

文章编号:1001-9081(2009)03-0675-03

基于 H. 264 标准的码率控制算法

杨 宇,何小海,吕 瑞

(四川大学 电子信息学院,成都 610004)

(cias@263.net)

摘 要:在视频通信过程中,为了使编码端输出码流更好地适应不同的信道,基于 H. 264 视频编码标准提出一种改进码率控制算法。该算法针对标准算法中对基本单元的 MAD 值预测时只考虑时间上相关性的缺点,提出了一种新的结合时间预测和空间预测的预测方法。实验结果表明,与标准算法相比较,该算法能更准确地控制码率,同时视频输出序列的峰值信噪比也得到了提高,编码端输出码流能够更加平稳。

关键词:图像传输;码率控制;平均绝对差;H. 264

中图分类号: TN919.8 **文献标志码:** A

Rate control algorithm based on standard H. 264

YANG Yu, HE Xiao-hai, Lü Rui

(School of Electronics and Information Engineering, Sichuan University, Chengdu Sichuan 610004, China)

Abstract: During the video communication process, an improved rate control algorithm was developed based on H. 264 video coding standard in order to make the output data stream better adapted to different channels. Contrary to the MAD prediction of the basic unit only considering the relationship of the time in the coding standard, a new prediction method was introduced combining time prediction and space prediction. Compared with coding standard, the experimental results illustrate that the new algorithm can control rate more precisely and increase the peak signal noise ratio of the output video stream effectively. It proves that the new algorithm can make the output data stream of the encoder more stable.

Key words: image transmission; rate control; Mean Absolute Difference (MAD); H. 264

0 引言

在图像通信过程中,视频信号经过离散余弦变换、量化、熵编码后输出到信道上进行传输。在编码过程中,如果采用相同的参数设置,就会使得每帧图像编码后的比特数差距较大。造成这种码率波动的原因有两方面:第一,图像帧的类型不同会导致码率的变化。I 帧由于只采用了帧内预测,压缩率一般比较低;P 帧由于采用了帧间预测,通过运用运动搜索和运动补偿等技术之后,压缩率一般较高;而 B 帧采用前后双向预测,压缩率一般比 P 帧更高。第二,场景的复杂程度也决定了码率波动的剧烈程度,一般复杂场景所需的比特数比简单场景要大得多。在相同场景中码率变化不大,但是一旦出现场景转换等情况时,码率就会剧烈增加。然而当比特流在各种有线网和无线网中传播时,网络带宽却是有限的,这显然对网络传输极为不利。这就要求编码器的输出比特流必须控制在一定的范围内以适应网络传输,从而决定了码率控制算法成为视频编码的关键技术之一。

常用的码率控制算法一般是通过调整量化参数 QP 来实现的。当信源的输出比特率大于我们设定的目标比特率,可以增大 QP 值来降低输出比特率,避免信道中出现拥塞现象;当信源的输出比特率小于我们设定的目标比特率,可以减小 QP 值来增大输出比特率,提高传输图像质量。虽然在视频传输过程中增大 QP 值会导致图像质量下降,但是比起信道拥塞引起的图像传输延迟甚至掉包的严重后果,牺牲一部分图像质量是可以接受的。以往的视频压缩标准,如 MPEG-2^[1],

MPEG-4^[2], H. 263^[3] 等都对如何调整 QP 值使码率尽可能和目标码率一致,并且同时能获得较高视频质量进行了广泛而深入的研究。

新一代的 H. 264 视频编码标准^[4]在码率控制中加入了最新的率失真技术真优化(Rate-Distortion Optimization, RDO)^[5]技术,这项技术可以在给定码率的条件下,尽量使得解码的失真度达到最低。然而由于采用了率失真优化技术,导致在实际的编码过程中出现一个新的难题:如图 1 所示,进行率失真优化必须先设定量化参数,而量化参数是在码率控制过程中用平均绝对差(Mean Absolute Difference, MAD)计算得到,但是当前宏块的 MAD 值只能在率失真优化以后才能确定,这就是 H. 264 码率控制当中著名的蛋鸡悖论^[6]。正是因为这个难题,导致了 H. 264 中的码率控制比以往的视频编码标准要难许多。

