第9卷第11期
1998年11月

软 件 学 报
JOURNAL OF SOFTWARE

Vol. 9, No. 11
Nov. 1998

# MPEG 视频编码比特速率控制算法的改进*

陈维强 高 文

(哈尔滨工业大学计算机科学与工程系 哈尔滨 150001)

**摘要** 针对MPEG-2测试模型5(TM5)比特速率控制算法进行了以下改进:①引入一种智能积分作用,有利于提高比特速率的控制精度;②采用DCT交流系数定义宏块和图象的编码难度测度,然后提出了改进的目标编码量分配策略,即根据编码难度测度自适应调整图象和宏块的目标编码量,有利于提高包括场景变化在内的图象解码质量及其一致性.模拟实验结果表明:采用智能积分作用能达到很高的控制精度,与TM5相比,相对误差降低了约一个数量级;采用目标编码量的改进分配策略,平均峰值信噪比(PSNR)有明显提高,而且也能改善由于场景切换所带来的对图象解码质量的影响.

**关键词** MPEG视频编码,比特速率控制,智能积分作用,编码难度测度,目标编码量分配.

**中图法分类号** TP391

输出比特速率的控制是MPEG(moving picture experts group)[1]视频编码的一个重要问题,控制的好坏直接关系到解码图象的质量与编码器输出比特速率的稳定性.而在MPEG标准中只提供了比特速率的控制机制,具体如何进行比特速率的控制,标准中并没有作定义.到目前为止,已发表了一些有关比特速率控制算法的文章.从控制角度看,可分为两类:①前馈比特速率控制[2];②反馈比特速率控制.前馈控制为了达到控制目标,通常都需要精确计算编码器的设置、输出比特速率和图象解码质量之间的关系,这样就必须作许多预处理和预分析.因此,一个控制质量好的前馈比特速率控制算法一般需要很大的计算量,这是前馈比特速率控制的一个致命弱点.正是因为这一弱点,使其应用大受限制,目前基本上不采用这类算法.相对来说,反馈控制算法的计算量是可以接受的,并已广泛应用在目前的MPEG编码器中.[3]MPEG-2 Test Model 5(TM5)[4]算法就是一种反馈比特速率控制算法,它是MPEG进行测试和比较实验的算法.

实际应用中,通常要求MPEG编码器的输出比特速率能达到很高的控制精度,这一点具有重要意义.例如,一方面可以满足码流在给定通道速率上的传输需要,另一方面如果每个图象组GOP的编码量恒定的话,则可便于实现可放的特技模式,如快进、快退等.然而,TM5算法只是一般意义上的平均比特速率控制,并不是精确的控制.文献[5]采用重编码方法来达到比特速率的精确控制,所谓的重编码是指对每幅图象都通过多次反复编码直到实际编码量与要求的目标编码量最接近为止.显然,文献[5]提出的方法计算量很大,实时实现困难.另外,TM5是根据上一个同类型的图象复杂度来分配当前图象的目标编码量,因此当图象序列发生场景切换或局部变化时,图象解码质量就会下降;同时,TM5为同一幅图象内的各个宏块分配相等的编码量,而事实上各宏块的编码难度可能差别较大,因而这一做法是不合理的,会影响同一幅图象内各宏块的解码质量一致性.

本文针对这些问题,在TM5算法的基础上进行了以下改进:①在控制器中引入了一种智能积分作用,用以提高控制精度,而TM5采用的只是比例控制策略;②对目标编码量的分配策略进行改进,主要表现在对$P$,$B$图象能根据其编码难度自适应调整图象目标编码量的分配,并能根据当前图象内各宏块的编码难度自适应调整宏块的目标编码量.

## 1 比特速率控制算法(TM5)的改进

MPEG-2测试模型5中所建议的比特速率控制算法,即本文所说的TM5算法,前一节已提到该算法属于反馈控

* 本文研究得到国家863高科技项目基金和国家"九五"科技攻关计划基金资助.作者陈维强,1968年生,在职博士生,讲师,主要研究领域为运动图象压缩技术,系统控制.高文,1956年生,博士,教授,博导,主要研究领域为多媒体压缩与处理技术,智能计算机接口技术,虚拟现实技术,人工智能应用.

