

基于内容复杂度的视频编解码率控制

唐田野 刘钊

(电子科技大学电子工程学院 成都 610054)

摘要 该文根据线性模型提出信源内容复杂度的一种度量方式,并以此进行帧级比特分配和宏块级的视频码率控制。仿真表明该方法与 MPEG2 中的 TM5 相比,能比较显著地提高信噪比,实现更加准确的码率控制,同时可以在帧间和帧内获得更一致的视觉效果。该算法在 MPEG4 及低码率视频编码中也是适用的。

关键词 视频编码, 码率控制, 内容复杂度

中图分类号: TN919.8

文献标识码: A

文章编号: 1009-5896(2006)06-1115-05

Rate Control Based on Content Complexity in Video Coding

Tang Tian-ye Liu Zhao

(Electronic Engineering School, University of Electronic Science and Technology of China, ChengDu 610054, China)

Abstract A new video rate control scheme which uses linear source model as content complexity measure of video frames is proposed in this paper to implement bit allocation between frames and rate control over Micro Btoctc(MB) level in one frame. Comparing with TM5, simulation results show that the proposed method can improve PSNR obviously and achieve more accurate rate control as well as more smooth objective and subjective visual quality even when scenes change. This algorithm is also applicable in MPEG4 and low rate video coding.

Key words Video coding, Rate control, Content complexity

1 引言

随着计算机、通信等技术的不断发展,视频应用越来越广泛,出现了许多国际视频编码标准,例如已经广泛应用的 ISO/IEC 的 MPEG2^[1], MPEG4^[2] 以及 ITU 的 H.263^[3] 等。

码率控制是视频编码中一个重要的研究内容,其目的是在满足带宽约束条件下,防止缓冲器的溢出,并且尽量提高图像的质量。一般的码率控制涉及两个方面:帧级目标比特分配和宏块级的速率控制。例如应用在 MPEG2 中的 TM5^[4] 就是这样一种典型的码率控制方案。这是一种反馈类型的码率控制方法,运算复杂度低,在实际应用中具有很好的效果,但也存在某些不足:如不能有效地处理场景变换,同时宏块级速率控制不能保证精确的码率控制,可能导致缓冲器的溢出。

本文提出的码率控制方案依据待编码帧和待编码宏块的内容复杂度进行码率控制。即根据帧的内容复杂度分配相应的帧目标比特;根据宏块的内容复杂度分配相应的宏块目标比特,以此确定宏块的量化参数。仿真表明,本文的算法具有更高的信噪比,更一致的图像质量,同时码率控制更加准确。

2 基于内容复杂度的码率控制

2.1 算法思想

我们知道, TM5 是 MPEG2 中采用的码率控制方案。它

的编码参数强烈依赖已编码帧的信息,性能在帧间会出现波动,这在出现场景变化的时候尤其明显,因为此时过去的参考信息不再有效;另外,在宏块级的速率控制中,它依赖缓冲的空满程度确定初始量化参数,再根据宏块活动特性,确定最终量化参数,这种两步调节的方法也不能实现准确的码率控制。本文的码率控制算法在上述两方面作出改进:(1)在帧级比特分配上,同时考虑待编码帧和已编码帧的复杂度信息,并根据帧类型进行比特分配,以补偿帧间场景变化带来的复杂度变化而造成的算法性能降级;(2)在宏块级的速率控制上,同时考虑待编码宏块和整帧的复杂度信息,根据宏块的相对复杂度进行比特数分配,确定合适的量化参数,达到速率控制的目的。

获得一致平滑的图像质量是很多码率控制方案的一个实现目标^[5,6]。本文提出的码率控制方案,对内容复杂度高的帧和宏块分配较多的比特数,而对内容复杂度低的帧和宏块分配较少的比特数。通过这种“按需分配”的方法,一方面提高对付视频序列帧间场景变化的能力,另一方面也能使得帧内图像的量化参数分布更加集中,达到更一致的视觉效果。

2.2 内容复杂度的度量准则

在 TM5 中,内容复杂度是以编码帧实际产生的比特数及其平均量化参数的乘积来表征的,即 $X = S \times QP$ 。S 表示已编码帧的实际比特数, QP 是量化参数,在采用线性量化类型时(quantization_scale_type=0),有 $1 \leq QP \leq 31$ 。由 MPEG2