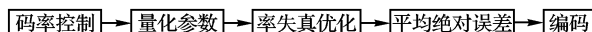


图 1 H. 264 中码率控制示意图

为了解决这个鸡蛋悖论问题,有关专家提出了往返流量模型算法^[7],最终被 H. 264 标准采纳。这种算法引入了基本单元和线性预测模型的概念。码率控制过程中,通过前一帧相同位置基本单元的 MAD 来预测当前基本单元的 MAD,算法复杂度相对较小,对码率的控制效果也比较好。

1 现有的码率控制算法

目前 H. 264 视频编码标准的 JM(Joint Model)中码率控制是从三个层次来实现的:GOP 层、帧层和基本单元层。

收稿日期:2008-09-08;修回日期:2008-10-24。

作者简介:杨宇(1983-),男,四川成都人,硕士研究生,主要研究方向:图像处理、图像通信;何小海(1964-),男,四川盐亭人,教授,博士生导师,主要研究方向:图像处理、图像通信;吕瑞(1982-),男,四川绵阳人,博士研究生,主要研究方向:图像处理、图像通信。

1.1 GOP 层的码率控制

GOP 层的码率控制主要有两个任务:根据网络带宽、帧率以及 GOP 中帧的数目分配当前 GOP 序列中所有帧的目标比特数,以及设置 GOP 层的量化参数。在 H. 264 的参考软件 JM 中,允许用户手动设置初始量化参数 QP_0 ,即图像序列中第一个 GOP 的量化参数。该 GOP 中的 I 帧和第一个 P 帧都使用 QP_0 进行编码。通常情况下,如果可用带宽较大,可将 QP_0 设置较小;可用带宽较小,则可将 QP_0 设置较大。在带宽一定的情况下,如果 GOP 的长度增加 15,则 QP_0 减小 1。

1.2 帧层的码率控制

帧层的码率控制也主要是两个任务:计算各帧的量化参数和进行率失真优化 RDO。由于 B 帧不用作为后续帧的预测图像,为了节省更多比特数来编码相对更重要的 I 帧和 P 帧,它的量化参数可以比相邻的 I 帧和 P 帧更大。为了保证图像视觉质量波动不会过大,相邻两个 B 帧的量化参数的差值最好不能大于 2。而 P 帧的量化参数则是由目标缓冲区满溢程度、帧率、可用信道带宽和实际缓冲区占有程度通过相应公式计算得到的。进行率失真优化时,当前帧中各个宏块的 RDO 则由前面已得的量化参数计算得来。

1.3 基本单元层码率控制

基本单元层码率控制与帧层码率控制类似,首先必须先为基本单元确定目标比特数。由于未编码的基本单元的 MAD 同样未知,所以在分配目标比特时,是将可用的比特数平均分配给所有未编码的基本单元。然后通过线性预测模型,利用前一帧中相同位置基本单元的 MAD 来预测当前基本单元 MAD。记当前要预测的基本单元的 MAD 值为 MAD_C ,前一帧中相同位置基本单元的 MAD 值为 MAD_P ,于是基本单元层的 MAD 预测机制可以由下式表示^[8]:

$$MAD_C = a1 * MAD_P + a2$$

其中 $a1, a2$ 为预测参数,初始值分别为 1, 0。每次编码完成一个基本单元,就对 $a1$ 和 $a2$ 分别进行更新。

最后使用二次率失真模型为当前的基本单元计算量化参数 QP。

2 对 H. 264 中的码率控制模型的改进

2.1 H. 264 中码率控制模型的缺点

由二次率失真模型可以发现为了控制码率 R 保持在一个稳定的状态,各个帧层或者基本单元层的量化差数 Q 值必须随相应的 MAD 值增大而增大,减小而减小,说明对 MAD 预测的准确性是决定码率控制是否准确的一个重要因素。通过考查 H. 264 中的 MAD 线性预测模型,我们发现每一个基本单元的 MAD 值直接由前一帧中相同位置基本单元的 MAD 来预测。这种只考虑时间相关性的预测模型,在图像场景变化平缓的情况下会取得较好的效果,但是一旦图像出现运动剧烈或者是场景转换的情况,其预测的效果就不准确了,从而直接影响码率控制的准确度。

2.2 针对 H. 264 中线性预测模型的改进

针对线性预测模型中只利用了时间相关性的局限性,本文提出的 MAD 预测算法通过综合利用各个基本单元的时间相关性和空间相关性,提出了 4 种 MAD 帧内预测模式:横向 LINE、纵向 ARRAY、斜向左 CROSSRIGHT 和斜向右 CROSSLEFT。

