

一种自适应的 AR-FGS 漏因子选择算法

邓 云^{1,2}, 彭 强¹, 诸昌铃¹

(1. 西南交通大学 信息科学与技术学院, 四川 成都 610031;

2. 四川理工学院, 四川 自贡 643000)

摘要: AR-FGS(Adaptive Reference for Fine Granular Scalable)使用漏预测在编码效率和码流鲁棒性间取得平衡. 针对漏因子对漏预测效果影响至关重要的问题, 提出了一种自适应的漏因子选择算法. 分析发现 JSVM(Joint Scalable Video Model)中对每帧按相同比例进行码流截取的方式存在着输出码率在帧级上不平滑的问题, 为此提出按固定帧总数据量的方式进行截断的码流截取算法, 保证输出码率在帧级上的平滑. 根据当前参考帧基本层数据量与第一帧基本层数据量的比值为每帧单独决定最优的漏因子, 并根据短期平均基本层数据量与当前参考帧基本层数据量的比值, 对选择的漏因子进行调节. 实验结果表明, 算法产生的码流在大部分码率范围内 PSNR 值高于使用固定漏因子的最高 PSNR 值. 此外, 修改的码流截取方式产生的码流在帧级数据量上平滑.

关键词: 视频编码; AR-FGS 漏预测; 漏因子; 码流截取; JSVM

中图分类号: TP391 **文献标识码:** A **文章编号:** 1001-2400(2008)06-1115-06

Adaptive leaky factor selection algorithm for AR-FGS

DENG Yun^{1,2}, PENG Qiang¹, ZHU Chang-qian¹

(1. School of Information Sci. and Tech., Southwest Jiaotong Univ., Chengdu 610031, China;

2. Sichuan Inst. of Tech., Zigong 643000, China)

Abstract: AR-FGS (Adaptive Reference for Fine Granular Scalable) adopts leaky prediction to achieve a better tradeoff between coding efficiency and robustness. The leaky factor is crucial to the performance of leaky prediction. This paper proposes an adaptive method for determining the optimal leaky factor for each frame. First, the disadvantage of the bit-stream extraction method used in the current JSVM (Joint Scalable Video Model) is analyzed that the bit-rate of extracted sub-stream is not smooth at the frame level. To guarantee the smoothness, a modified extraction method is presented which truncates each frame at the fixed total bit-rate. The proposed leaky factor determination algorithm sets the optimal leaky factor for each frame according to the ratio of current reference frame's base layer bit-rate to that of the first I-frame's. The optimal leaky factor is further adjusted according to the ratio of several previous frames' average base layer bit-rate to that of current reference frame's. Simulation results show that the proposed algorithm can further improve the PSNR over a wide range of bit-rate, compared with the use of the fixed leaky factor. Additionally, the bit-rate of sub-stream extracted by the modified bit-stream extraction method is smooth at the frame level.

Key Words: video coding; AR-FGS; leaky prediction; leaky factor; bit-stream extraction; JSVM

视频数据传输由于使用了混合视频编码技术, 容易受到带宽波动、随机错误、突发错误以及包丢失等多种问题的影响而造成差错扩散, 并最终导致图像质量严重下降. 精细的可伸缩性视频编码 FGS(Fine Granular Scalable), 因其增强层码流可在任意位置被截断而不影响保留码流的正确解码, 是对抗网络带宽波动及提高视频数据鲁棒性的有效手段. 然而, MPEG-4 FGS 由于在运动补偿中对基本层和增强层都只使

收稿日期: 2007-12-07

基金项目: 国家自然科学基金资助(60672099); 西南交通大学交通运输工程研究生创新实践基地资助

作者简介: 邓 云(1976-), 男, 四川理工学院讲师, 西南交通大学博士研究生, E-mail: dragon_dy@163.com.