标准^[1]知道, S 中包含量化后的DCT系数和边带信息(运动矢量和头信息等), 而量化参数QP用来调节输出码率, 控制输出质量, 它的变化直接影响DCT系数(图像内容)。由于边带信息等与量化参数无关, 这种复杂度不尽合理。为此, 本文寻找一种新的复杂度度量方式。在文献[7]中提到, 编码帧产生的比特数与量化后的DCT系数的非零个数成正比, 即满足下式:

$$E_S(QP) = E(QP) + 0bit \quad (1)$$

$$E(QP) = \sum_{qt} K_{qt} \times NZC_{qt}(QP) \quad (2)$$

其中, $E_S(QP)$ 是编码帧所有比特数的估计值; $E(QP)$ 是由量化后的DCT系数产生的比特数的估计值; $0bit$ 为码流中边带信息的估计值, 如运动矢量、头信息和相关指示信息等, 不受QP直接影响, 根据帧类型一次估计获得; K_{qt} 是比例系数; qt 是宏块类型, 分为帧内宏块(intra_MB)和非帧内宏块(nonintra_MB)两种, $NZC_{qt}(QP)$ 是量化参数为QP时, qt 类型宏块中非零系数个数的总和。

由上可知, 量化后DCT系数的非零个数表征了内容复杂度: 非零的DCT系数越多, 表明该帧(或该宏块)的复杂度越高; 反之, 非零的DCT系数越少, 该帧(或该宏块)的复杂度越低。但是由于非零DCT系数的个数由量化参数确定, 帧和宏块的复杂度的确定必须基于某一参考量化参数 QP_r , 该参数由下节中的式(5)确定。宏块内容复杂度定义为宏块中非零DCT系数个数的加权值, 加权因子即为式(2)中的 K_{qt} , 这实际上是根据宏块类型对宏块复杂度的一种修正。帧的内容复杂度定义为帧中宏块内容复杂度之和。下面的式(3), 式(4)给出宏块内容复杂度和帧内容复杂度的表达式:

$$X_{mb}^j = K_{qt} \times nzc^j(QP_r) \quad (3)$$

$$X_{frame} = \sum_{j=1}^N X_{mb}^j \quad (4)$$

其中 j 是宏块序号, $nzc^j(QP_r)$ 是在 QP_r 为量化参数时, 第 j 个宏块中非零DCT系数的个数, N 是一帧中宏块的数目。可以看到, 按照这种复杂度准则进行的码率控制内在要求不同帧及宏块之间的量化参数具有趋同性, 使得图像视觉质量更加一致。

2.3 级目标比特分配

帧级比特分配是指在编码前, 对视频序列中各帧图像确定一个目标比特数, 目标比特数的选取应当使得编码码率匹配输出(信道)速率, 这一步是实现码率控制的前提。一般来说, 目标比特数的选取与应用有关。在TM5中, 目标比特选取主要考虑的是GOP的结构、帧类型以及已编码帧的复杂度信息; 在面向低延迟或交互式应用中的码率控制方案, 目标比特的选取几乎是一个常数, 也可结合缓冲器的状态, 对目标比特进行适度修正, 例如TMN8^[8]。

针对前面提到的TM5的缺点, 结合2.2节的复杂度准则, 本文提出一种帧级比特分配算法。算法流程描述如下:

步骤1 初始化 为了消除初始化的影响, 便于比较算

法性能, 对于视频序列的前3帧图像(I, P, B)采用TM5方法进行码率控制。

步骤2 码后处理 (a)已编码帧复杂度估计对于已编码图像, 由下式求出参考量化参数 QP_r :

$$QP_r = \arg \min_{1 \leq QP \leq 31} |E(QP) - (S - 0bit)| \quad (5)$$

其中, $E(QP)$ 按式(2)估计获得, S 是当前帧实际比特数, $0bit$ 是帧的边带信息所用比特数。用此 QP_r 对编码图像(这里编码图像是指经过DCT变换后, 并经MPEG2定义的量化矩阵量化后的图像数据)进行逐宏块编码后量化, 得到每个宏块的非零DCT系数的个数 nzc^j , 可求出 qt 类型宏块非零DCT系数个数的总和 NZC_{p-qt} , 令 $X_{p-pt} = \sum_{qt} K_{qt} \times NZC_{p-qt}$ 作为已编码帧的复杂度估计, 其中 pt 表示图像类型, 有 $pt \in \{I, P, B\}$ 。

