

MPEG-4验证模型中的位率控制策略^{*}

周兵 李晓强 李炜

(郑州大学信息工程学院计算机应用研究室 郑州450052)

(北京航空航天大学计算机学院数字媒体室 北京100083)

Rate Control Strategy of MPEG-4 Verification Model

ZHOU Bing LI Xiao-Qiang LI Wei

(School of Information Engineering, Zhengzhou University, Zhengzhou 450002)

(Digital Media Laboratory, School of Computer, Beihang University, Beijing 100083)

Abstract As a new generation of multimedia coding standard, MPEG-4 aims to satisfy all kinds of low bitrate applications. Accordingly some new coding ideas and technologies are introduced into MPEG-4. This paper discusses the new rate control strategy recommended by MPEG-4 Verification Model—a rate control strategy based on quadratic rate distortion model, and then analyzes and evaluates the strategy, finally some suggestions for improving the strategy are proposed.

Keywords MPEG, Rate control, Quadratic modeling, Buffer, Quantization

1 引言

始于20世纪70年代的数字视频处理技术经过三十多年的发展,取得了很大的进展,其中国际标准化组织 ISO/IEC 下属的 MPEG 专家组制定的标准系列包括已经制定完成的 MPEG-1、MPEG-2、MPEG-4、MPEG-7 标准以及正在制定当中的 MPEG-21 标准。严格说来, MPEG 系列标准中只有前三个属于多媒体编码标准。其中 MPEG-1 标准已经在 VCD 应用中取得了成功, MPEG-2 标准也在 DVD、DVB(数字视频广播)以及 HDTV(高清晰度电视)中得到了广泛应用。MPEG-4 作为一个较新的多媒体编码标准,其目标是通过采用一些新的编码思想和技术实现更高的压缩效率。为此, MPEG-4 相对于 MPEG-1 和 MPEG-2 做了较大的改动,本文将讨论其中的位率控制策略。

位率控制的精度将直接影响编码效率以及重建图像的质量,因此位率控制与缓冲区调整问题一直是视频编码中需要研究的一个重要问题^[6]。在 MPEG-1、MPEG-2 视频编码器中采用的位率控制策略是基于一个一阶率失真模型^[7],它们通过改变编码器的量化参数来获得位率的自适应。这些控制策略中因为对图像内容考虑不够充分,并且使用固定的帧率,因此位率控制精度不够高,特别是在场景变化时质量的下降更加明显。

为了提高位率控制的精度和鲁棒性, Ya-Qin Zhang 和 Tihao Chiang 引入了一个二阶率失真模型^[8],并在该模型的基础上提出了一种新的位率控制策略,该策略已被 MPEG-4 标准采纳。这一位率控制策略具有较广泛的适用性:在应用方面,既可满足无延时限制的 VBR 应用也可满足缓冲区受限的低延时 CBR 应用;在位率方面,既可适用于 10kbps 的低带宽也可适用于 10Mbps 的高带宽;在编码器方面,既可采用基于 DCT 变换的编码器也可以采用基于 Wavelet 变换的编码器^[9]。

2 MPEG-4验证模型中的位率控制策略

MPEG-4 验证模型中的位率控制策略同时实现了帧级、

对象级和宏块级的位率控制^[10],在低延时和缓冲区大小受限条件下,能够获得比一阶率失真模型更精确的目标位数分配。

虽然在 MPEG-4 标准中定义了基于对象的编码方式,但是在实际编码过程中,实时分割视频序列中的对象难度较大,因此可以简单地将整幅图像帧作为一个对象,这样既利用了标准中对象的概念,又保留了 MPEG-1/2 中的图像帧概念,所以本文在讨论位率控制策略时将略过对象级的位率控制。