用了低质量的基本层重建图像作为参考,因而编码效率较低^[1].为改善编码效率并保证码流的鲁棒性,文献[2~5]提出了多种改进算法.MC-FGS(Motion-Compensated FGS)通过对基本层和增强层同时使用高质量的重建图像作为预测参考来去除时域冗余,提高了编码效率;然而当增强层被截断时,错误将同时扩散到基本层和增强层,牺牲了较大的鲁棒性^[2].渐进精细的可伸缩性视频编码 PFGS(Progressive FGS)^[3]对增强层使用一个单独的高质量预测环路,并保留一条从基本层到最高质量层的预测路径,使增强层可以从错误中逐渐恢复,在编码效率和鲁棒性间取得了较好的折中.鲁棒、精细的可伸缩性视频编码 RFGS(Robust FGS)^[4]则使用漏预测技术(Leaky Prediction),将增强层部分位平面用漏因子缩放后与基本层重建图像加权平均,作为增强层参考,使增强层错误可经漏预测后逐帧减弱.EFGS(Efficient FGS)^[5]针对 RFGS 中至关重要的漏因子取值,根据增强层数据量为每个位平面独立设置漏因子,进一步提高了编码效率.

最新制定的精细可伸缩视频编码标准 JVT FGS^[6],采用 AR-FGS 在编码效率和鲁棒性间取得平衡.AR-FGS 使用了漏预测技术^[7],其高质量参考图像由前一帧对应 FGS 层的重建图像和基本层重建图像经漏因子加权平均后产生.其中漏因子的取值对编码效率和鲁棒性的影响至关重要.最优漏因子的取值紧密依赖于序列内容,然而,由于 JVT FGS 使用 Cyclical Block Coding 熵编码^[8]代替了 MPEG-4 FGS 中的位平面编码,MPEG-4 FGS 中的位平面被替换为连续的 FGS 层,因此针对 MPEG-4 FGS 的漏预测技术(如 EFGS),并不能直接应用于 JVT FGS.

针对上述不足,笔者提出了一种自适应的漏因子选择算法.考虑到码流截取方式对 FGS 的性能影响很大,首先分析了 JVT FGS 测试模型(JSVM)^[9]中 FGS 层被截取的过程,并针对该截取过程在码率控制方面的不足,提出修改的截取方式,有效满足了码流截取算法作为码率控制机制的需要.然后分析了 FGS 层被截取后,截取比例、序列运动程度及纹理复杂度对预测效率和预测漂移的影响,提出根据基本层数据量比值逐帧选择最优漏因子,并根据短期平均基本层数据量对最优漏因子进行调节的算法.算法不仅能对不同序列逐帧设置合适的漏因子,且在大部分码率范围内优于使用固定漏因子的最优效果.由于仅使用基本层数据量作为判断条件,算法计算复杂度极低.

1 JSVM 中的码流截取过程与改进

1.1 JSVM 的码流截取方式及其不足

FGS 码流分为两部分:基本层码流与增强层码流.基本层码流通常被保证正确接收,增强层码流则可在任意位置被截断而不影响解码器对实际接收到的部分码流正确解码,且重建的图像质量随实际解码的码率增加而逐渐改善.这既能对抗网络带宽波动,也可作为一种码率控制手段.目前 JSVM 中使用一种非常简单的方式对增强层码流进行截取:首先通过码流分析计算出基本层码率 R_b 与增强层码率 R_e (增强层码率不包括基本层码率),然后计算出目标带宽 R 超出基本层码率的部分 $R_r = R - R_b$.如果 $R_r \geq R_e$,则增强层可完整传输;否则,计算出增强层可被传输的比例 $\lambda = R_r / R_e$,并对每一帧增强层码流,在比例为 λ 的位置进行截取,使保留码流的码率满足目标带宽的要求.

上述码流截取方式虽然在总体码率上满足了目标码率的要求,然而,作为码率控制机制,其在帧级数据量控制上却不够平稳.原因在于:为简化编码过程,目前 JSVM 对基本层没有使用码率控制机制,仅使用足够小的固定量化参数来保证其码率满足最低目标带宽的要求;然后对增强层码流通过码流截取进行码率控制.因此,基本层数据量在帧级上波动明显.由于增强层实际编码的是原始图像与基本层重建图像的差值,即基本层量化误差.而量化误差的大小主要取决于量化步长,因此当基本层使用固定量化步长时,基本层量化误差变化较小,即增强层实际编码的信息量变化较小,进而编码产生的增强层码流相对基本层码流变化较小,其按相同比例截取后保留的增强层码流波动也较小.因为截取后的总体码流由完整的基本层码流和保留的增强层码流构成,前者在帧级波动明显,后者相对平稳,故总体码流在帧级波动明显.针对此问题,提出下面改进的码率截取方式.