本文通讯联系人:陈维强,哈尔滨150001,哈尔滨工业大学计算机科学与工程系

本文1997-06-02收到原稿,1997-11-04收到修改稿

© 中国科学院软件研究所 http://www.jos.org.cn

制算法. 实际上,TM5 算法可分为 3 个主要步骤:①目标编码量分配:对给定的通道比特速率,根据图象复杂度分配每幅图象的目标编码量(比特数);②确定量化参数的参考值:为切片和宏块确定量化参数的参考值,用以匹配目标编码量,即匹配目标比特速率;③导出实际的量化值:根据宏块视觉活性值(Macroblock Activity)确定每个宏块的实际量化值. 采用 TM5 算法构成的反馈控制系统结构主要由以下 4 个单元组成:①目标编码量分配单元,它的功能是根据编码历史信息分配目标编码量;②控制器,采用比例控制决策控制信号 $q$,保证编码器的比特流按所要求的速率输出;③自适应量化单元,根据控制信号 $q$ 和宏块视觉活性值计算量化级 MQUANT;④MPEG 编码器. 本文对 TM5 算法的改进主要表现在控制器与目标编码量分配单元上.

### 1.1　智能积分作用的引入

TM5 比特速率反馈控制系统经过简化以及线性化后可用一个比例线性控制模型表示. 根据控制理论知识可知,TM5 算法(比例控制)存在静态误差,即 TM5 算法只是一般意义上的平均比特速率控制. 显然,引进积分作用是减小系统静态误差、提高系统控制精度的重要途径. 积分作用在一定程度上模拟了人的记忆特性,它"记忆"了误差的存在及其变化的全部信息. 然而,如果将积分作用直接引入到 TM5 算法中,由于这种积分作用只要误差存在就一直进行积分,所以针对性不强,容易导致积分饱和,且其参数不易选择. 因此,这里提出将一种智能积分作用引入到 TM5 算法中,以减小或消除静态误差.

设 $e$ 为实际编码量与目标编码量的误差,$\Delta e$ 为 $e$ 的一阶偏差. 当 $e\cdot\Delta e>0$ 时,表明实际编码量正在偏离分配的目标编码量,这时加入积分作用可使实际编码量尽快向目标比特分配量回归;而当 $e\cdot\Delta e<0$ 时,表明实际编码量正在向目标编码量回归,此时如果加入积分作用将会减慢回归速度,因此这种情况下可不加入积分作用. 系统通过依靠其惯性向目标编码量回归. 这样,引入积分控制作用的方法实质上是模仿人对目标的控制,故可称为智能积分作用.

在 TM5 算法中引入上述智能积分作用,并结合边界条件,则比特速率控制器的控制原理规则可描述如下(称为改进控制规则).

if $(e[n]\cdot\Delta e[n]<0)$ or $(e[n]=0)$ then

　　$d[n]=d+e[n]$

else if $(e[n]\cdot\Delta e[n]>0)$ or $(\Delta e[n]=0$ and $e[n]\neq 0)$ then

　　$d[n]=d+\alpha\cdot e[n]+\beta\cdot\sum_{i=0}^{n}e[i]$

$q[n]=\frac{d[n]\cdot 31}{r}$　/* 与 TM5 相同,导出控制信号,其中 $r=\frac{2\cdot bit_rate}{picture_rate}$ */

规则中,$d$ 为虚拟缓冲区填满度,$d[n]$ 为当前虚拟缓冲区填满度;$e[n]$ 为当前实际编码量与分配的目标编码量之差,$\Delta e[n]=e[n]-e[n-1]$;$q[n]$ 为控制器输出的控制信号;$\alpha,\beta$ 为常数,可由实验确定;$bit_rate$ 是编码器给定的输出比特速率,即目标比特速率;$picture_rate$ 为图象的帧速率(下同).

### 1.2　目标编码量分配策略的改进

在 TM5 算法中,目标编码量的分配可分成 3 个层次:①分配图象组 GOP 的目标编码量;②分配当前图象的目标编码量;③分配宏块的目标编码量. 下面提出对当前图象的目标编码量和宏块的目标编码量的改进分配策略,即根据当前图象(或宏块)的编码难度来修改目标编码量的分配. 由于 MPEG 是以块为单位对原始图象数据或运动补偿数据作 DCT 变换,然后再进行量化、变长编码,因此,DCT 系数能直接反映编码难度,尤其是交流(AC)系数.

**定义 1.** 宏块编码难度的测度(记为 $D_{MB}$,只以 $Y$ 分量计算)按如下方式定义.

$$D_{MB}=\min_{blk=1\sim 4}(abs_sum) \tag{1}$$

$$abs_sum=\sum_{i=1}^{7}\sum_{j=1}^{7}|32\times coeff_{ij}/W_{ij}| \tag{2}$$

其中 $blk$ 为宏块中的 4 个 8*8 像素块;$coeff_{ij}$ 为 DCT 系数;$W_{ij}$ 是加权矩阵系数.