设当前正进行预测的基本单元为 C,那么将 C 对应的 MAD 记为 MAD_C 。如图 2 所示,将 C 的前一个基本单元标为

D, C 的上一行同一列的基本单元标为 B, B 的前一个基本单元标为 A, 后一个标为 F, A 的前一个基本单元标为 E。那么同样将相对应的 MAD 记为 $MAD_D, MAD_B, MAD_A, MAD_F, MAD_E$ 。同时将 JM 标准中对基本单元 C 预测得到的 MAD 值记为 MAD_P 。

E	A	B	F
	D	C	

图 2 相邻基本单元的相互位置

2.2.1 选择预测模式

新算法提供四种可选的预测模式: LINE, ARRAY, CROSSRIGHT 和 CROSSLEFT。设 $Line = \text{abs}(MAD_A - MAD_B)$, $Array = \text{abs}(MAD_A - MAD_D)$, $Crossright = \text{abs}(MAD_E - MAD_D)$, $Crossleft = \text{abs}(MAD_B - MAD_D)$ 。分别计算 Line, Array, Crossright, Crossleft 四个值,然后将比较四个值,选出值最小的那一个。如果 Line 最小,说明 C 单元的横向相关性较好,那么选择 LINE 模式;如果 Array 最小,说明 C 单元的纵向相关性较好,那么选择 ARRAY 模式;如果 Crossright 最小,说明 C 单元的斜向右相关性较好,那么选择 CROSSRIGHT 模式;如果 Crossleft 最小,说明 C 单元的斜向左相关性较好,那么选择 CROSSLEFT 模式。

2.2.2 进行预测

如果选择 LINE 模式:

$$MAD_C = a * MAD_P + b * MAD_B + c * MAD_B + (1 - a - b - c) / 2 * MAD_A + (1 - a - b - c) / 2 * MAD_F$$

如果选择 ARRAY 模式:

$$MAD_C = a * MAD_P + b * MAD_B + c * MAD_D + (1 - a - b - c) / 2 * MAD_A + (1 - a - b - c) / 2 * MAD_F$$

如果选择 CROSSRIGHT 模式:

$$MAD_C = a * MAD_P + b * MAD_A + d * MAD_D + d * MAD_B$$

如果选择 CROSSLEFT 模式:

$$MAD_C = a * MAD_P + b * MAD_F + d * MAD_D + d * MAD_B$$

考虑到当前帧基本单元和前一帧对应预测单元的时间相关性在大多数情况下要大于这个基本单元及其帧内相邻单元的空间相关性,以及在帧内进行空间预测时,相关性好的方向的预测系数肯定应该大于相关性差的方向,于是我们可以给出如下不等式: $a > b > c$ 以及 $a > b > d$ 。在这个范围内,本文通过大量的试验,发现当 $a = 0.6, b = 0.2, c = 0.1, d = 0.1$ 时码率控制效果最好,于是设定 $a = 0.6, b = 0.2, c = 0.1, d = 0.1$ 。

需要说明的是,如果 C 是当前帧的第一个基本单元的话,那么 $MAD_C = MAD_P$ 。如果 C 的位置在当前帧的第一行,那么直接选择 LINE 模式,并且 $MAD_C = a * MAD_P + (1 - a) * MAD_D$ 。如果 C 的位置在当前帧的第一列,那么直接选择 ARRAY 模式,并且 $MAD_C = a * MAD_P + (1 - a) * MAD_B$ 。如果 C 的位置在当前帧的第二列,那么从 Line, Array, Crossleft 三个值中比较出最小的那一个,选择对应的模式进行预测。如果 C 的位置在当前帧的最后一列,那么从 Line, Array, Crossright 中比较出最小值,选择对应的模式进行预测。