MPEG-4 验证模型中的位率控制策略可分为四个阶段:初始化阶段、预编码阶段、编码阶段以及后处理阶段^[11],其执行流程如图 1 所示。

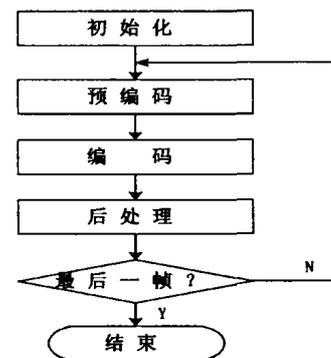


图1 MPEG-4位率控制策略流程图

2.1 初始化阶段

在初始化阶段,编码器主要完成一些与位率控制相关的准备工作,这些工作包括:(1)采用初始量化参数编码第一个 I 帧。(2)从总位数减去第一个 I 帧的编码位数,得到编码当前帧时的剩余位数。(3)基于延时的需求设置缓冲区大小。(4)将缓冲区初始化到半满状态。

2.1.1 编码第一个 I 帧 由于在编码任一视频序列时,第一帧无一例外都是 I 帧,而第一帧在编码时又没有任何的参考信息,因此对序列中的第一帧直接使用初始量化参数进行编码。初始量化参数的选取可以根据具体应用的目标位率

^{*} 国家高技术研究发展计划专项经费(2001AA114142)基金资助。周兵 博士,副教授,研究方向:数字视频处理。李晓强 硕士研究生。李炜 博士,研究方向:视频压缩。

和目标帧率而定。因为量化参数的范围是1~31,所以初始量化参数一般取其中值15。

2.1.2 计算当前帧编码时的剩余位数 不失一般性,假设视频序列的编码顺序是第一个为I帧,随后都是P帧,在该阶段编码器使用初始量化参数QP编码第一个I帧,此时编码随后的P帧可用位数为:

$$P_0 = T \cdot r - I \quad (1)$$

其中 P_0 是编码随后的P帧可用位数, T 是视频序列持续的周期(以秒为单位), r 是序列的输出位率(以位/秒为单位), I 是用于第一个I帧的位数。这样信道输出率是 P_0/N ,其中 N 是一个序列或GOP中的P帧数。

2.1.3 缓冲区设置 在这一步中将设置缓冲区空间的大小,缓冲区空间的设置是根据用户指定的延迟,在MPEG-4标准中缺省的缓冲区大小 $\gamma = r \cdot 0.5$ (即最大累积延时是500ms)。

2.1.4 缓冲区初始化 为了更好地进行缓冲区管理,使其在编码过程中不会频繁地出现上溢和下溢的情况,初始缓冲区占空比被设置成半满状态(即 $r \cdot 0.25$)。

2.2 预编码阶段

在预编码阶段,位率控制策略的任务是根据时域相关性、缓冲区充满度等参数估算出当前帧编码所需的目标位数,并根据这一目标位数为随后的编码阶段计算出合适的量化参数,具体步骤包括:(1)目标位数的初步估计;(2)目标位数调整;(3)量化参数计算。

2.2.1 目标位数的初步估计 这种估计可以分为三级,由粗到细分别为:帧一级的位数估计;对象一级的位数估计(可选项);宏块一级的位数估计(可选项)。本文为简单起见,仅讨论MPEG-4标准中的必选项——帧一级的目标位数估计。

在帧一级, $t+1$ 时刻P帧的目标位数 B_{t+1} 可用下面公式估计:

$$B_{t+1} = \frac{P_t}{N_t} \cdot (1-S) + A_t \cdot S \quad (2)$$

$$P_t = P_0 - \sum_{i=0}^{t-1} A_i \quad (3)$$

其中 P_t 是 t 时刻的剩余位数,使用(3)式进行计算, N_t 是 t 时刻剩下的P帧数, A_t 是 t 时刻P帧使用的实际位数(即前一个P帧), S 是决定前一帧对当前帧目标位数影响的权重因子,可以动态改变,也可以设为一个常量。一般的标准算法中 S 的缺省值被设为0.05。

2.2.2 目标位数的调整 根据(2)式估计的目标位数只考虑了每帧的平均分配位数及两帧之间的相关性对编码的影响,为了获得更精确的目标位数,我们还需要考虑缓冲区的充满度,因此目标位数可以使用下面的公式进一步调整。

$$B_t = \frac{F_t + 2 \cdot (\gamma - F_t)}{2 \cdot F_t + (\gamma - F_t)} \cdot B_t \quad (4)$$