本文的加权矩阵系数 $W_{ij}$ 取编码器设定的量化矩阵系数 $quant_mat_{ij}$,这样可减少由于计算宏块编码难度测度而引起的编码器计算量增加. 因为式(2)的中间结果,即 $32\times coeff_{ij}/W_{ij}|_{ij\in[1,7]}$ 就是下一步量化过程的中间结果,因此,计算宏块编码难度测度所得到的中间结果可在下一步量化过程中直接利用,而不需重新计算.

**定义 2.** 假设一幅图象的宏块总数为 $N_{MB}$,对某宏块 $i(i\in[0,N_{MB}-1])$,其编码难度的测度值记为 $D_{MB}[i]$,$\Delta D_{MB}$ 为一门限值,且门限 $\Delta D_{MB}$ 对 $I$,$P$ 和 $B$ 图象一般取不同的值,则图象的编码难度测度(记为 $D_{PIC}$)定义如下. 运算 $\lceil x,y\rceil$ 表示:如果 $x>y$,$\lceil x,y\rceil=1$,否则 $\lceil x,y\rceil=0$.

$$D_{PIC}=\left(\sum_{i=0}^{N_{MB}-1}\lceil D_{MB}[i],\Delta D_{MB}\rceil\right)/N_{MB} \tag{3}$$

© 中国科学院软件研究所　http://www.jos.org.cn

1.2.1 当前 $P$ 或 $B$ 图象目标编码量的调整方法

当序列图象的场景发生变化甚至切换时，那么，对 $P$，$B$ 图象来讲，不仅运动补偿较差的宏块数量会增多，而且还会出现一定数量的没有运动补偿的 $I$ 宏块，所以其编码难度就会提高，为此应该适当地增加该图象的目标编码量. 设 $D_{PIC}^{P,B}$ 为当前 $P$ 或 $B$ 图象的编码难度测度值，$\Delta D_{PIC}^{P,B}$ 为 $P$ 或 $B$ 图象的编码难度测度门限值，$X_{P,B}$ 为前一帧 $P$ 或 $B$ 图象的复杂度. 调整方法是：当 $D_{PIC}^{P,B}<\Delta D_{PIC}^{P,B}$ 时，则不做调整(即与 TM5 算法相同)；否则，首先采用式(4)对复杂度进行调整，其中 $X'_{P,B}$ 表示调整后的复杂度，$\Delta X_{P,B}$ 为常数，$\gamma$ 为比例因子；然后，根据 $X'_{P,B}$ 求出调整后 $P$，$B$ 图象的目标编码量 $T_{P,B}$，见式(5)，其中 $R$ 是当前 GOP 的编码余量，$N_P$，$N_B$ 分别是当前 GOP 中包括当前帧在内的剩余 $P$，$B$ 图象数，$K_P$，$K_B$ 为常数(与 TM5 相同). 由式(5)可知，由于当前图象的复杂度增加，其目标编码量 $T_{P,B}$ 也将随之增大.

$$X'_{P,B}=X_{P,B}+\gamma\cdot\Delta X_{P,B} \tag{4}$$

$$T_{P,B}=\max\left\{R/\left(N_{P,B}+\frac{N_{B,P}K_{P,B}X_{B,P}}{K_{B,P}X'_{P,B}}\right),\frac{bit_rate}{8\cdot picture_rate}\right\} \tag{5}$$

1.2.2 宏块的目标编码量调整方法

设 $T_{I,P,B}$ 为 $I$，$P$ 或 $B$ 图象的目标编码量，当前图象各宏块的编码难度测度记为 $D_{MB}^{I,P,B}[i](i\in[0,N_{MB}-1])$，则图象的平均宏块编码难度测度为 $\overline{D}_{MB}^{I,P,B}$，可由式(6)求得.

$$\overline{D}_{MB}^{I,P,B}=\left(\sum_{i=0}^{N_{MB}-1}D_{MB}^{I,P,B}[i]\right)/N_{MB} \tag{6}$$

首先，对各宏块的编码难度测度值进行如下操作(其中 $\Delta MB_{I,P,B}$ 为一个大于 0 的阈值).