2.2.3 计算 QP 值

将 MAD_C 代入 H. 264 标准中的二次率失真公式计算出当前基本单元的 QP 值。

3 码率控制效果测试

3.1 对无场景转换图像序列的测试

本文分别选用了4个 QCIF 格式和4个 CIF 格式的无场景转换图像序列来进行新算法的测试,首先将测试所得的码率 R_2 和用 JM 算法测试得到的码率 R_1 进行比较,如果 R_2 比 R_1 更接近设置码率,那么说明新算法将码率控制在了更精确的范围内。由于图像质量的评价指标是亮度峰值信噪比,本文再将新算法测试所得的亮度峰值信噪比 (Peak Signal to Noise Ratio, PSNR) P_2 和用 JM 算法测试得到亮度峰值信噪比 P_1 进行比较,如果 P_2 比 P_1 更高,那么说明新算法提高了传输图像的质量。用于测试的 QCIF 格式的图像序列有 foreman, highway, grandma, bridge-far; CIF 格式的图像序列有 foreman, mobile, tempete, highway。在这些序列中,局部运动较为剧烈有:forman, highway 和 mobile, 而 bridge_far 运动较为缓慢, grandma 和 tempete 则属于中等运动强度的图像序列。在对 QCIF 序列进行码率控制算法测试时,我们将码率分别设定为 64 Kbps、128 Kbps、192 Kbps;而对于 CIF 序列,我们将码率分别设定为 256 Kbps、512 Kbps、1024 Kbps。在测试条件方面,我们设置帧率为 30 fps, 初始 QP 值为 28, 参考帧数为 1, 编码帧数为 100, 采用 Hadamard 编码, 熵编码形式采用 CAVLC, GOP 结构为 IPPP。

表1 QCIF 序列测试结果

序列名称	设置码率/Kbps	实际码率/Kbps			亮度 PSNR/dB		
		R_1	R_2	$R_2 - R_1$	P_1	P_2	$P_2 - P_1$
foreman	64	65.92	65.74	-0.18	33.87	33.93	0.06
	128	131.36	131.14	-0.22	36.69	36.78	0.09
	192	196.75	196.61	-0.14	38.30	38.40	0.10
highway	64	66.25	65.20	-0.05	37.28	37.28	0.00
	128	133.52	132.49	-1.03	40.01	40.05	0.04
	192	196.81	196.77	-0.04	40.80	41.10	0.30
grandma	64	65.63	65.48	-0.15	39.87	39.94	0.07
	128	131.16	131.01	-0.15	42.55	42.62	0.07
	192	196.74	196.61	-0.13	44.28	44.32	0.04
bridge-far	64	65.55	65.42	-0.13	40.20	40.31	0.11
	128	131.17	131.08	-0.09	40.69	40.77	0.08
	192	196.85	196.72	-0.13	41.20	41.30	0.10

从表1~3中可以看出在不同的图像格式和不同码率条件下,用新算法测试得到的码率 R_2 比用 JM 原算法测试得到码率 R_1 更接近设置码率。这说明改进后的 MAD 预测算法

提高了码率的控制精度,在实时通信中可以更好地进行码率控制。用新算法测试得到的亮度峰值信噪比 P_2 比用 JM 原算法测试得到的亮度峰值信噪比 P_1 更高,说明改进后的 MAD 预测算法在更精确控制码率的同时,还提高了传输图像质量。

表2 CIF 序列测试结果

序列名称	设置码率/Kbps	实际码率/Kbps			亮度 PSNR/dB		
		R_1	R_2	$R_2 - R_1$	P_1	P_2	$P_2 - P_1$
foreman	256	259.50	259.45	-0.05	33.19	33.27	0.08
	512	518.86	518.76	-0.10	36.91	37.00	0.09
	1024	1036.37	1036.35	-0.02	40.68	40.71	0.03
mobile	256	260.44	259.82	-0.62	21.86	21.94	0.08
	512	520.04	518.72	-1.32	26.13	26.24	0.11
	1024	1040.79	1040.29	-0.50	31.17	31.35	0.18
tempete	256	258.35	257.92	-0.43	27.87	27.97	0.10
	512	517.85	517.54	-0.31	31.21	31.32	0.11
	1024	1039.58	1038.27	-1.31	35.21	34.35	0.14
highway	256	262.98	262.26	-0.72	38.86	38.95	0.09
	512	525.11	525.08	-0.03	40.87	40.97	0.10
	1024	1048.81	1046.88	-1.93	43.19	43.30	0.11

表3 QCIF 序列和 CIF 序列改进前后的平均码率差和 PSNR 差

设置码率/Kbps	$R_2 - R_1$ 的均值	$P_2 - P_1$ 的均值
64	-0.1275	0.0600
128	-0.3725	0.0700
196	-0.1100	0.1350
256	-0.4550	0.0875
512	-0.4400	0.1025
1024	-0.9400	0.1150