(b) K_{qt} 的估计 物理含义是DCT系数量化后, 每一个基本编码单位(游程, 类别, 非零值)的平均熵。可以采用线性回归分析的方法估计如下:

$$K_{qt} = \frac{\sum_{i=1}^n (Sbit_{qt}^i \times NZC_{qt}^i)}{\sum_{i=1}^n (NZC_{qt}^i)^2} \quad (6)$$

这里, n 是过去已编码帧数, $Sbit_{qt}^i$ 和 NZC_{qt}^i 分别表示第 i 帧中 qt 类型宏块实际净比特数(不含边带信息比特)和实际非零DCT系数个数。仿真显示, K_{qt} 在编码过程中相当稳定。

步骤3 帧级比特数分配 (a)待编码帧复杂度估计 当前编码图像用 QP_r 实行预量化, 得到非零DCT系数的个数 NZC_{c-qt} , 与步骤2(a)相似, 用加权和 $X_{c-pt} = \sum_{qt} K_{qt} \times NZC_{c-qt}$ 作为待编码帧的复杂度估计。

(b)确定帧级目标比特 在帧级比特数分配中, 既考虑过去同类型编码帧的复杂度, 也结合当前编码帧自身的复杂度。因此, 我们把作为目标比特分配依据的复杂度表示为先前已编码帧复杂度和待编码帧复杂度的加权和, 即

$$X_i = a \times X_{p-i} + b \times X_{c-i} \quad (7)$$

$$X_p = c \times X_{p-p} + d \times X_{c-p} \quad (8)$$

$$X_b = e \times X_{p-b} + f \times X_{c-b} \quad (9)$$

其中 a, c, e 的大小表示过去同类型帧复杂度对比特分配的影响, b, d, e 的大小表示当前帧复杂度对比特分配的影响。对系数的选择基于以下认识: (1)在GOP结构中, 从参考意义上讲, I帧是最重要的帧, 其次是P帧, 最后是B帧; (2)P帧的重要性按其在GOP中的序号确定, 即P1比P2重要, P2比P3重要, 以此类推; (3)所有B帧重要性相同; (4)越重要的帧受当前复杂度影响越大, 即 X_{c-pt} 所占比重越大。得到 X_{pt} 后, 按照TM5相同的方式计算得到目标比特 T ^[4]。

2.4 宏块级的速率控制

帧目标比特分配之后, 接下来确定宏块的量化等级。宏块级的速率控制就是在宏块层次调整量化参数, 使得每帧消耗的实际比特数与目标比特相当, 这一步是实现码率控制的保证。流程如下:

步骤1 初始化 $j=1, T_1=T, NZC_{qt}^1 = \sum_j nzc_{qt}^j$ 。

nzc_{qt}^j 在第 j 个宏块类型是 qt 时, 值为 nzc^j (第 j 个宏块的非零 DCT 系数的个数, 在帧级比特分配的复杂度估计时已经得到); 如果第 j 个宏块类型不为 qt 时, 值为 0。可见 NZC_{qt}^1 获得初始时该帧中 qt 类型宏块非零 DCT 系数个数的总数。

步骤2 宏块复杂度估计 帧中第 j 个宏块的复杂度通过下式计算:

$$X^j = K_{qt} \times nzc^j / \sum_{qt} K_{qt} \times NZC_{qt}^j \quad (10)$$

其中加权因子 K_{qt} 由该宏块的类型确定。

步骤3 宏块目标比特分配 根据宏块的复杂度, 以及剩余比特 T_j , 第 j 个宏块分配的比特数为

$$T^j = T_j \times X^j \quad (11)$$

步骤4 宏块量化参数的确定 根据模型, 第 j 个宏块的量化参数按下式确定:

$$QP_j = \arg \min_{1 \leq QP \leq 31} |K_{qt} \times nzc^j(QP) - (T^j - C^j)| \quad (12)$$