if $(D_{MB}^{I,P,B}[i]-\overline{D}_{MB}^{I,P,B})\le-\Delta MB_{I,P,B}$ then $D_{MB}^{I,P,B}[i]=\overline{D}_{MB}^{I,P,B}-\Delta MB_{I,P,B}$

else if $(D_{MB}^{I,P,B}[i]-\overline{D}_{MB}^{I,P,B})\ge\Delta MB_{I,P,B}$ then $D_{MB}^{I,P,B}[i]=\overline{D}_{MB}^{I,P,B}+\Delta MB_{I,P,B}$

else $D_{MB}^{I,P,B}[i]=\overline{D}_{MB}^{I,P,B}$

然后，求出当前图象内宏块的编码难度测度和，记为 $Sum^{I,P,B}$，则有

$$Sum^{I,P,B}=\sum_{i=0}^{N_{MB}-1}D_{MB}^{I,P,B}[i]. \tag{7}$$

最后，当前图象内各宏块的目标编码量 $T_{MB}^{I,P,B}[i](i\in[0,N_{MB}-1])$ 可由下式确定

$$T_{MB}^{I,P,B}[i]=(D_{MB}^{I,P,B}[i]/Sum^{I,P,B})\cdot T_{I,P,B}. \tag{8}$$

由式(8)可以看出，图象各宏块的目标编码量可根据其编码难度自适应调整.

## 2 实验结果与分析

为了检验本文所提出的改进算法的效果，我们组织了 3 组与 TM5 算法的对比实验. 进行实验时采用的测试图象序列是：flower garden 的前 31 帧图象与带场景切换的序列 mobile&calendar-flower garden 共 31 帧，其中前 3 帧是 mobile&calendar 的前 3 帧图象，后 28 帧是 flower garden 的前 28 帧图象. 编码时，图象组 GOP 的长度 $N$ 为 15，相邻参考图象的帧距离 $M$ 为 3，前两组实验分别在输出比特速率为 3Mbits/s，5Mbits/s 和 8Mbits/s 下进行，后一组实验是在输出比特速率为 8Mbits/s 下进行. 限于篇幅，PSNR 曲线仅给出比特速率为 8Mbits/s 下的对比.

(1) 只采用对宏块目标编码量的改进分配策略，其他与 TM5 相同，记为改进算法 1. 在 flower garden 序列上与 TM5 算法的 PSNR 对比实验结果表明(如图 1 所示)，改进算法 1 的 PSNR 值普遍高于 TM5 算法，在 3Mbits/s，5Mbits/s 和 8Mbits/s 下分别平均约高出 0.1、0.3 和 0.4dB，并且平均量化级 MQUANT 普遍下降. 同时，主观实验也表明，改进算法 1 的解码图象质量优良.

(2) 在改进算法 1 的基础上，进一步采用对 $P$，$B$ 图象目标编码量的改进分配策略，记为改进算法 2. 在场景切换序列 mobile&calendar-flower garden 上与 TM5 算法的 PSNR 对比实验结果表明，图象的解码质量得到了改善，一致性得到了提高，在 3Mbits/s，5Mbits/s 和 8Mbits/s 下分别平均约高出 1.0、0.3 和 0.3dB. 图 2 是 8Mbits/s 下的 PSNR 对比实验结果，从中观察到：①在场景切换处，与 TM5 相比，改进算法 2 的 PSNR 得到了提高，并且主观实验也表明解码图象质量得到了改善. ②在 10～15 帧图象上，改进算法 2 的 PSNR 相对于 TM5 有所降低，这主要是因为在场景切换处分配了较多的比特数，导致在后续编码中图象组 GOP 的剩余目标编码量比 TM5 要少. ③在其他地方，改进算法 2 的 PSNR 都普遍高于 TM5，这主要得益于对宏块目标编码量分配策略的改进.

© 中国科学院软件研究所 http://www.jos.org.cn

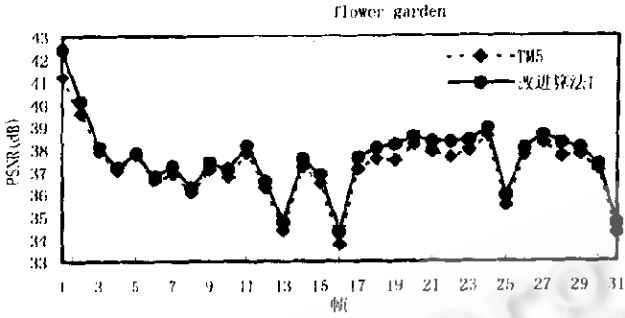


图 1　改进算法 1 与 TM5 算法的 PSNR 对比曲线(8Mbits/s)