3.2 对场景转换图像序列的测试

本文最后进行了场景转换图像序列的新算法测试。测试前,我们首先将 CIF 格式的序列 foreman 与 mobile 拼接成一个序列 foreman&mobile, bridge-far 与 news 拼接为一个序列 bridge-far&news,把场景转换均设置在第6帧。其中 foreman 与 mobile 均为运动较剧烈的图像,而 bridge-far 与 news 均为运动缓慢的图像。接下来将我们用新算法和 JM 算法将两个拼接序列分别进行测试。本节的测试条件同 3.1 节测试条件,只不过将编码帧数改为 10, 目标码率定为 400 Kbps。foreman&mobile 序列和 bridge-far&news 序列在新算法和 JM 算法中编码每一帧所用字节数如图3与4所示。

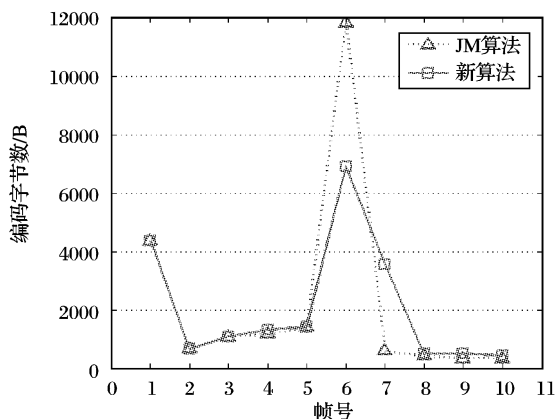


图3 foreman&mobile序列的编码比特数

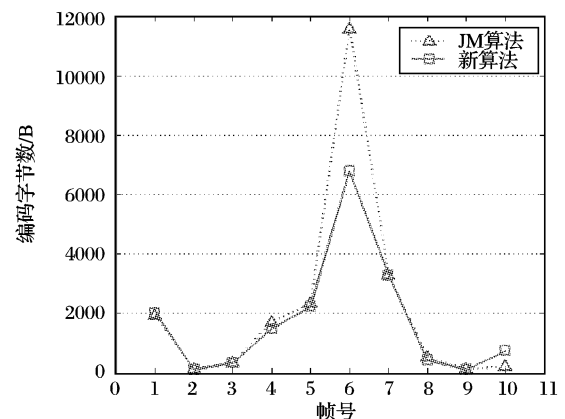


图4 bridge-far&news序列的编码比特数

线模型的可扩展性差及资源利用率低的缺点,同时解决了并行模型对于某些处理开销大且设计复杂的问题,可以针对不同特性的处理选择不同的处理方式。

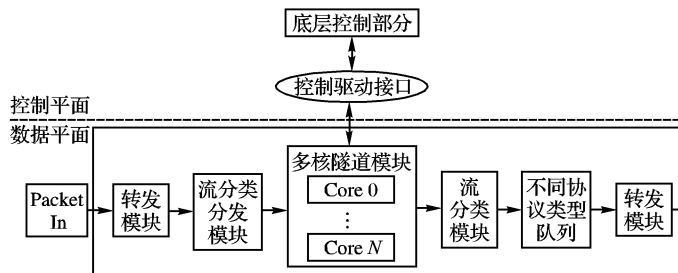


图3 多核隧道系统仿真模型

3.2 测试数据与分析评价

依据上述系统模型与实现原理,利用图4组网进行测试,在PC机A与PC机B之间用SMB打双向包流量,以0丢包时的吞吐量作为转发流量,SMB测试时间为60s,数据精度要求0.1%,其中SMB带宽1000Mbps,多核CPU有8核32线程(VCPU),型号为RMIXLR732。获得数据见表1。

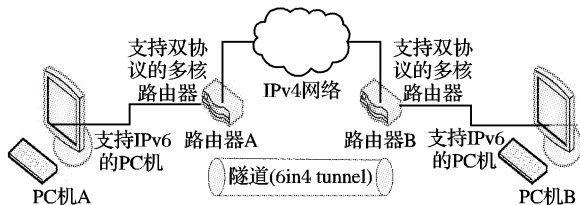


图4 隧道测试组网图

表1 6in4隧道单核与多核下吞吐量(双向)

测试包长/B	吞吐量/pps	
	多核方式	单核方式
128	1332875.840	295113.774
256	842108.243	292792.652
512	450099.859	292739.887
1024	234412.428	233781.754
1280	188251.212	188189.372
1400	173580.546	172962.593