1.2 改进的码流截取方式

为使截取后的保留码流在帧级变化平稳,修改了 JSVM 中的截取过程.对指定目标码率,计算出平均每

帧的目标数据量,并按此目标数据量对增强层码流逐帧截取.修改后的截取过程如下:

- 1) 根据目标带宽 R 与帧率 F 计算出每帧的目标数据量 $R_t = R/F$;
- 2) 若当前帧基本层数据量 $R_{bi} > R_t$,则只传输完整的基本层部分,并对超出目标数据量的部分平均到其余帧中扣除;
- 3) 若当前帧总体数据量 $R_{bi} + R_{bi} < R_t$,则传输完整码流,并将剩余的目标数据量平均加至其余帧;
- 4) 若 $R_{bi} + R_{bi} > R_t > R_{bi}$,则对增强层数据在 $R_t - R_{bi}$ 点进行截取,使得保留的帧数据量恰好为 R_t .

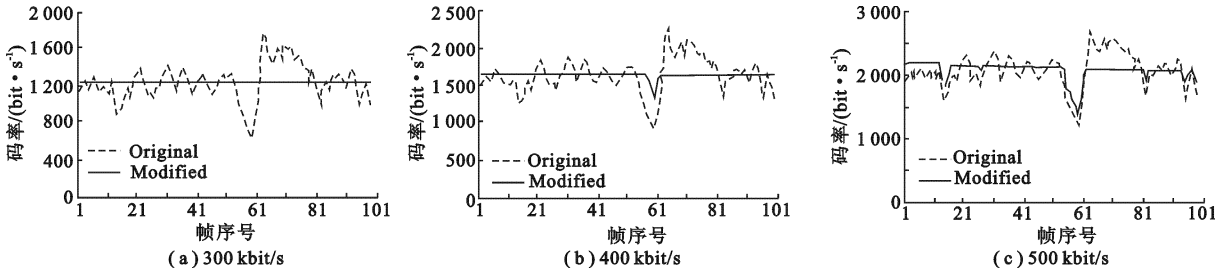


图 1 原截取算法与本文截取算法在帧数据量控制上的实验结果对比(基本层码率为 135 kbit/s,总码率为 569 kbit/s)

图 1 为 JSVM 中的截取过程与修改后的截取过程对“Foreman”序列分别在 300 kbit/s, 400 kbit/s, 500 kbit/s 截断后,统计的每帧数据量波动情况(不计第一个 I 帧).从图中可以看出,修改后的截取过程对序列帧数据量有非常平稳的控制效果.图 1(b), (c) 中本算法出现的数据量向下波动的情况正是 $R_{bi} + R_{bi} < R_t$ 时的特殊情况.

2 自适应的 AR-FGS 漏因子选择算法

2.1 AR-FGS 中的漏预测

JVT FGS 允许在一个基本层之上编码至多三个 FGS 增强层.为提高编码效率,允许 FGS 层在预测时使用前面帧的基本层或 FGS 增强层重建图像作为参考,并同时使用 AR-FGS 控制预测漂移. AR-FGS 基于漏预测技术,根据基本层已编码信息,以 4×4 块为单元,自适应地将参考帧增强层重建图像经漏因子缩放后与基本层重建图像加权产生一个预测参考图像,并以此预测参考图像进行运动补偿,最后对预测残差进行熵编码.这样,预测漂移经过多次漏预测缩放后被逐渐消除.自适应参考图像的具体产生过程如下:

- 若当前块在基本层中被量化后的系数全部为 0,则当前块的参考块为当前帧基本层重建块与参考帧增强层中运动补偿对应的重建块的权值和,即

$$R_e^n(i, j) = (1 - \alpha)X_b^n(i, j) + \alpha R_e^{n-1}(i, j) \quad , \quad 0 \leq i < N, \quad 0 \leq j < N \quad , \quad (1)$$

其中, R_e^n 为待产生的预测参考块, X_b^n 为当前帧基本层对应块的重建图像, R_e^{n-1} 为参考帧增强层中运动补偿对应的重建图像, (i, j) 为像素点坐标, N 为坐标范围. α 为针对 4×4 块的漏因子,表示增强层图像在实际产生的预测参考图像中所占的权重.

