“线性方案”的编码质量(以PSNR衡量)的提高(ΔPSNR 的值为正)或下降(ΔPSNR 的值为负)的程度。从表中可以看出,在测试范围内,大多数测试序列的 ΔPSNR 为正值, ΔRATE 为负值。这表明在相同的测试条件下,与“二次方案”相比,本文所设计的“线性方案”能够以较少的编码比特数获得较高的恢复视频质量,并且其码率控制的精度较高。统计结果表明,与“二次方案”相比,“线性方案”对qcif格式的序列进行编码得到的恢复视频质量平均提高了0.15 dB,码率控制精度平均提高了0.46%;对cif格式的序列进行编码得到的PSNR平均提高了1.43dB,码率控制精度平均提高了1.04%。这说明本文所设计的码率控制方案在码率较低的情况下,能够在一定程度上改善码率控制的性能,进而提高编码器的编码效率。

表3.3 “线性方案”和“二次方案”PSNR和码率的比较

序列 (qcif)	目标 码率 (kbps)	线性方案			二次方案			比较结果	
		PSNR (dB)	RATE (kbps)	相对误 差(%)	PSNR (dB)	RATE (kbps)	相对误 差(%)	ΔPSNR (dB)	ΔRATE (%)
silent		30.95	36.27	0.75	30.20	36.56	1.56	0.75	-0.81
news		31.54	36.10	0.28	31.40	36.65	1.81	0.14	-1.53
carphone		31.76	36.55	1.53	31.17	36.78	2.17	0.59	-0.64
foreman	36	29.07	36.15	0.42	27.21	36.75	2.08	1.86	-1.66
container		34.05	36.32	0.89	34.01	36.57	1.58	0.04	-0.69
highway		35.16	36.30	0.83	35.10	36.64	1.78	0.06	-0.95
susie		32.96	36.23	0.64	32.09	36.25	0.69	0.87	-0.05
silent		32.10	48.18	0.37	31.89	48.74	1.54	0.21	-1.17
news		32.98	48.14	0.29	33.00	48.37	0.77	-0.02	-0.48
carphone		33.22	48.19	0.40	32.87	48.84	1.75	0.35	-1.35
foreman	48	30.65	48.09	0.19	29.81	48.93	1.94	0.84	-1.75
container		35.08	48.19	0.40	35.13	48.43	0.90	-0.05	-0.50
highway		36.00	48.53	1.10	35.98	48.30	0.62	0.02	0.48
susie		34.28	48.52	1.08	33.90	48.16	0.33	0.38	0.75

silent		33.36	64.17	0.27	33.49	64.46	0.72	-0.13	-0.45
news		34.50	63.96	-0.06	34.35	64.37	0.58	0.15	-0.64
carphone		34.49	64.32	0.50	34.48	64.62	0.97	0.01	-0.47
foreman	64	32.10	64.03	0.05	31.77	64.95	1.48	0.33	-1.43
container		36.11	64.22	0.34	36.02	64.09	0.14	0.09	0.20
highway		36.82	64.14	0.22	36.80	64.36	0.56	0.02	-0.34
susie		35.39	64.18	0.28	35.41	64.05	0.08	-0.02	0.20
silent		35.27	80.74	0.92	35.25	80.97	1.21	0.02	-0.29
news		36.85	80.06	0.08	36.88	80.43	0.54	-0.03	-0.46
carphone		35.64	80.13	0.16	35.64	80.77	0.96	0.00	-0.80
foreman	80	33.13	80.01	0.01	33.02	81.12	1.40	0.11	-1.39
container		38.00	80.29	0.36	38.06	80.63	0.79	-0.06	-0.42
highway		37.96	80.46	0.57	37.94	80.33	0.41	0.02	0.16
susie		36.50	80.31	0.39	36.28	80.21	0.26	0.22	0.13
silent		35.87	96.53	0.55	36.38	96.55	0.57	-0.51	-0.02
news		37.93	95.67	-0.34	38.00	96.52	0.54	-0.07	-0.89
carphone		36.51	96.09	0.09	36.62	96.86	0.90	-0.11	-0.80
foreman	96	34.08	95.97	-0.03	34.33	96.81	0.84	-0.25	-0.88
container		38.58	96.22	0.23	38.74	96.56	0.58	-0.16	-0.35
highway		38.51	96.18	0.19	38.48	96.13	0.14	0.03	0.05
susie		37.28	96.59	0.61	37.17	95.94	-0.06	0.11	0.68
silent		37.30	112.28	0.25	37.29	112.33	0.29	0.01	-0.04
news		38.85	111.70	-0.27	38.97	112.08	0.07	-0.12	-0.34
carphone		37.29	112.36	0.32	37.32	112.63	0.56	-0.03	-0.24
foreman	112	34.82	111.77	-0.21	34.35	112.30	0.27	0.47	-0.47
container		39.25	112.53	0.47	39.30	112.42	0.38	-0.05	0.10
highway		38.92	112.29	0.26	38.91	112.34	0.30	0.01	-0.04
susie		38.01	112.53	0.47	38.00	112.30	0.27	0.01	0.21
Average	—	—	—	0.38	—	—	0.84	0.15	-0.46

表 3.4 “线性方案”和“二次方案”PSNR 和码率的比较

序列 (cif)	目标 码率 (kbps)	线性方案			二次方案			比较结果	
		PSNR (dB)	RATE (kbps)	相对误 差(%)	PSNR (dB)	RATE (kbps)	相对误 差(%)	Δ PSNR (dB)	Δ RATE (%)
coastguard	64	20.65	65.46	2.28	20.03	66.38	3.72	0.62	-1.44
foreman		23.19	64.55	0.86	22.83	65.34	2.09	0.36	-1.23
waterfall		26.41	64.33	0.52	25.43	64.80	1.25	0.98	-0.73
Carphone		29.93	64.30	0.47	27.92	65.74	2.72	2.01	-2.25
coastguard	72	21.87	72.35	0.49	20.25	73.14	1.58	1.62	-1.09
foreman		27.54	72.37	0.51	24.11	73.79	2.49	3.43	-1.98
waterfall		27.08	73.03	1.43	24.83	73.21	1.68	2.25	-0.25
Carphone		30.63	72.21	0.29	28.75	73.34	1.86	1.88	-1.57
coastguard	96	24.65	96.20	0.21	21.74	96.33	0.34	2.91	-0.13
foreman		29.46	96.81	0.84	27.12	97.74	1.81	2.34	-0.97
waterfall		29.17	96.46	0.48	26.29	97.09	1.14	2.88	-0.66
Carphone		32.19	96.43	0.45	30.55	97.50	1.56	1.64	-1.11
coastguard	120	25.70	120.10	0.08	22.91	121.30	1.08	2.79	-1.00
foreman		30.95	120.02	0.02	29.44	122.11	1.76	1.51	-1.74
waterfall		30.03	120.43	0.36	28.31	121.58	1.32	1.72	-0.96
Carphone		33.21	120.20	0.17	32.11	122.05	1.71	1.10	-1.54
coastguard	144	26.53	144.61	0.42	24.12	145.91	1.33	2.41	-0.91
foreman		31.92	144.22	0.15	31.11	145.79	1.24	0.81	-1.09
waterfall		30.81	145.12	0.78	29.29	146.22	1.54	1.52	-0.76
Carphone		34.05	144.53	0.37	33.62	146.33	1.62	0.43	-1.25
coastguard	168	27.10	168.39	0.23	25.56	170.07	1.23	1.54	-1.00
foreman		32.73	169.07	0.64	32.31	169.53	0.91	0.42	-0.27
waterfall		31.43	168.63	0.37	30.56	170.16	1.29	0.87	-0.92
Carphone		34.70	168.28	0.17	34.47	169.66	0.99	0.23	-0.82
coastguard	192	27.39	192.34	0.18	26.29	194.47	1.29	1.10	-1.11
foreman		33.38	193.01	0.53	33.10	193.66	0.86	0.28	-0.33
waterfall		31.73	192.77	0.40	31.41	194.50	1.30	0.32	-0.90
Carphone		35.19	192.48	0.25	35.15	194.34	1.22	0.04	-0.97
average	—	—	—	0.50	—	—	1.53	1.43	-1.04