其中 F_t 是 t 时刻缓冲区的充满度, γ 是缓冲区大小。式(4)所进行的调整可以保持缓冲区基本处于半满状态,从而减小上溢和下溢的危险。

为了使每一帧图像获得恒定的质量,编码器还必须保证每一帧有一个最小的目标位数 r/r_f ,其中 r 和 r_f 分别是编码的位率和源视频的帧率,即

$$B_t = \max \left\{ \frac{r}{r_f}, B_t \right\} \quad (5)$$

最后一步的调整是用来预测估计目标位数 B_t 对未来缓冲区充满度的影响,即为缓冲区确保一个安全的空白区域 m ,用来避免缓冲区可能的上溢和下溢,该值在编码前预置。

为了避免缓冲区上溢,需要判断缓冲区中原有的位数 F_t 与当前帧编码后产生的位数 B_t 之和是否超出了缓冲区的安全区域,即不等式 $(B_t + F_t) > (1-m) \cdot \gamma$ 是否成立?如果成立,那么目标减小到:

$$B_t = (1-m) \cdot \gamma - F_t \quad (6)$$

反之,为了避免缓冲区下溢,需要判断不等式 $(B_t + F_t - C) < m \cdot \gamma$ 是否成立?如果成立,那么目标位数增加为:

$$B_t = C - F_t + m \cdot \gamma \quad (7)$$

其中 $C (= P_0/N)$ 是信道输出率。

2.2.3 量化参数计算 采用Ya-Qin Zhang和Tihao Chiang引入的R-D模型。

$$\frac{B_t - H_t}{M_t} = \frac{a_1}{Q_t} + \frac{a_2}{(Q_t)^2} \quad (8)$$

其中 t 表示某一时刻, B_t 是由(4)~(7)式得到的当前帧编码的目标位数, H_t 是 t 时刻图像的语法、运动向量等编码使用的头信息位数, M_t 是 t 时刻图像的MAD值。

当 $a_2 = 0$ 或二阶无解时,退化为一阶模型时,

$$Q_t = \frac{M_t}{B_t - H_t} \cdot a_1 \quad (9)$$

否则

$$Q_t = \frac{M_t \cdot a_1 + \sqrt{(M_t \cdot a_1)^2 + 4 \cdot (B_t - H_t) \cdot M_t \cdot a_2}}{2 \cdot (B_t - H_t)} \quad (10)$$

然后对 Q_t 值进行修整(clip)使其在1-31之间。

在使用(9)(10)两式计算量化参数时,因为无法得到当前编码帧的 H_t ,所以在实际计算过程中,一般使用前一编码帧的 H_{t-1} 。

2.3 编码阶段

在编码阶段,编码器需要完成的任务主要包括:(1)编码视频帧并记录实际的编码位数;(2)如果需要的话,激活宏块级的位率控制。

在编码阶段,帧一级的编码直接采用预编码阶段计算的QP值,量化方法同MPEG-1/2。但是一些低延时的应用可能需要对缓冲区和累积延时进行更加严格的控制,这时就需要进行宏块一级的位率控制,宏块一级位率控制的主要缺点是:①计算复杂度增加;②在低位率应用中,位率控制开销太大,这是因为在一帧内量化参数的变化会带来额外的头信息。

2.4 后处理阶段

在后处理阶段,编码器需要完成如下任务:(1)更新二阶R-D模型和(2)跳帧控制来防止潜在的缓冲区上溢和下溢。

2.4.1 R-D模型更新 在当前帧编码阶段完成之后,编码器必须为下一帧的编码更新R-D模型,为了使模型能够更准确地反映视频内容的变化,R-D模型的更新过程包括如下三个步骤:

步骤1(数据点选择) 这里所谓的数据点:就是被用作更新R-D模型的数据集,该数据集的质量与数量对于模型的精度非常重要。从数据集的数量来说数据集中数据点的数量越多,通常得到的模型也就越精确,但是运算复杂度也随之增加;从数据集的质量来说,当前数据点的一些客观特征值(如MAD或SAD)可以被使用,综合考虑这些因素,给出一个基于滑动窗口的数据选择机制。

该滑动窗口机制被用来自适应地平滑场景变化对R-D模型更新的不利影响。如果场景变化剧烈,则使用一个只包括最近数据点的小窗口。通过使用这种机制,编码器能够智能地为R-D模型更新过程选择那些具有代表性的数据点。数据点的选择是基于当前帧的“场景变化”比例,为定量场景的变化,可以使用一些特征值如MAD或SAD。