C^j 是第 j 个宏块的边带信息比特数, 可由已编码的同类型宏块的边带信息比特数统计获得。

步骤5 更新参数 第 j 个宏块量化编码后, 统计实际产生的比特数 S^j , 更新相关参数。

$$T_{j+1} = T_j - S^j, NZC_{qt}^{j+1} = NZC_{qt}^j - nzc_{qt}^j, j = j + 1。$$

步骤6 循环条件 如果 $j \leq N$ (N 表示每帧图像中宏块的数目), 转至步骤2; 否则结束本帧编码。

可以看出, 上述算法的运算复杂度主要体现在对目标量化参数(包括计算复杂度所需的参考量化参数和宏块的实际量化参数)的确定。在实现中, 在确定初始参考量化参数时, 可采用前一同类型编码帧的平均量化参数作为基准, 由于图像序列的相关性, 此量化参数非常接近最终要确定的参考量化参数, 这样, 只需要很少的搜索次数, 就可以获得初始参考量化参数; 在确定宏块实际量化系数时, 基准是上一已编码宏块的量化参数, 在此附近搜索, 也可以很快获得所需量化参数。在仿真测试中, 几乎 70% 的目标量化参数分布在基准量化参数 ± 1 的范围内, 95% 以上分布在 ± 2 的范围内。在确定量化参数时, 这种方法比在文献[5]中采用直方图统计的方法在计算量上有较大减少。

3 仿真结果

实验选取具有较快运动场景和(或)较高空间复杂度的 3 个标准 MPEG 视频序列 (“football”, “flower garden” 和 “mobile”) 作为测试序列^[2], 格式为 SIF, 码率 R_{CBR} 分别为 3, 4, 5 Mbps, 测试帧数为 115。测试时, 采用 $N=12, M=2$ 的固定 GOP 结构, 缓冲器大小设为 $R_{CBR} \times 0.2$, 式(7)-式(9)中参数分别为 $a=0.3, b=0.7, c_i=0.8+0.5 \times (i-1), d_i=0.2-0.5 \times (i-1)$ (i 是 P 帧序号, $i \in \{1, 2, 3\}$), $e=1, f=0$ 。另外, 为避免初始化条件不同的影响, 客观比较两

种算法性能, 在测试中, 本文的算法对所有序列的前 3 帧的编码也采用 TM5 的方式。

我们从 PSNR、帧级比特分配性能、宏块级速率控制性能以及帧内视觉质量一致性 4 个方面给出本文的码率控制算法与 TM5 的比较结果。

表 1 是两种算法的峰值信噪比测试结果, 可以看到, 本文的算法比 TM5 性能上有较大提高。

表 1 3 个标准视频序列(码率为 3,4,5 Mbps)平均 PSNR(dB)的测试结果比较

Tab.1 Average PSNR comparison between two rate control algorithms at 3,4,5 Mbps

| 测试序列 (Mbps) | TM5 | 本文方法 |
|---------------|-----|-------|
| football | 3 | 34.86 |
| | 4 | 36.92 |
| | 5 | 38.91 |
| flower garden | 3 | 33.59 |
| | 4 | 35.89 |
| | 5 | 37.98 |
| mobile | 3 | 32.12 |
| | 4 | 34.41 |
| | 5 | 36.41 |

图 1 和图 2 是本文的帧级比特分配算法与 TM5 之间的比较结果。图 1 中, 测试序列 fb_mo 是用 “football” 的前 60 帧和 “mobile” 的后 55 帧拼接而成。图 2 中的测试序列 mo_fg 是用 “mobile” 的前 60 帧和 “flower garden” 的后 55 帧拼接而成。图 1 可以认为由简单场景切换到复杂场景, 在切换点, 本文的目标比特分配算法较 TM5 信噪比有 2dB 左右的提高。这是因为 TM5 在比特分配时对复杂场景的复杂度仍然以简单场景的复杂度进行估计, 造成在切换点性能陡降。图 2 可以认为由复杂场景切换到简单场景, 在切换点, 本文的目标比特分配算法信噪比较 TM5 略有损失(约为 0.2dB)。这是因为 TM5 在比特分配时把简单的场景当作复杂场景, 分配了过多比特, 这会造成后续帧可分配比特数减小, 引起信噪比损失。因此, 本文的方法能有效地对付场景变换, 维持更加一致的帧间视觉效果。