3.5 小结

在 H.264/AVC 视频编码标准的基本档次中, P 帧出现的概率可达 90%以上。因此, 为 P 帧确定合适的量化参数对编码器的性能有着重要影响。本章首先介绍了一些经典的码率控制方案, 并分析了在低码率情况下码率控制性能下降的原因。然后通过理论分析, 发现编码比特率的对数与量化参数之间存在线性关系。对此, 本章通过实验进行了进一步的验证。在此基础上, 研究了一种基于码率-量化参数的码率控制方案, 并给出了确定量化参数的具体方法。最后, 将设计的码率控制方案应用于 H.264/AVC 视频编码标准的参考软件 JM15.1 中, 并与 H.264/AVC 视频编码标准推荐使用的码率控制方案进行比较。

在低码率情况下, 与参考方案相比, 本文所研究的方案可以明显提高码率控制的性能, 进而提高编码器的效率。对于 qcif 和 cif 两种格式的视频序列, 本文研究方案的恢复视频质量的平均增益可分别提高 0.15dB 和 1.43dB, 码率控制的精度可分别提高 0.46%和 1.04%。

第四章 低码率跳帧算法的改进

4.1 引言

视频序列中，每一帧内容的运动剧烈程度以及图像复杂度的不断变化，导致编码器产生的编码比特数也在不断变化。为使可变的码流能够在恒定比特率 (Constant Bit Rate, CBR) 的通信信道中进行传输，在编码器与信道之间设置了缓冲区，如图 4.1 所示。编码器将视频压缩码流存放于缓冲区中，同时，缓冲区按照恒定的比特率将压缩码流输出到信道中，从而保证了码流的恒定比特率传输。

缓冲区大小是有限的，其值由最大时延决定。时延越大，相应的缓冲区越大；反之，时延越小，相应的缓冲区越小。为了防止缓冲区溢出，码率控制机制将“缓冲区占有度”作为调整编码参数的考虑因素之一。当缓冲区占有度较低时，码率控制为编码器确定一个较小的量化参数以增加压缩的码流，防止缓冲区发生下溢而浪费带宽；当缓冲区占有度较高时，码率控制器为编码器确定一个较大的量化参数以减小压缩的码流，甚至采用跳帧技术以防止缓冲区出现上溢而导致信息丢失。跳帧是一种能够有效避免缓冲区上溢的策略。但是，跳帧使得解码端恢复视频的质量产生较大的波动，从而影响恢复视频的主观质量^[54]，如图 4.2 所示。尤其在给定目标码率较低的情况下，由于缓冲区设置的相对较小，使编码过程中发生跳帧的概率相对较大，严重影响编码器的性能^[24]。因此，如何设计跳帧策略以降低跳帧发生的概率、提高恢复视频的质量是码率控制中一个非常关键的问题。

本章内容安排如下：4.2 小节介绍了一些已有的跳帧算法并对其进行了相应的分析；4.3 小节在分析了帧间预测模式与编码效率之间统计特性的基础上，探索性地研究了一种降低跳帧发生概率的算法。4.4 小节给出了实验结果及相应的分析；4.5 小节是本章的小结。

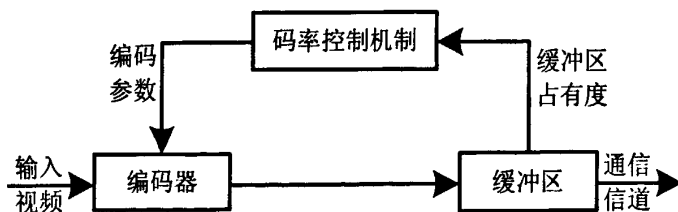


图 4.1 码率控制基本原理框图

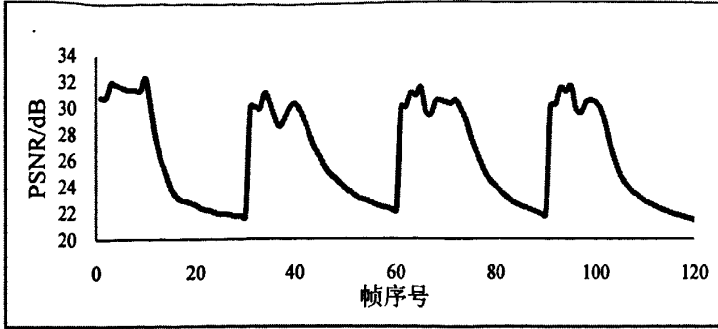


图 4.2 跳帧导致解码端视频质量的波动(waterfall 序列, 目标比特率: 96kbps)

4.2 已有的跳帧策略

跳帧能够有效避免缓冲区上溢, 然而, 其引起的恢复视频质量的波动是不容忽视的。较早的跳帧算法直接将缓冲区的 80% 作为跳帧阈值^[23,49,50,55]。当缓冲区占有度超过该阈值时, 不考虑当前帧的复杂度, 直接跳过不对其进行编码, 跳帧之后更新缓冲区的状态, 直至缓冲区占有度小于所设定的阈值才停止跳帧。在解码端, 直接将与该帧相邻最近的未发生跳帧的解码帧作为该跳帧的恢复帧。这种跳帧方式虽然可以有效避免缓冲区上溢, 但是, 盲目的跳过运动剧烈的图像势必会影响解码端图像的连续性, 从而导致恢复视频质量的下降。