- 否则,先对 X_b^n 和 R_e^{n-1} 进行前向 4×4 整数变换,分别得到变换系数块 $F_{X_b^n}$ 和 $F_{R_e^{n-1}}$.再按式(2)计算出 R_e^n 的变换系数块 $F_{R_e^n}$,

$$F_{R_e^n}(u, v) = \begin{cases} F_{X_b^n}(u, v) & , \quad Q_b^n(u, v) \neq 0 \quad , \\ (1 - \beta)F_{X_b^n}(u, v) + \beta F_{R_e^{n-1}}(u, v) & , \quad Q_b^n(u, v) = 0 \quad . \end{cases} \quad (2)$$

最后将 $F_{R_e^n}$ 进行反向 4×4 整数变换得到 R_e^n .其中, $Q_b^n(u, v)$ 表示当前块在基本层位置 (u, v) 上的量化系数, β 为针对单个系数的漏因子,同样表示增强层图像在实际产生的预测参考图像中所占的权重.

α 和 β 的设置以片为单位.漏因子越小,预测更多来自基本层,码流的抗差错能力较强,但编码效率较低;漏因子越大,预测更多来自增强层,编码效率提高,但码流的抗差错能力下降.

2.2 自适应的 AR-FGS 漏因子选择算法

AR-FGS 中的漏因子对编码效率和抗差错能力的影响至关重要.然而,现有 AR-FGS 中虽指出最优漏

因子依赖于序列内容,却并未提出确定最优漏因子的有效方法.基于 MPEG-4 FGS 的漏预测也因为与 JVT FGS 的差异不能直接应用于 AR-FGS. JVT FGS 测试模型——JSVM 中使用固定漏因子对序列所有图像进行编码,尽管实现简单,却存在明显不足:(1)虽然可通过将所有可用漏因子逐一编码并模拟传输和解码后找出该序列的最优因子,但因计算复杂度巨大而无法应用于实时应用.(2)漏因子以序列为单位选择,没有针对序列每一帧进行单独优化.(3)码流的截取方式对实际解码的增强层图像影响至关重要,进而影响到图像在下一帧漏预测中可占的权重,这一点也没有考虑到.

针对此问题,笔者提出一种简单灵活的 AR-FGS 漏因子选择算法.算法思想是:在漏预测中,从参考帧增强层进行预测的准确程度决定了漏因子的选择,而预测的准确程度主要由以下两个因素决定.

1) 参考帧增强层截取后保留的数据量.增强层保留的数据量越多,其重建图像失真越小,以其作为预测参考时,预测图像中包含的预测漂移越小,应使用较大漏因子;反之,保留的数据量越少,参考图像中包含的预测漂移越大,应使用较小漏因子.在本文的码流截取方式中,因为截取后每帧保留的总数据量固定不变,故,基本层数据量越小,增强层截取后保留的增强层越大,漏因子应越大;反之,基本层数据量越大,漏因子应越小.

2) 参考帧图像的运动程度和纹理复杂度.运动程度和纹理复杂度虽然不产生预测漂移,但影响到时域预测的效率.图像运动程度越大,纹理越复杂,用参考帧增强层作为预测参考时,预测效率越低,相对使用当前帧基本层重建图像作为预测图像的优势下降,故应选择较小漏因子;反之,漏因子应取较大值.然而,直接计算图像运动程度和纹理复杂度将增加计算复杂度,另外,对个别根据率失真原则使用帧内模式的宏块,需设定一个等价运动程度,也带来实现上的难度.因为当量化参数固定时,图像运动程度和纹理复杂度决定着编码产生的数据量,故为避免额外的计算开销,本文同样通过基本层数据量来估计图像运动程度和纹理复杂度.编码产生的基本层数据量越大,图像运动程度和纹理复杂度越大,应使用较小漏因子;反之,使用较大漏因子.对使用帧内模式的宏块,因为编码产生的数据量普遍大于帧间模式产生的数据量,故运动程度已经反映在数据量中,无需另行计算.