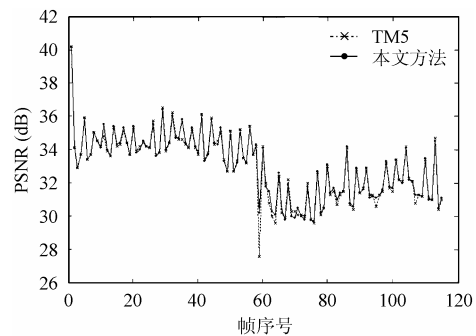


图 1 在 3M 码率时, 由 football 序列向 mobile 序列切换时目标比特分配算法 PSNR 的比较

Fig.1 PSNR comparison between two bit allocation algorithms when scenes change from “football” to “mobile” at 3Mbps

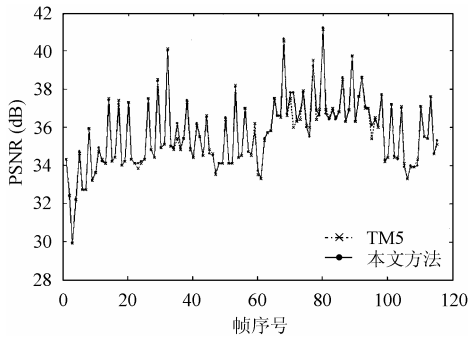


图2 在5M码率时,由mobile序列向flower garden序列切换时目标比特分配算法PSNR的比较

Fig.2 PSNR comparison between two bit allocation algorithms when scenes change from “mobile” to “flower garden” at 5 Mbps

图3是两种算法对4Mbps的“flower garden”序列的速率控制比较结果。定义rate_offset如下:

$$\text{rate_offset} = (T - S) / S \quad (13)$$

T 是编码帧目标分配比特数, S 是编码帧实际产生比特数。图中的纵坐标rate_ratio定义为 $1 + \text{rate_ratio}$,可以看到,本文算法产生的曲线更加接近1,说明该算法的宏块级速率控制能实现更精确的码率控制。实验表明,在其它码率和其它序列情况下也有类似结果。

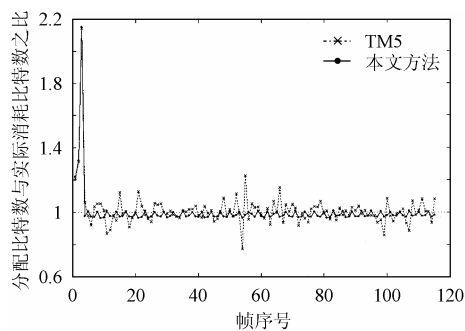


图3 flower garden序列在4M码率时,比特分配与实际比特数之比的比较

Fig.3 Comparison of the ratios of target bits to consumed bits between two rate control algorithms(“flower garden” at 4Mbps)

图4-图6是两种算法D(MSE)的比较结果,其中D(MSE)定义如下:

$$D(\text{MSE}) = \frac{1}{N} \sum_{i=1}^N (\text{MSE}_i - \overline{\text{MSE}})^2 \quad (14)$$

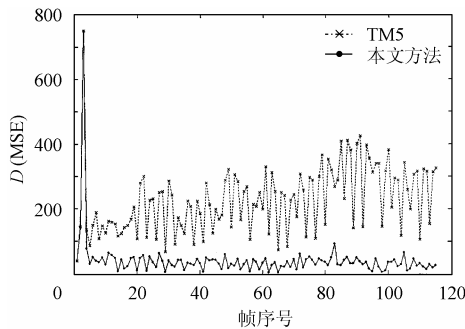


图4 football序列在4M码率时,帧内宏块失真度方差的比较

Fig.4 D(MSE) comparison of distortions of MBs within one frame between two rate control algorithms(“football” at 4Mbps)