为了在不出现缓冲区上溢的情况下, 尽量保持解码端视频的连贯性, 文献[56-60]提出了不同的衡量图像运动剧烈程度的方法, 并结合序列的运动剧烈程度与缓冲区的不同状态, 判断是否对当前帧进行跳帧处理。其基本思想为: 尽量跳过运动较为平缓的帧, 为运动剧烈的图像保留缓冲区空间, 以避免对运动剧烈的图像跳帧。从而, 保证视频的连贯性, 提高恢复视频的质量。文献[56]采用当前帧 MAD 的预测值与前 n 个已编码 P 帧实际 MAD 的平均值之间的比值作为衡量图像运动剧烈程度的量。文献[57]认为文献[61]提出的衡量图像运动剧烈程度的方法较为准确。文献[58]利用编码比特数与图像运动剧烈程度、纹理复杂度的线性关系, 提出了利用当前帧编码比特数的预测值与前 n 个已编码帧的实际编码比特数的平均值之间的比值作为衡量图像运动剧烈程度的量。文献[59]采用 Jiang 提出的 PSNR 与 MAD 相结合的方法^[62]估计当前帧的运动剧烈程度。由上述可知, 文献[56-60]提出的方法无一例外的将当前帧的运动剧烈程度与当前缓冲区的占有度结合起来, 为不同运动剧烈程度的帧预先确定不同的阈值, 尽量跳过运动平缓的帧, 为运动剧烈的图像保留缓冲区空间, 以降低对运动剧烈的图像发生跳帧的概率, 从而提高了视频序列的连续性。这种跳帧机制控制效果较好, 但是计算量较大。

基于此, 本章探索性地研究了一种低码率情况下, 复杂度较低的降低跳帧发生概率的算法。

4.3 改进的跳帧算法

4.3.1 基本思想

当给定目标比特率很低时, 设置的缓冲区容量相对较小, 跳帧发生的概率相对较大。为了尽量避免由跳帧导致恢复视频质量不连贯而使视频质量下降的现象, 本文根据缓冲区占有度大小, 实时地更新码率控制确定的量化参数和率失真优化准则选择的最佳编码模式, 即当缓冲区占有度较低时, 按照原有方法进行编码; 当缓冲区占有度较高时, 对当前帧宏块的预测模式以及帧层的量化参数进行进一步调整, 以减少编码当前帧所用的编码比特数, 降低缓冲区的占有度, 从而尽量降低后续未编码帧发生跳帧的概率。与此同时, 还可以为后续未编码帧节省一定的比特数, 提高解码端恢复视频的质量; 当缓冲区占有度过高时, 采用跳帧策略以防止缓冲区溢出而丢失编码比特流。考虑低码率的情况, 本文所设计的方法能够在一定程度上降低跳帧发生的概率, 从而提高编码器的性能。

4.3.2 帧间预测模式与编码效率特性分析

为了提高视频编码质量, H.264/AVC 视频编码标准提供了更多可供选择的预测模式^[63], 其主要包括帧间和帧内两大类。根据宏块的不同划分方式可以将帧间模式分为四种, 如图 4.3 所示。其中, 模式 4 又可以分为 4 个子模式。帧内模式主要有两种: I16×16 和 I4×4。

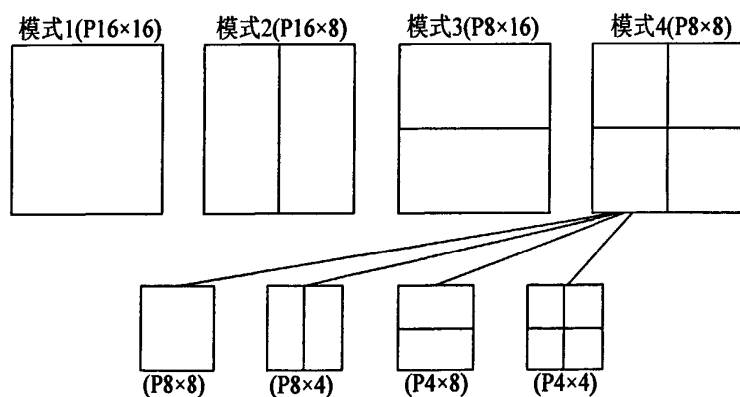


图 4.3 H.264/AVC 的帧间宏块划分模式及其子模式

为了得到当前宏块的最佳编码模式, H.264/AVC 编码器采用了率失真优化准则^[52]。基于拉格朗日乘数的率失真代价函数的定义如下,

$$J_{mode} = D + \lambda_{mode} R \quad \text{式(4-1)}$$

其中, J_{mode} 表示率失真代价值(Rate Distortion Cost, RDC), D 为当前宏块与解码

端恢复宏块之间的失真度, R 表示编码当前宏块所需要的编码比特数, λ_{mode} 为拉格朗日乘数, 对于 I 帧和 P 帧, 可以通过式(4-2)求得,

$$\lambda_{mode} = 0.85 \times 2^{(QP-1)/3} \quad \text{式(4-2)}$$

对于 B 帧, 可通过下式求得,

$$\lambda_{mode} = 4 \times 0.85 \times 2^{(QP-1)/3} \quad \text{式(4-3)}$$

利用式(4-1)为当前宏块计算每种编码模式的率失真代价值之后, 选择使得 J_{mode} 最小的编码模式作为当前宏块的最佳编码模式。

在低码率情况下, 不同的帧间预测模式对编码效率的贡献不同, 这可以通过实验进行验证。统计索引号为 x 的帧间预测模式, 对编码效率贡献的具体方法为:

- 1) 统计原方案的测试结果。按照 H.264/AVC 编码器所推荐使用的率失真优化准则为每一个宏块确定最佳的编码模式, 并对输入的视频序列进行编码。记录编码得到的编码比特率 R_{ate_org} 和恢复视频质量 P_{snr_org} 。
- 2) 统计缺省索引号为 x 的帧间预测模式的测试结果。利用率失真优化准则为当前宏块选择最佳的编码模式。如果选择的最佳预测模式的索引号为 x , 则采用 skip 模式对当前宏块进行编码; 否则, 则根据所确定的最佳预测模式进行编码。按照这种方式进行编码, 就可以得到缺省索引号为 x 的预测模式的编码结果。记录编码得到的编码比特率 R_{ate_x} 和恢复视频质量 P_{snr_x} , 并记录在编码过程中, 根据率失真优化准则确定的最佳预测模式的索引号为 x 的宏块个数 M_{ode_num} 。
- 3) 由 1) 和 2) 的编码结果可以得到, 当没有索引号为 x 的预测模式时, 编码比特率的变化量 ΔR_{ate} (为正则表示没有索引号为 x 的预测模式时, 降低的码率; 否则, 则表示增加的码率) 和恢复视频质量的变化量 ΔP_{snr} (为正则表示没有索引号为 x 的预测模式时, 降低的 PSNR 值; 否则, 则表示提高的 PSNR 值) 分别为,

$$\Delta R_{ate} = R_{ate_org} - R_{ate_x} \quad \text{式(4-4)}$$

$$\Delta P_{snr} = P_{snr_org} - P_{snr_x} \quad \text{式(4-5)}$$

- 4) 本文通过定义帧间预测模式影响因子 (F_{actor}) 来衡量索引号为 x 的预测模式对编码效率影响的大小, 其定义为,

$$F_{actor} = \frac{\Delta R_{ate}}{\Delta P_{snr} \times M_{ode_num}} \quad \text{式(4-6)}$$