以上两个因素都通过基本层数据量进行估计,且数据量都与漏因子成反正关系,可结合在一起处理,因而计算复杂度极低.然而,在实际实现过程中,还有如下两个重要问题需要注意.

1) 量化参数同样影响了序列编码产生的数据量.当使用不同量化参数对同一序列编码时,产生的数据量变化较大,很难判断数据量变化是受量化参数影响还是序列本身内容变化所致.因此,如果直接使用数据量作为漏因子选择依据,将影响到选择的准确性.然而,由于相同量化参数影响到序列每帧数据量的程度相同,因此,本文使用每帧基本层数据量与第一个 I 帧的基本层数据量的比值作为漏因子的实际选择依据,可以消除不同量化参数带来的影响.

2) 当前帧的参考帧使用的漏因子也影响到当前帧从参考帧增强层预测的准确程度.参考帧使用的漏因子越大,参考帧增强层在它的漏预测中使用的参考图像更多地来自它的参考帧增强层,因此它的重建图像中包含的预测漂移越大,导致当前帧从参考帧增强层预测的准确度下降,故当前帧的漏因子应适当减小.反之,参考帧使用的漏因子越小,参考帧重建图像中包含的预测漂移越小,当前帧的漏因子应适当增加.在实现上,因为漏因子由基本层数据量决定,故对根据当前基本层数据量与第一个 I 帧数据量的比值计算出的漏因子,再根据当前帧基本层数据量与前面几帧的平均基本层数据量的比值进行调整,得到最后优化的漏因子.

AR-FGS 中 α 和 β 对应的语法元素可有 $0, 1, \dots, 32$ 共 33 个取值,对应的漏因子分别为 $0, 1/32, \dots, 1$.通过大量实验发现:漏因子取值低于 16 时,绝大多数序列的 PSNR 几乎在整个码率范围内低于取值大于 16 时的情况.当漏因子取值介于 24 与 32 之间时,序列变化程度对 PSNR 影响极大,运动平缓的序列 PSNR 提升较大.当漏因子大于 28 时,只有运动程度极平缓的序列才在较高码率段有所获益.当漏因子为 32 时,几乎所有序列的 PSNR 都在较大码率范围内明显下降,仅在极高码率处有所提高.为保证极低的计算复杂度,笔者使用简单的分段线性关系来选择适当的漏因子.

算法具体步骤如下:

步骤 1 编码第一个 I 帧,记录其基本层数据量 R_{1b} 作为衡量后续 P 帧数据量大小的标准.对第一个 I 帧,因其数据量很大,从码率控制角度出发,传输时只传输基本层,不传输增强层.

步骤 2 编码第一个 P 帧. 因为第一帧只传输基本层, 故此帧只能从 I 帧基本层预测, 漏因子设为 0.

步骤 3 对后续 P 帧, 当编码增强层时, 先计算出参考帧基本层数据量 R_{rb} 与 R_{lb} 的比值 $p = R_{rb}/R_{lb}$. 显然, p 越小, 基本层数据量越小, 漏因子应越大; p 越大, 漏因子越小. 经过大量实验统计, 使用以下简单的线性关系来计算预设的漏因子:

$$\alpha' = \begin{cases} 27 + 3 \times (0.15 - p) / 0.14 & , \quad p < 0.15 \\ 22 + (0.3 - p) / 0.03 & , \quad p \geq 0.15 \end{cases} \quad (3)$$

步骤 4 根据前 5 帧平均基本层数据量 \bar{R}_{b5} 对 α' 进行调节. 先计算调节系数 $c = (\bar{R}_{b5}/R_{rb})^{1/4}$, 再将 c 限制在范围 1.1 ~ 0.9 内, 最后得到优化的漏因子 $\alpha = c \times \alpha'$

步骤 5 对 β , 因其仅调整非全零块中的零系数权重, 通过大量实验发现, 固定 α 时, β 的不同取值对编码效率产生的影响甚微. 因此, 简单起见, 令 $\beta = \alpha$.