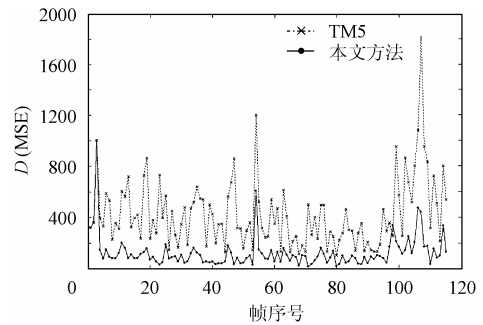


图5 flower garden序列在4M码率时,帧内宏块失真度方差的比较

Fig.5 D(MSE) comparison of distortions of MBs within one frame between two rate control algorithms(“flower garden” at 4Mbps)

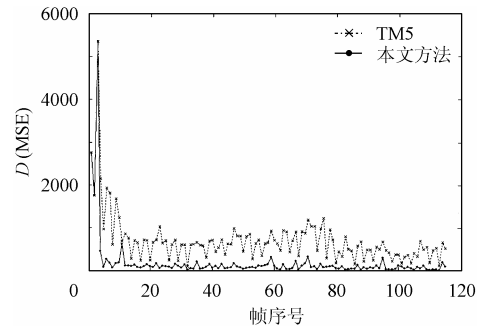


图6 mobile序列在4M码率时,帧内宏块失真度方差的比较

Fig.6 D(MSE) comparison of distortions of MBs within one frame between two rate control algorithms(“mobile” at 4Mbps)

N 是每帧中宏块的数目, MSE_i 是第*i*个重建宏块与原始宏块之间的均方误差,反映出第*i*个宏块的失真度。可以看出,本文的算法具有较小的D(MSE),即帧内宏块失真度分布更加集中,说明视觉效果更加一致。事实上,由于MPEG2中DCT系数预先已经过符合HVS特性的量化矩阵量化,如果量化参数分布相对集中,更符合人眼视觉特性,本文的码率控制算法可以很好地做到这一点。

4 结束语

本文的码率控制方案,在帧级和宏块级都对比特数实现基于内容复杂度的“按需分配”,在提高信噪比的同时,隐含了对量化步长的平滑,使得各帧的量化参数分布更加集中,图像视觉质量更加一致。如果把宏块的码率控制方法应用到视频对象,本文的方法同样适用于MPEG4的编码。另外,由于宏块间量化参数变化对编码比特数有“惩罚”作用^[8],即这种变化需要消耗额外的比特数来表示,这在低码率应用中对编码性能的影响尤为明显。因此,在此类应用中,具有分布集中的量化参数是一个期望的特性,本文的方法也可以较好地解决这个问题。

参考文献

[1] ISO-IEC/JTC1/SC29/WG11. Generic coding of moving pictures and associated audio information: Video. ISO-IEC 13818-2, Nov. 1994.

- [2] ISO-IEC/JTC1/SC29/WG11. Coding of Audio/Visual Objects: Vide. ISO-IEC 14496-2, 1999.
- [3] ITU-T, Study Group 15. Draft of recommendation H.263: Video coding for low bit rate communication. May 2, 1996.
- [4] Document ISO-IEC/JTC/SC29/WG11. Test Model 5. Draft, Apr. 1993.
- [5] Hong S H, *et al.*. Rate control of MPEG video for consistent picture quality. *IEEE Trans. on Broadcasting*, 2003, 49(1): 1–12.
- [6] Hoang D T, Vitter J S. *Efficient Algorithms for MPEG Video Compression*. New York, USA: John Wiley & Sons, 2002: 212–213.
- [7] He Z H, Mitra S K. A linear source model and a unified rate control algorithm for DCT video coding. *IEEE Trans. on Syst. Video Technol.*, 2002, 12(11): 970–982.
- [8] Corbera J R, Lei S M. Rate control in DCT video coding for low-delay communications. *IEEE Trans. on Syst. Video Technol.*, 1999, 9(1): 172–185.
- 唐田野: 男, 1976 处年生, 博士生, 研究方向为视频信号处理及多媒体通信.
- 刘 钊: 男, 1943 年生, 教授, 博士生导师, 从事图像处理及传输等方面的研究工作.