由上述定义可见, F_{actor} 表示采用索引号为 x 的预测模式时, 恢复视频质量每提高 1dB 用于单位宏块上的编码比特数。 F_{actor} 数值越大, 表明该预测模式对编码效率的提高越小。反之, 则说明该预测模式对编码效率的

贡献越大。

按照上述方法, 采用 H.264/AVC 视频编码标准所推荐的参考软件 JM15.1, 利用公共测试序列, 对 H.264/AVC 所提供的帧间预测模式的 F_{actor} 进行统计。图 4.4 给出了不同序列的测试结果。从图 4.4 可以看出, P8×8 预测模式的 F_{actor} 数值最大。这说明 P8×8 预测模式需要更多的编码比特数才能获得与其它预测模式同样的编码效率。这也意味着用 skip 模式代替 P8×8 模式进行编码, 可以牺牲最小的 PSNR 来节约最多的编码比特数, 保留更大的缓冲区空间。从而, 尽可能避免跳帧的发生。

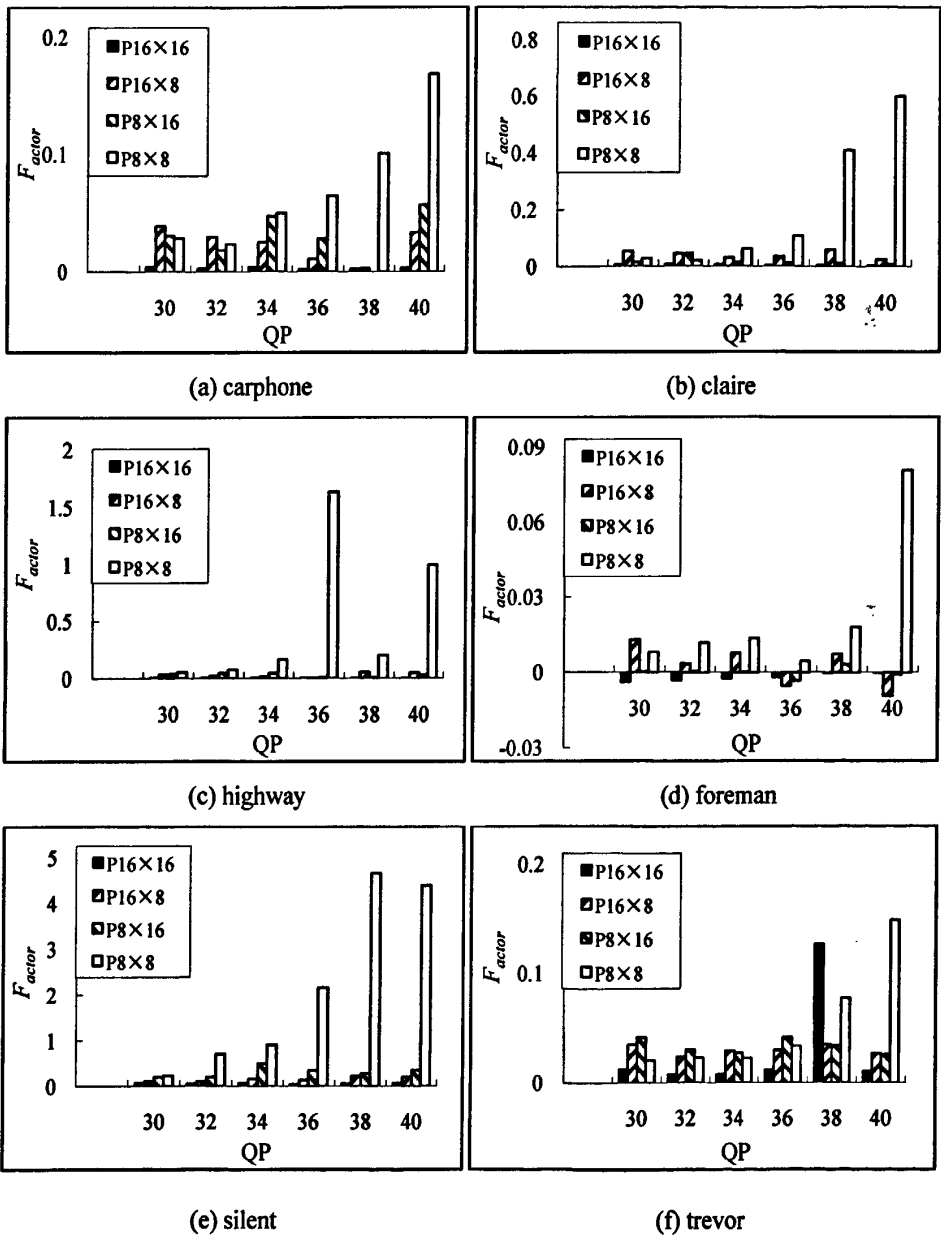


图 4.4 帧间预测模式对编码效率的影响

4.3.3 跳帧算法的实现

当给定目标比特率较低时,为了尽量减少由跳帧导致的视频质量下降的现象,本文根据缓冲区的不同状态和宏块的帧间预测模式与编码效率之间的统计特性,实时地为当前编码单元更新编码参数。本文设计的降低跳帧算法的具体实现方法如下:

- 1) 为缓冲区设置两个阈值 B_{min} 和 B_{max} , B_{min} 的数值小于 B_{max} 。 B_{min} 的数值设置不能太小,太小则影响恢复视频的质量; B_{min} 的数值也不能设置过大,过大则不能起到降低跳帧发生概率的作用。在本文实验中, B_{min} 设置为缓冲区的 65%, B_{max} 的数值按照以往的经验设置为缓冲区的 80%。
- 2) 在编码过程中,如果当前缓冲区占有度小于 B_{min} ,则按照原来的方式进行编码;如果缓冲区占有度介于 B_{min} 和 B_{max} 二者之间,则将码率控制机制所确定的量化参数进行适当放大(在实验中,将量化参数加 1),并将放大之后的量化参数作为当前帧的编码参数。对当前帧进行编码的过程中,根据宏块帧间预测模式与编码效率之间的统计特性更新编码模式,即当率失真优化准则确定的最佳预测模式为 P8×8 时,强制对当前宏块进行 skip 模式编码;否则,则按照率失真优化准则确定的最佳编码模式进行编码。如果缓冲区占有度大于预定义的阈值 B_{max} ,必须对当前帧进行跳帧处理。一方面为了节省编码时间,另一方面为了防止缓冲区溢出导致图像数据丢失的情况发生。

上述给出了本文降低跳帧算法的实现过程。可见,该方法简单,易于实现。

4.4 实验与分析

将本文设计的减少跳帧的方法应用在 H.264/AVC 视频编码标准的参考软件 JM15.1 中,在 H.264/AVC 视频编码标准的基本档次下,使用 container、foreman、mthr_dotr 和 silent 四个 qcif 格式的标准测试序列,在给定的不同目标码率条件下,比较该方法(标记为“本文方法”)与 JM15.1 中推荐使用的方法(标记为“原有方法”)的性能。