3 实验结果及分析

模拟实验以 JVT SVC 标准测试模型 JSVM8 为平台, 实现了笔者提出的码流截取算法以及自适应漏因子选择算法. 因为 AR-FGS 针对低延迟应用, 不使用 B 帧, 故 GOP 大小设为 1. 基本层量化参数固定为 28, 编码一个 FGS 增强层, 增强层在时域预测时直接使用基本层运动信息(参考帧号和运动矢量). 第一帧为 I 帧, 其余为 P 帧. 运动矢量搜索范围限制为 ± 16 个整像素点, 帧率为 30 f/s, 每个序列编码 100 帧(CIF 格式的 Football 只有 90 帧). 对原 JSVM 中使用固定漏因子方式, 令 $\alpha = \beta$, 对 0, 8, 16, 24, 32 共 5 个值分别进行试验, 并与本算法进行比较.

先对 Foreman, Miss_am 和 Container 3 个 QCIF 序列进行实验. 对编码产生的码流, 从基本层码率到完整码率, 以 25 kbit/s 为增量在各个码率点依次截取, 对每次截取后的保留码流解码并计算序列重建图像的平均 PSNR. 图 2 为实验结果. 再对 Football, Waterfall, 和 News 共 3 个 CIF 序列进行实验. 实验条件与 QCIF 序列相同, 只是码率等级增量改为 100 kbit/s. 图 3 为实验结果. 从图中可以看出, 不同漏因子对编码结果影响很大, 而本文算法因为综合考虑了影响漏预测的多个因素, 对漏因子进行了逐帧优化, 所以在整个码率范围内都有很好的效果, 且在大部分码率范围内, PSNR 值高于使用固定漏因子时的最高 PSNR 值, 最多高出 0.9 dB, 仅在极高码率点处略低于漏因子为 32 的效果. 算法对具有中等运动程度的序列效果尤其明显, 如 QCIF 的 Miss_am, Container, CIF 的 Football, 因为此类图像时域预测的效率适中, 特别需要在效率和鲁棒性间达到平衡, 所以对漏因子进行逐帧优化后效果特别明显. 对运动缓慢序列在中低码率点效果仍然明显, 但在高码率点处有所下降, 如 CIF 的 News. 原因在于码率较高时增强层被正确解码的比例较大, 从增强层预测的获益较大, 有利于较大的漏因子参数, 如 32. 故码率较高时, 有利于较大的固定漏因子取值, 在最高码率点处, 使用 32 的固定漏因子始终有最高 PSNR 值, 但在实际应用中, 最高码率常常是不能满足的. 本算法保证了实际应用中可以使用的大部分码率范围内的更好的编码性能.