依表1中6in4隧道单核与多核方式下测试数据,得两者的数据折线比较图,如图5所示。

从表1与图5可见,相对于单核方式,多核下隧道对小报文与大报文在转发处理性能上成倍提高,这表明多核方式下

IPv6隧道技术能更高效地解决过渡时期问题。同时采用多核来实现隧道技术更易于进行业务功能的扩展,适应未来网络对高效与业务复杂的要求。

4 结语

多核处理器具有高性能、设计和验证周期短、易扩展、通信延迟低等优点,在推动PC安全性和虚拟技术方面起到关键作用^[6]。IPv6彻底解决了地址空间耗尽和路由表爆炸等问题,协议的设计使路由器处理报文更加简便^[1]。

本文采用多核与隧道技术相结合的方案,能满足目前高速网络处理的同时适应未来网络的发展方向,更好地解决IPv4/IPv6过渡问题。

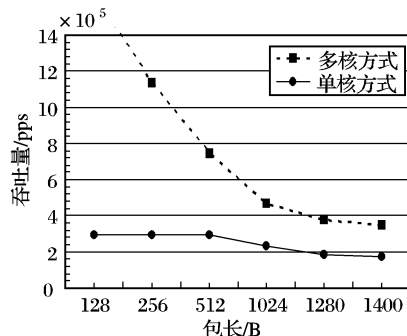


图5 6in4隧道单核与多核方式下测试数据折线图

参考文献:

- [1] DESMEULES R. Cisco ipv6 网络实现技术[M]. 王玲芳, 张宇, 李颖, 等译. 北京: 人民邮电出版社, 2004.
- [2] 王忠培, 周健, 李勇. 基于隧道和 NAT-PT 相结合的 IPv6 过渡方案[J]. 微型计算机, 2008, 24(1/3): 109-111.
- [3] 沈庆伟, 张霖. 基于隧道的 IPv4/IPv6 过渡技术分析[J]. 计算机技术与发展, 2007, 17(5): 171-176.
- [4] 杜慧军. 基于双协议栈的 6to4 隧道技术的应用[J]. 广东技术师范学院学报, 2007(12): 16-20.
- [5] 杭州华三通信技术有限公司. 多核分布式 Net Stream 技术白皮书[Z]. 杭州华三通信技术有限公司, 2007: 6-8.
- [6] 吴继雁. 多核技术与发展趋势[J]. 哈尔滨轴承, 2007, 28(2): 57-58.
- [7] 戴彬, 钱德沛, 刘铁, 等. IPv4/IPv6 协议过渡机制的实验研究[J]. 微电子学与计算机, 2004, 21(9): 64-67.

(上接第 677 页)

从图3和4中可以看出,在出现场景转换的位置, JM 算法中剧烈增加的编码字节数在新算法中得到了很好的控制。说明当图像序列出现场景转换时,和 JM 算法相比,新算法可以更加有效地进行码率控制。

4 结语

本文基于 H.264 视频编码标准的码率控制方案,提出了一种改进的码率控制算法。实验结果表明,改进后的新算法不仅能在各种信道传输中对码率的控制更加准确,而且能在相同的码率下得到比原有算法更高的 PSNR 值,较为明显地提高了解码图像的质量。

参考文献:

- [1] ISO/IEC JTC1/SC29/WG11 MPEG93/457, MPEG-2 Video Test Model, Version 5.0[S]. Genevese: ISO, 1993.

- [2] ISO/IEC JTC1/SC29/WG11 W1796, MPEG-4 Video Verification Model, Version 8.0[S]. Genevese: ISO, 1999.
- [3] ITU-T TMN8, Video Codec Test Model, Near-Term, Version 8.0[S]. Genevese: ISO, 1997.
- [4] WIEGAND T, SULLIVAN G, BJONTEGAARD G. Overview of the H.264/AVC video coding standard[J]. IEEE Transaction on Circuits and Systems for Video Technology, 2003, 13(7): 560-576.
- [5] 葛晓妮, 王宇, 郝重阳, 等. H.264/AVC 率失真优化(RDO)策略研究[J]. 无线通信技术, 2006, 15(2): 14-22.
- [6] 蒋建国, 袁炜, 王德宝. H.264 码率控制算法分析[J]. 电视技术, 2007(5): 10-22.
- [7] 王友炳, 蔡灿辉. 一种 H.264/AVC 码率控制的改进算法[J]. 计算机工程与应用, 2008, 44(11): 199-202.
- [8] KIM M-J, KIM K-H, HONG M-C. Adaptive rate control in frame-layer for real-time H.264/AVC[J]. Advanced Communication Technology, 2008, 3(2): 1875-1880.