表 4.1 给出了测试结果。从表中可以看出,与“原有方法”相比,“本文方法”虽然产生的实际编码比特率稍高,但可以获得更高质量的恢复视频,还能够有效地减少跳帧的帧数。

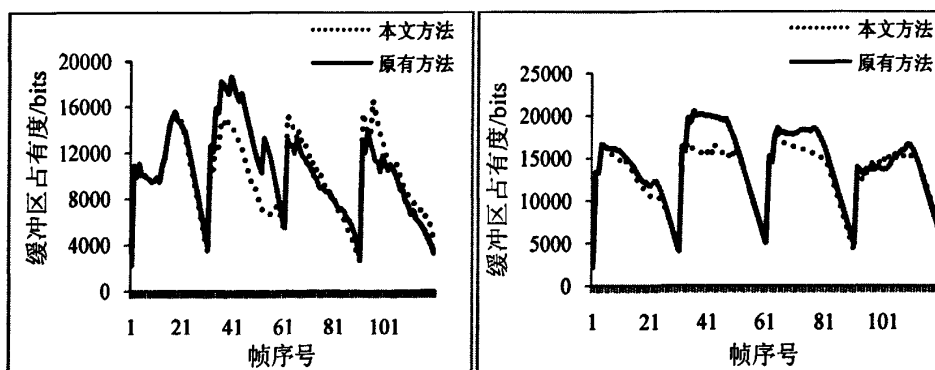
图 4.5 比较了“原有方法”和“本文方法”的缓冲区占有度曲线。从图中可以看出,在大多数情况下,使用本文设计的方法进行编码得到的缓冲区占有度要小于“原有方法”的情况。从而,能够降低跳帧的概率,提高恢复视频质量。

图 4.6 给出了两种方法的 PSNR 曲线。从图中可以看出,对所测试的大多数帧,“本文方法”编码得到的 PSNR 高于“原有方法”的 PSNR,并且通过“本文方法”得到的 PSNR 曲线更为平稳,这表明恢复视频质量的波动相对较小。

综上,与原有的跳帧方案相比,在低码率情况下,本文设计的方法可以有效地降低跳帧发生的概率。从而,在一定程度上改善编码器的性能。

表 4.1 与“原有方法”的性能比较

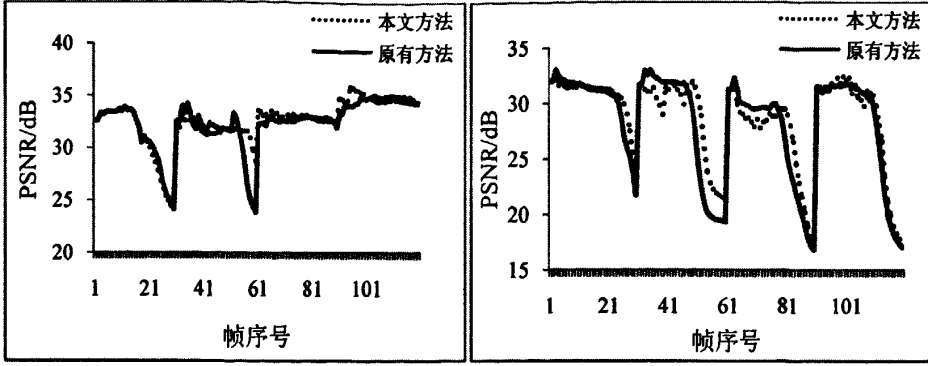
序列	原有方法			本文方法			比较结果	
	Rate (kbps)	PSNR (dB)	NUM (帧)	Rate (kbps)	PSNR (dB)	NUM (帧)	PSNR 差 值(dB)	跳帧帧数 差值(帧)
container	24.37	31.36	33	24.65	31.69	29	0.33	-4
foreman	24.65	23.49	73	25.44	25.83	36	2.34	-37
mthr_dotr	24.51	30.05	41	24.62	30.36	30	0.31	-11
silent	24.83	27.95	50	24.86	28.25	39	0.30	-11
container	36.50	33.86	7	36.45	33.75	6	-0.11	-1
foreman	36.95	28.25	33	37.02	28.66	26	0.41	-7
mthr_dotr	36.28	32.32	10	36.37	32.59	7	0.27	-3
silent	36.58	30.41	18	36.63	30.47	16	0.06	-2
container	48.39	35.03	0	48.38	35.04	0	0.01	0
foreman	49.08	30.67	14	48.71	30.61	14	-0.06	0
mthr_dotr	48.14	33.97	0	48.34	33.99	0	0.02	0
silent	48.63	31.66	13	48.60	31.99	7	0.33	-6



(a) mthr_dotr.qcif

(b) foreman.qcif

图 4.5 两种方法的缓冲区占有度比较曲线



(a) mthr_dotr.qcif

(b) foreman.qcif

图 4.6 两种方法的 PSNR 比较曲线

4.5 小结

本章分析了跳帧对恢复视频质量的影响，并介绍了一些已有的跳帧方法。在低码率情况下，通过实验统计并分析了宏块的不同帧间预测模式对编码效率的影响。实验结果表明， $P8 \times 8$ 预测模式对编码效率的贡献最小。在此基础上，本文设计了一种降低跳帧发生概率的算法。该算法根据缓冲区的不同状态以及宏块预测模式与编码效率之间的统计特性，实时地为当前编码单元更新编码参数。

将本文设计的算法应用于 H.264/AVC 的参考软件 JM15.1 中，并与 H.264/AVC 视频编码标准所推荐使用的方案进行比较。结果表明，在低码率情况下，本文设计的方法能够有效降低跳帧发生的概率，进而提高编码器的性能。

第五章 总结与展望

为了在资源受限的条件下,获得质量最优的恢复视频,码率控制为编码器确定合适的编码参数。由于这些编码参数直接应用于视频编码器中,其准确度直接影响着编码器的性能。因此,码率控制技术是视频编码器的重要组成部分。本文以 MPEG 和 VCEG 共同提出的 H.264/AVC 视频编码标准为基础,对其推荐使用的码率控制算法进行了研究与优化,具体的研究工作总结如下:

1. 序列初始量化参数的确定。确定序列的初始量化参数是码率控制的关键技术之一。本文通过实验分析了初始量化参数对整个视频编码质量的影响,介绍了一些已有的确定序列初始量化参数的方法,并对其进行了相应的分析。考虑到初始量化参数不仅受视频格式、目标比特率的影响,而且还受到图像复杂度的影响,本文通过实验推导了一种确定初始量化参数的方案。通过对该方案中的相关参数进行分析,发现不同的图像复杂度将导致相同测试条件下,不同序列的最佳初始量化参数之间出现偏差。为了定量该偏差,本文利用 Sobel 算子研究了一种衡量图像复杂度的方法,并对该方法进行了相应的实验验证。在此基础上,深入地分析了不同序列的图像复杂度与最佳初始量化参数偏差之间的定量关系。在低码率的情况下,与 JVT-G012 推荐的方法相比,利用本文所设计的确定初始量化参数的方法进行编码,可以在一定程度上改善编码器的性能。