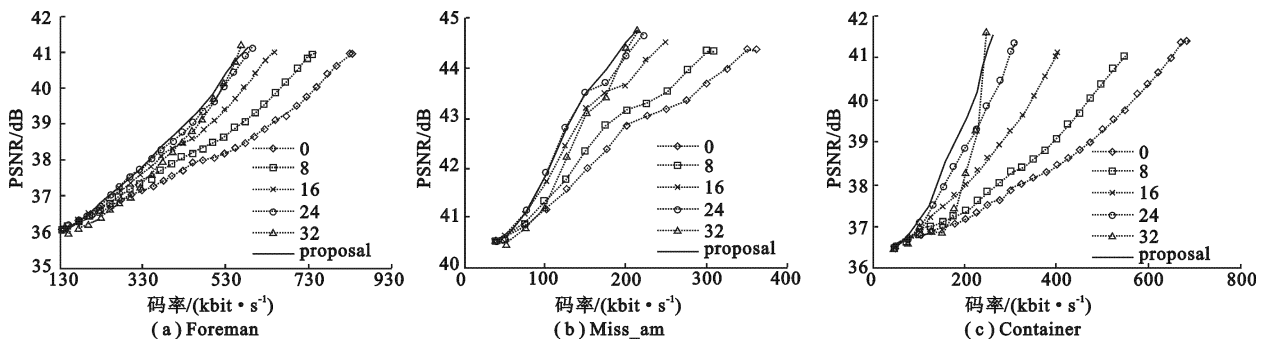


图 2 QCIF 序列在固定因子算法与本算法的实验结果对比

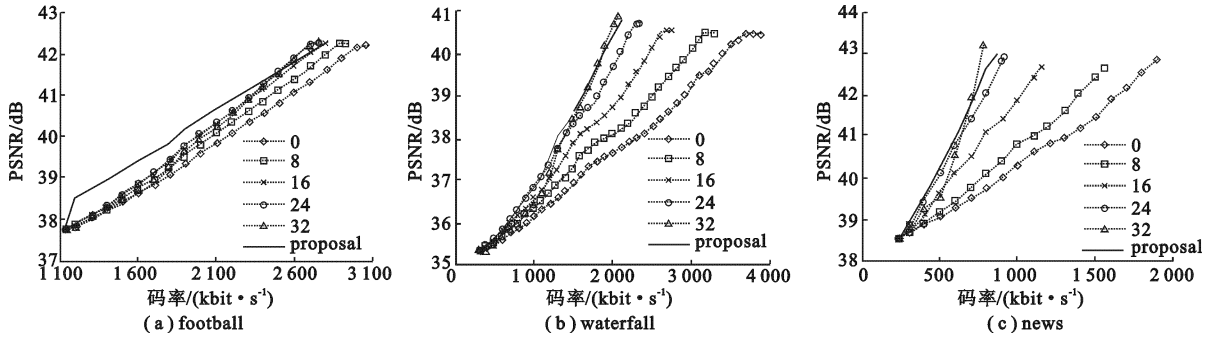


图 3 CIF 序列在固定因子算法与本算法的实验结果对比

4 结 论

AR-FGS 使用漏预测有效提高了 FGS 的编码效率. 然而, 固定的漏因子没有将序列本身信息和码流截取特点考虑在内, 存在参数选择困难和效率不高两个缺点, 且无法针对实时视频通讯应用进行优化. 笔者提出的自适应漏因子选择算法, 结合修改的码流截取过程, 充分考虑了序列的运动程度和纹理复杂度, 根据编码产生的基本层数据量, 自适应地对每一帧选择最优的漏因子, 并根据短期基本层平均数据量进行调整. 算法计算复杂度极低, 且在大部分码率范围内优于最优固定漏因子. 此外, 笔者提出的修改后的码流截断方式在帧级数据量控制上效果非常平稳.

参考文献:

- [1] Li W. Overview of Fine Granularity Scalability in MPEG-4 Video Standard [J]. IEEE Trans on Circuits Syst Video Technol, 2002, 11(3): 301-317
- [2] Van Der Schaar M, Radha H. Adaptive Motion-Compensation Fine-Granular-Scalability (AMC-FGS) for Wireless Video [J]. IEEE Trans on Circuits Syst Video Technol, 2002, 12(6): 360-371.
- [3] Wu F, Li S, Zhang Y Q. A Framework for Efficient Progressive Fine Granularity Scalable Video Coding [J]. IEEE Trans on Circuits Syst Video Technol, 2001, 11(3): 332-344.
- [4] Huang H C, Wang C N, Chiang T. A Robust Fine Granularity Scalability Using Trellis-based Predictive Leak [J]. IEEE Trans on Circuit Syst Video Technol, 2002, 12(6): 372-385.
- [5] Gao Y L, Chau L P. Efficient Fine Granularity Scalability Using Adaptive Leaky Factor [J]. IEEE Trans on Broadcasting, 2005, 51(4): 512-519.
- [6] ITU-T, ISO/IEC JTC1. JVT-V201, Scalable Video Coding-Joint Draft 8 [S]. Marrakech, Morocco: JVT, 2007.
- [7] ITU-T, ISO/IEC JTC1. JVT-O039, CE7 Report: FGS Coding for Low-delay Applications [S]. Nice, France: JVT, 2005.
- [8] ITU-T, ISO/IEC JTC1. JVT-V203, Joint Scalable Video Model JSVM 8 [S]. Marrakech, Morocco: JVT, 2007.

(编辑: 高西全)