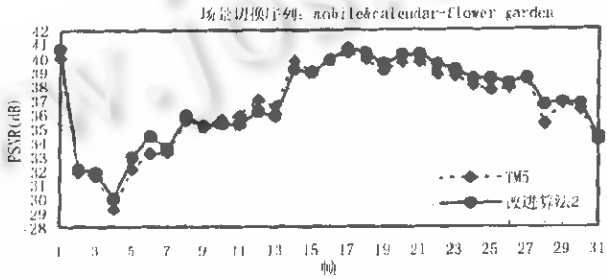


图 2　改进算法 2 与 TM5 算法的 PSNR 对比曲线(8Mbits/s)

(3) 在改进算法 2 的基础上,再引进智能积分作用,记为改进算法 3. 表 1 是改进算法 3 与 TM5 算法的控制精度的比较结果. 表中绝对值误差是指一个图象组 GOP 的实际编码量与理想目标编码量的绝对值差;相对误差=(绝对值误差/理想目标编码量)*100%. 在表中给出的是测试序列的第 2 个 GOP 结果. 从表 1 可以看出,改进算法 3 具有很高的控制精度,与 TM5 相比,相对误差降低了约一个数量级. 实验时发现,虽然在改进算法 3 中也采用了目标编码量的改进分配策略,但其 PSNR 仍然有一定程度的降低. 其原因是:①由表 1 中可知,TM5 算法的 GOP 实际编码量高于改进算法 3;②追求高的比特速率控制精度与追求较好的解码质量是一对互相制约的矛盾,产生这对矛盾的主要原因是现有的编码难度测度定义、编码量分配策略(包括本文所提出的)并不是十分理想,或者说还不够智能,在这方面还有许多工作要做. 我们的下一步工作将是在比特速率的控制精度与图象的解码信噪比(PSNR)上寻求进一步的提高.

表 1　改进算法 3 与 TM5 算法的控制精度比较

| 测试序列 | 理想目标编码量(bits/GOP) | 实际编码量(bits/GOP) 改进算法 3/TM5 | 绝对值误差(bits/GOP) 改进算法 3/TM5 | 相对误差(%) 改进算法 3/TM5 |
|---|---|---|---|---|
| flower garden | 4 000 000 | 3 999 844/4 003 631 | 156/3 631 | 0.004/0.091 |
| mobile&calendar-flower garden | 4 000 000 | 4 002 229/4 020 331 | 2 229/20 331 | 0.056/0.508 |

## 参考文献


1　ISO/IEC 13818-2: Information Technology-Generic Coding of Moving Pictures and Associated Audio: Video. ISO/IEC JTC1/SC29/WG11, 1994. 1～229

2　Shoham Y, Gersho A. Efficient bit allocation for an arbitrary set of quantizers. IEEE Transactions on Acoustics Speech Signal Process, 1988,36(9):231～237

3　Keesman G, Shah I, Gunnewiek R K. Bit-rate control for MPEG encoders. Signal Processing: Image Communication, 1995,6: 545～560

4　MPEG2 Test Model 5. Document ISO/IEC JTC1/SC29/WG11/93-400. Test Model Editing Committee, ISO/IEC, 1993. 5～32


© 中国科学院软件研究所　http://www.jos.org.cn

5 Ding W, Liu B. Rate control of MPEG video coding and recording by rate-quantization modeling. IEEE Transactions on Circuits and Systems for Video Technology, 1996,6(1):12~20

## Improvements on Bit Rate Control Algorithm for MPEG Video Coding

CHEN Wei-qiang GAO Wen

(*Department of Computer Science and Engineering Harbin Institute of Technology Harbin* 150001)

**Abstract** In this paper, based on MPEG-2 Test Model 5 (TM5) bit rate control algorithm, there are some improvements proposed: ① an intelligent integral action is proposed for improving the control precision of bit rate; ② after using DCT AC coefficients to define the measure of macroblock and picture coding difficulty, an improvement on target bits allocation is proposed for the decoded picture quality and conformance including scene change, which can adjust the target bits allocation adaptively by their measure. Simulation results show that it can achieve high control precision by using intelligent integral action, its relative error is reduced one order of magnitude compared with TM5; using the improvement on target bits allocation, the average decoding PSNR is increased visibly, and the decoded picture quality is improved when scene changes.

**Key words** MPEG video coding, bit rate control, intelligent integral action, measure of coding difficulty, target bits allocation.

© 中国科学院软件研究所 http://www.jos.org.cn