2. 低码率线性码率控制方案。H.264/AVC 视频编码标准所推荐使用的码率控制机制通过码率-量化步长之间的关系,为编码器确定合适的量化步长。由于量化步长不是直接的编码参数,必须将其转换为相应的量化参数,才能对当前的编码单元进行编码。然而,量化步长转换为量化参数的过程中存在着不可避免的舍入误差,而且该误差随着量化步长的增大而增大。从而,影响了码率控制的性能。本文首先对一些经典的码率控制机制进行了简要的介绍,并分析了低码率下码率控制性能下降的原因。在此基础上,通过理论推导研究了一种基于码率-量化参数之间关系的码率控制方案,并利用 H.264/AVC 的参考软件 JM15.1 对其进行了实验验证,实验结果证明了该方案的正确性。基于以上的分析和研究,给出了本文方案确定编码参数的具体步骤。最后,将本文研究的方案应用在 H.264/AVC 的参考软件 JM15.1 中进行测试。测试结果表明,与 H.264/AVC 所推荐使用的码率控制方案相比,本文设计的方案可以用较少的编码比特数获得较高质量的恢复视频。另外,通过本文设计的方案获得的恢复视频质量的波动更小,码率控制的精度更高。

3. 低码率跳帧算法的改进。在视频编码过程中,跳帧策略可以有效避免缓冲区溢出。但是,当给定目标码率较低时,由于设置的缓冲区容量较小,发生跳帧的概率增大。这大大降低了解码端恢复视频的质量。针对这一问题,本文首先统

计了 H.264/AVC 支持的各种帧间预测模式与编码效率之间的关系。统计结果表明, 给定目标码率较低时, P8×8 预测模式对编码效率的贡献最小。这意味着以 skip 模式代替 P8×8 模式可以牺牲较小的 PSNR 节省更多的编码比特数, 从而降低缓冲区占有度。此外, 量化参数也可以有效地控制编码器输出码流的比特率。在此基础上, 本文利用帧间预测模式与编码效率的统计特性以及量化参数, 探索性地研究了一种低码率情况下降低跳帧发生概率的算法。该算法根据缓冲区的不同状态实时地更新编码参数: 如果当前缓冲区占有度介于两个预定义的阈值之间, 该算法首先将码率控制确定的量化参数进行调整, 然后, 通过率失真优化准则为当前宏块选择最佳编码模式。最后, 根据帧间预测模式与编码效率的统计特性, 为当前宏块调整最终的预测模式; 对于其它缓冲区状态, 按照原有方法进行编码。将本文所设计的方法应用于 H.264/AVC 的参考软件 JM15.1 中, 在给定目标比特率较低的情况下, 该方法可以在一定程度上降低跳帧发生的概率, 提高编码器的编码效率。

本论文主要研究了给定目标码率较低情况下的码率控制算法, 从序列初始 QP 的确定、P 帧 QP 的确定和降低跳帧发生的概率三个方面对 H.264/AVC 视频编码标准所推荐使用的码率控制算法进行了改进。其中, 序列初始 QP 除了受编码第一帧图像复杂度的影响之外, 还受时域图像复杂度的影响。如果能够准确提取衡量时域图像复杂度的方法, 进而为当前序列确定初始 QP, 则可能会进一步提高编码器的编码效率。除此之外, 在不同的目标比特率, 码率控制方案的性能不同。目前, 尚未存在一个码率控制算法能够适应所有的情况。因此, 作为编码器的关键技术之一, 码率控制机制还需要进行进一步的研究。

致谢

随着硕士毕业论文的完成，我的研究生生活也渐渐地接近尾声。在整个研究生期间，我在学习上、生活上、科研工作中的每一份收获都离不开老师的孜孜教诲、同学们的热情帮助和家人朋友们默默无闻的支持。在这里，我要向所有关心我、帮助我、支持我的老师、同学、朋友和家人献上我最诚挚的谢意。

首先，我要向敬爱的常义林教授致以由衷的感谢。研究生期间，常老师在各个方面都给我很大的支持和帮助。在学习上，常老师严谨的治学态度为我树立了学习的榜样，激励着我要踏实、努力学习；在生活上，常老师无微不至的关怀让我体会到了慈父般的温暖，让我可以无后顾之忧的将全部精力投入到学习和科研工作中；在科研工作中，具有渊博的专业知识、敏锐的洞察力、极强的科学预见性和实事求是工作精神的常老师耐心地对我的科研工作中遇到的问题给予分析和指导，对我的学术研究有极大的启发。同时也非常感谢师母对常老师工作的支持以及对我慈母般的关怀。在此，谨向常老师以及师母致以最真诚的感谢！

同时，我也非常感谢霍俊彦老师。感谢霍老师对我工作能力的信任，让我参与科研项目进行锻炼。感谢霍老师在科研项目中耐心的指导我，开拓我思考问题的思维方式；感谢霍老师在我生活上遇到困难的时候给予帮助。谢谢你，霍老师。

非常感谢元辉博士后在生活、学习以及科研工作中给予的帮助。元辉师兄刻苦努力的学习精神、一丝不苟的治学态度一直激励着我要踏踏实实的学习工作；当科研中遇到问题时，元辉师兄总是以敏锐的思维能力和饱满的工作热情给予无私的帮助。在生活中，元辉师兄总是耐心的给我解答各种困惑。谢谢你，师兄。

感谢任光亮老师给我提供就读西安电子科技大学通信与信息系统专业硕士的机会。感谢杨付正老师，马彦卓老师和任光亮老师在科研工作以及生活上的关怀和帮助。你们严谨的治学态度和一丝不苟的工作作风为我树立了学习的楷模，激励着我要踏踏实实工作。与你们的每次讨论都让我受益匪浅。

感谢冯妮娜博士，李明博士，贺竞师弟对论文所做的工作。感谢李世光、马祥、赵海霞、张泉枝、亢丽娟、杨旭、刘晓仙、李鹏、冯妮娜、李志斌、李明、江亮亮、苏洪磊、李维、宋佳润、李蕊、李睿歌、史慧敏、李凡、孟勇、邹娜、王珂等已经毕业或者未毕业的兄弟姐妹们，他们无私地和我分享各种经验，激励着我克服困难，勇于进取。和你们在一起的日子，让我终生难忘。谢谢你们！

感谢我的父母，他们给予了我生命，教会我做人的道理。在我遭受挫折困难的时候，激励我，安慰我。感谢我的男友安立刚对我的关爱和照顾，谢谢你在我不失落的时候，默默的支持我，鼓励我。他们的期望是我最大的动力。