在MPEG-4位率控制中,为了降低实现的复杂度,一般使

用 MAD 作为场景变化的一个度量。如果一段视频有较剧烈的运动场景(即编码复杂度增加),那么只有最近的少量数据点被选择;反之,对于缓慢运动的视频,较多的数据点被选中。因此得到下面的算法:

滑动窗口的大小 =

$$\begin{cases} \frac{MAD_t}{MAD_{t-1}} \cdot \text{MAX_SLIDING_WINDOW} & \text{若 } MAD_{t-1} > MAD_t \\ \frac{MAD_{t-1}}{MAD_t} \cdot \text{MAX_SLIDING_WINDOW} & \text{其它} \end{cases}$$

其中 t 是编码的某一时刻, MAX_SLIDING_WINDOW 是一个预置的常数,可以根据需要设置,一般最大不超过 20。

步骤 2(模型参数 a_1 和 a_2 的计算) 根据被选数据点的实际量化值和编码位数,使用线性递归技术可以得到两个模型参数 a_1 和 a_2 。

$$a_2 = \frac{n \sum_{i=1}^n \frac{A_i - H_i}{M_i} - (\sum_{i=1}^n Q_i^{-1}) (\sum_{i=1}^n Q_i \cdot \frac{A_i - H_i}{M_i})}{n \sum_{i=1}^n Q_i^{-2} - (\sum_{i=1}^n Q_i^{-1})^2} \quad (11)$$

$$a_1 = \frac{\sum_{i=1}^n Q_i \cdot \frac{A_i - H_i}{M_i} - a_2 \cdot Q_i^{-1}}{n} \quad (12)$$

其中 n 是被选中的帧数, Q_i 、 H_i 、 M_i 和 A_i 分别是各帧实际的平均量化值、头信息与运动向量的编码位数、图像复杂度特征值和各帧的实际编码位数。

步骤 3(从数据集中去掉外点) 在得到新的模型参数 a_1 和 a_2 之后,编码器通过步骤 3 来去除“坏”的数据点。从数据集中去除这些外点,能够减少“坏”数据点对模型更新过程的影响。“坏”数据点的定义是:“实际位率与目标位率的差别大于某个阈值”。目标位率可以使用步骤 2 中获得的新模型参数进行计算,实际编码位数是 A_i ,实际量化级别为 Q_i 。通过应用这一方法,更多具有代表性的数据点被选择,最终的模型参数可以基于这些新的数据点计算(注:为避免删除所有的数据点,最新的数据点总是在数据集中)。在实际的编码过程中,为了提高模型的精度可以多次迭代执行步骤 2 和 3。

2.4.2 跳帧控制 其目的是防止缓冲区上溢。一旦编码器预测出下一帧将引起缓冲区上溢,编码器将跳过下一帧的编码。缓冲区的充满度将因帧率的降低而减小。虽然跳帧是防止缓冲区上溢的一个有效方法,但视频的整体感知质量可能会因此而下降,特别是在连续跳帧的情况下,为了既能够有效地防止缓冲区上溢,同时又能够使因连续跳帧引起的视频质量的下降最小化,可采用以下跳帧机制。

在编码下一帧之前,首先检查当前的缓冲区占用情况,并且为下一帧估计编码目标位数。如果当前缓冲区的充满度加上下一帧的估计目标位数大于某一个预定的阈值,例如缓冲区大小的 80%,下一帧将被跳过。使用预定阈值的目的是减少连续的跳帧,预定阈值的大小可以根据编码内容自适应地改变。

在实际的预测跳帧控制中,为了减少运算量,并不真正去估计下一帧的编码目标位数,而是使用上一帧的实际编码位数来替代。这样如果当前缓冲区的充满度与上一帧的实际位数之和减去两帧间隔时间内的信道输出超过了缓冲区的安全空白区,将跳过下一帧。在跳过下一帧之后,缓冲区的充满度还需要减去两帧间隔时间内的信道输出。跳帧的条件可用如下公式表示:

如果不等式 $(F_t + A_t - \frac{P_1}{N} \times \frac{1}{r_0}) \geq (1-m) \cdot \gamma$ 成立,那么跳过下一帧,并且缓冲区的充满度更新为:

$$F_t = F_t - \frac{P_1}{N} \times \frac{1}{r_0} \quad (13)$$

上式中 r_0 表示编码的输出帧率。

3 MPEG-4 验证模型中位率控制策略评价与分析

MPEG-4 验证模型中的位率控制策略通过引入一个全新的二阶率失真模型,实现了较 MPEG2 中 TM5 更加精确的位率控制,在该控制策略中提出了若干新颖的算法:

·在目标位数估计过程中充分考虑了每一帧的平均位数、相邻图像之间的相关性以及缓冲区状态等因素对编码位数的影响,对目标位数进行了数次调整,从而得到了一个更加合理的目标位数,为准确求解量化参数打下了良好的基础;

·在后处理阶段,该策略使用了滑动窗口的方法选择多个数据点对二阶模型进行更新,务求得到一个更加精确的模型,使位率控制的结果更加准确;

·该控制策略中还提出了一个新颖的跳帧方法,既避免了缓冲区的上溢,又保证了图像的主客观质量。

虽然该位率控制策略具有许多优点,但是其中也还有需要改进的方面,比如, MPEG-4 位率控制策略中假设的条件:第一帧为 I 帧,随后的所有帧都编码成 P 帧,这种情况并不符合实际应用。首先,为了便于随机存取和增加压缩码流的鲁棒性,实际编码时一般都需要将序列化分为图像组(GOP),从而使 I 帧图像以一定的频率出现;其次,在实际编码过程中,为了提高视频序列的时域分辨率,经常会使用到 B 帧编码。

而 MPEG-4 的位率控制策略中没有给出 I、B 帧的位率控制算法。因此在实际应用中,我们可以根据实际需要为 I 帧、B 帧分别建立一个类似于 P 帧二阶率失真模型,需要注意的是: B 帧的模型在初始化时,初始的参数可以从已编码的 P 帧中得到。

以上几点改进建议仅是笔者的一些初步想法,在 MPEG-4 编码器的设计中,读者可以根据实际的需要对位率控制中的相应部分进行必要的改进。

参考文献

- 1 Tzou K-H, Musmann H G, Aizawa K. IEEE Trans On Circuits and Systems for Video Technology. (Special Issue on Very Low Bit Rate Video Coding), 1994, 4: 213~367
- 2 Li W, Zhang Y-Q, Liou M L. Proc. IEEE (Special Issue on Advances in Image and Video Compression), 1995, 83: 135~340
- 3 Li H, Lundmark A, Forchheimer R. Image sequence coding at very low bitrates: A review. IEEE Trans On Image Processing, 1994, 3: 568~609
- 4 Girod B, et al. Recent advances in video compression. In: IEEE International Symposium On Circuits and Systems, Feb. 1996
- 5 Girod B. Advances in digital image communication. In: Proceing 2nd Erlangen Symposium, Erlangen, Germany, Apr. 1997
- 6 Ding Wei, Liu Bede. Rate control of MPEG video coding and recording by rate-quantization modeling. IEEE Trans On Circuits and Systems for Video Technology, 1996, 6(1): 12~20
- 7 MPEG-2 Video Test Model 5. ISO/IEC JTC1/SC29/WG11 MPEG93/457, Apr. 1993
- 8 Chiang T, Zhang Y-Q. A new rate control scheme using quadratic rate distortion model. IEEE Trans. On Circuits and Systems for Video Technology, 1997, 7(1): 246~250
- 9 MPEG-4 Video Verification Model version 15. 0. ISO/IEC JTC1/SC29/WG11 N3093, Maui, Dec. 1999
- 10 Information technology - coding of audio-visual objects, part 1: systems, part 2: visual, part 3: audio. ISO/IEC JTC1/SC29/WG11, FCD 14496, Dec. 1998
- 11 Wu Dapeng, Hou Y T, Zhu W, et al. On End-to-End Transport Architecture for MPEG-4 Video Streaming over the Internet. IEEE Trans On Circuits and Systems for Video Technology, 2000, 10(6): 923~941