最后，向参加论文审阅和答辩的各位专家和老师表示深深的感谢！

参考文献

- [1] K.R. Rao, Zoran S. Bojkovic, Dragorad A. Molovanovic. Multimedia communication system: techniques, standards, and networks[M]. London: Prentice Hall, 2002.
- [2] Yao Wang, Jorn Ostermann, Ya-Qin Zhang. Video Processing and Communications[M], Pearson Education, 2001.
- [3] ISO/IEC DIS 11172-2. Coding of moving pictures and associated audio for digital storage media up to 1.5 Mbits/s[S]. 1992.
- [4] ISO/IEC DIS 13818-2. Generic coding of moving pictures and associated audio information[S]. 1994.
- [5] ISO/IEC DIS 14496. Coding of audio-visual objects. 1999.
- [6] ITU-T Recommendation H.261. Video codec for audio visual services at Px64kbits/s. March, 1990.
- [7] ITU-T Recommendation H.263. Video coding for low bit rate communication. March, 1996.
- [8] ITU-T Recommendation H.263 Version 2. Video coding for low bit rate communication. January, 1998.
- [9] ITU-T Draft. Draft for “H.263++” Annexes U, V, AND W to recommendation H.263[S]. Telecommunication Standardization Sector. November, 2000.
- [10] Telecommunication Standardization Sector. ITU-T recommendation H.264 advanced video coding for generic audio visual services[S]. Geneva, Switzerland: Telecommunication Standardization Sector. May, 2003.
- [11] Telecommunication Standardization Sector. ITU-T recommendation H.264 advanced video coding for generic audio visual services[S]. Geneva, Switzerland: Telecommunication Standardization Sector. March, 2005.
- [12] Telecommunication Standardization Sector. ITU-T recommendation H.264 advanced video coding for generic audio visual services Amendment 1: Support of additional color spaces and removal of the High 4:4:4 Profile[S]. Geneva, Switzerland: Telecommunication Standardization Sector. June, 2006.
- [13] Telecommunication Standardization Sector. ITU-T recommendation H.264 advanced video coding for generic audio visual services Amendment 2: New profiles for professional applications[S]. Geneva, Switzerland: Telecommunication Standardization Sector. April, 2007.

- [14] Joint Video Team (JVT) of ISO/IEC MPEG and ITU-T VCEG, Joint Draft ITU-T Rec. H.264| ISO/IEC 14496-10 / Amd.3 Scalable video coding. JVT-X201. July, 2007.
- [15] Telecommunication Standardization Sector. ITU-T recommendation H.264 advanced video coding for generic audio visual services[S]. Geneva, Switzerland: Telecommunication Standardization Sector. November, 2007.
- [16] Telecommunication Standardization Sector. ITU-T recommendation H.264 advanced video coding for generic audio visual services[S]. Geneva, Switzerland: Telecommunication Standardization Sector. March, 2010.
- [17] Thomas Wiegand, Gray J. Sullivan, Gisle Bjontegaard, Ajay Luthra. Overview of the H.264/AVC video coding standard[J]. IEEE Transaction on Circuit and Systems for Video Technology. 2003, 7, 13(7): 560-576.
- [18] Gary J. Sullivan, Pankaj Topiwala, Ajay Luthra. The H.264/AVC Advanced Video Coding Standard: Overview and Introduction to the Fidelity Range Extensions[J]. SPIE Conference on application of Digital Image Processing. 2004, 8, 8: 1-21.
- [19] 沈兰荪, 卓力. 小波编码与网络视频传输[M]. 北京: 科学出版社, 2008.
- [20] 毕厚杰. 新一代视频压缩编码标准——H.264/AVC[M]. 北京: 清华大学出版社, 2006.
- [21] 余兆明, 查日勇, 黄磊, 周海娇. 图像编码标准 H.264 技术[M]. 北京: 人民邮电出版社, 2006.
- [22] ITU2T/SG15. Video Codec Test Model, TMN5[S]. 1995.
- [23] ITU2T/SG16. Video Codec Test Model, Near2Term, Version (TMN8) Release0[S], 1997.
- [24] ISO/IEC, J TC1/SC29/WG11. MPEG-4 video verification model V18.0, Coding of Moving Pictures and Audio N3908[S], 2001.
- [25] MA SW, GAO W, LU Y.. Rate Control on JVT Standard, Document JVT-D030 [A]. JVT 4th Meeting[C]. Klagenfurt, Austria, 2002.
- [26] MA SW, GAO W, LU Y, etc. Improved Rate Control Algorithm, Document JVT-E069 [A]. JVT 5th Meeting[C]. Geneva, CH, 2002.
- [27] MA SW, GAO W, LU Y, etc. Proposed draft description of Rate control on JVT standard, Document JVT-F086[A], JVT 6th Meeting[C]. Awaji, 2002.
- [28] Z. G. Li, F. Pan, K. P. Lim, G. Feng, X. Lin and S. Rahardja. Adaptive basic unit layer rate control for JVT, JVT-G012-r1, 7th Meeting Pattaya II, Thailand, Mar. 2003.
- [29] Zhengguo Li, Wen Gao, Feng Pan. Adaptive Rate Control with HRD Consideration.

- Document JVT-H014 [A]. JVT 8th Meeting[C]. Geneva, 2003.
- [30] Zhenzhong Chen, King Ngi Ngan. Recent advances in rate control for video coding[J]. *Signal Processing: Image Communication*. 2007, 1, 22 (1): 19-38.
- [31] ITU Telecom. Standardization Sector of ITU[S]. Geneva: ITU-T. Advanced Video Coding for Generic Audiovisual Services, ITU-T Recommendation H.264 and ISO/IEC 14496-10 AVC Standard. 2007.
- [32] Heinrich-Hertz-Institute. JVT Reference Software Version JM10.2[CP/OL]. [2011-11-10]. http://iphome.hhi.de/suehring/tml/download/old_jm/jm10.2.zip.
- [33] Leontaris A, Tourapis A L. Rate Control Reorganization in the Joint Model (JM) Reference Software[DB/OL]. California: Dolby Laboratories, Inc, April, 2007. [2011-11-10]. http://wftp3.itu.int/av-arch/jvt-site/2007_04_SanJose/JVT-W042.zip.
- [34] Wook Joong Kim, Jong Won Yi, Seng Dae Kim. A bit allocation method based on picture activity for still image coding[J]. *IEEE Transactions on Image Processing*. 1999, 7, 8(7): 974-977.
- [35] Xuan Jing, Lap-Pui Chau, Wan-Chi Siu. Frame Complexity-Based Rate-Quantization Model for H.264/AVC Intra-frame Rate Control[C]. *IEEE Signal Processing Letters*. 2008: 373-376.
- [36] Xuan Jing, Lap-Pui Chau. A Novel Intra-Rate Estimation Method for H.264 Rate Control[J], *IEEE International Symposium on Circuits and Systems*, 2006, 5. 5019-5022.
- [37] 王海婴, 张新发. I 帧 QP 设置算法及其在场景变换中的应用[J]. *北京邮电大学学报*, 2008, 31(16): 99-103.
- [38] 韩 霜, 林其伟, 张伟龙等. H.264/AVC 码率控制中初始量化参数设置的一种改进算法[J]. *微型机与应用*, 2010, 29(20): 36-38.
- [39] 常 侃, 覃团发, 门爱东. 一种帧内码率-量化模型及其应用[J]. *广西大学学报*, 2010, 35(6): 1018-1026.
- [40] 李 明, 常义林, 杨付正, 万 帅. 一种自适应的关键帧量化参数确定方法[J]. *西安电子科技大学学报(自然科学版)*. 2010, 6, 37(3): 396-404.
- [41] 周 贤, 姜 威. 基于边缘能量的自动聚焦算法[J], *光学技术*, 2006, 32(2): 213-218.
- [42] Bjøntegaard G. Calculation of Average PSNR Differences Between RD Curves [DB/OL]. [2011-11-10]. http://wftp3.itu.int/av-arch/video-site/0104_Aus/VCEG-M33.doc, 2001.
- [43] Pateux S, Jung J. An Excel Add-in for Computing Bjøntegaard Metric and Its Evolution[DB/OL]. [2009-04-15]. http://wftp3.itu.int/av-arch/video-site/0701_Ma

- r/VCEG-AE07.zip, 2007.
- [44]Tihao Chiang, Ya-Qin Zhang. A New Rate Control Scheme Using Quadratic Rate Distortion Model[J]. IEEE Transactions on Circuits and Systems for Video Technology. 1997, 2, 7(1):246-250.
- [45]Hsueh-Ming Hang, Jiang-Jone Chen. Source Model for Transform Video Coder and Its Application-Part I: Fundamental Theory[J]. IEEE Transactions on Circuits and Systems for Video Technology. 1997, 4, 7(2): 287-298.
- [46]Ning Wang, Yun He. A New Bit Rate Control Strategy for H.264[C]. CICS-PCM. 2003, 12, 3: 1370-1374.
- [47]万 帅, 常义林, 杨付正, 罗 忠. 一种新的视频编码二次率失真模型及其性能分析[J]. 电子与信息学报. 2007, 5, 29(5): 1136-1139.
- [48]Zhihai He, Yong Kwan Kim, Sanjit K. Mitra. Low-delay rate control for DCT video coding via ρ -domain source modeling[J]. IEEE Transactions on Circuits and Systems for Video Technology. 2001, 8, 11(8): 928-940.
- [49]Hung-Ju Lee, Tihao Chiang, Ya-Qin Zhang. Scalable rate control for MPEG-4 video[J]. IEEE Transactions on Circuits and Systems for Video Technology. 2000, 9, 10(6): 878-894.
- [50]Anthony Vetro, Huifang Sun, Yao Wang. MPEG-4 rate control for multiple video objects[J]. IEEE Transactions on Circuits and Systems for Video Technology. 1999, 2, 9(1): 186-199.
- [51]Jinho Choi. Distortion Policy of Buffer-Constrained Rate Control for Real-time VBR Coders[J]. IEEE Transactions on Image Processing. 1999, 4, 8(4): 537-547.
- [52]Gary J. Sullivan, Thomas Wiegand. Rate-Distortion Optimization for Video Compression[J]. IEEE Signal Processing Magazine. 1998, 11, 15(6): 74-90.
- [53]Hanli Wang, Sam Kwong. Rate-Distortion Optimization of Rate Control for H.264 With Adaptive Initial Quantization Parameter Determination[J]. IEEE Transactions on Circuits and Systems for Video Technology. 2008, 1, 18(1): 140-144.
- [54]戚怡宁. 视频“重帧”和“跳帧”对主观质量的影响[C]. 第十一届中国体视学与图像分析学术会议论文集: 10. 宁波: 中国体视学学会, 2006, 10: 402-406.
- [55]Jordi Ribas-Corbera, Shawmin Lei. Rate Control in DCT Video Coding for Low-Delay Communications[J]. IEEE Transaction on Circuits and Systems for Video Technology. 1999, 2, 9(1): 172-185.
- [56]蔡文霞, 杜明辉. 一种有效的 H.264 码率控制跳帧算法[J]. 计算机科学与技术, 2005, 10, 10(5): 52-55.
- [57]周骏华, 石旭刚, 宋 玮. 一种基于运动复杂度的新的 H.264 码率控制跳帧算法

- [J]. 浙江工业大学学报. 2006, 12, 34(6): 672-675.
- [58]王 晓, 邓 云, 朱 柳. 一种低码率下的 H.264 跳帧算法[J]. 计算机工程. 2010, 12, 36(23): 223-225.
- [59]杨国沐. H.264 中码率控制算法与率失真模型的研究[D]. 成都: 西南交通大学, 2008, 6.
- [60]王喜民, 杨鼎才, 闫俊海. H.264 码率控制跳帧算法研究[J]. 微处理机, 2010, 6, 6(3): 76-79.
- [61]Hongtao Yu, Feng Pan, Zhiping Lin. A new bit estimation scheme for H.264 rate control[J]. IEEE International Symposium on Consumer Electronics. 2004, 9, 9(3): 396-399.
- [62]Minqiang Jiang, Nam Ling. Low-Delay Rate Control for Real-time H.264/AVC Video Coding [J]. IEEE Transactions on multimedia. 2006, 6, 8(3): 467-477.
- [63]Thomas Wiegand. Draft ITU-T recommendation and final draft international standard of joint video specification[A]. (ITU-T Rec. H.264 | ISO/IEC 14496-10 AVC). 7th JVT Meeting and 19th VCEG Meeting [C]. Pattaya: 2003.

攻读硕士学位期间参加的科研项目以及研究成果

攻读硕士研究生期间参加的科研项目以及取得的研究成果如下：

一. 参加的科研项目

1. 国家自然科学基金(项目编码: 60772134), “提高多视点视频编码效率的技术研究”, 参加者。
2. 中兴合作项目, “变焦距相机的自动聚焦的研究”, 主要完成人。
3. 高通合作项目, “由 2 维视频虚拟生成 3 维视频的技术研究”, 主要完成人。
4. 高通合作项目, “3DV 标准化研究”, 参加者。

二. 发表论文情况

- [1] 任江涛, 元辉, 常义林, 任光亮. 视频编码线性码率控制模型[J]. 电讯技术, 2010.50(10): 23-27.
- [2] Hui Yuan, Ju Liu, Jiangtao Ren, Yujun Li. A Linear Model Based Frame Level Rate Control Method for Low Bit Rate Video Coding[C]. submitted to Journal of Electronic Imaging.

三. 专利

- [1] 杜昕, 史立庆, 陈涛, 任江涛, 刘克华. 一种基于图像处理的快速自动聚焦方法 申请中
- [2] 杜昕, 史立庆, 陈涛, 任江涛, 刘克华. 基于场景中感兴趣对象的自动聚焦与跟踪算法 申请中

