

提高 H.26L 压缩视频抗误码性能的 全局率失真优化方法¹

杨 华 余松煜 杨松岸

(上海交通大学图像通信与信息处理研究所 上海 200030)

摘 要: 该文从较新的 H.26L 视频压缩编码标准出发,并结合 Internet 信道的 Bernoulli 和 Gilbert 的包丢失模型,提出了一种全局率失真优化的编码模式判决方法,使编码的视频流具有更强的错误恢复能力.实验证明,在实际的 Internet 信道误码环境下,此方法具有更强的抗错误和实用性,得到了较好的效果.

关键词: Bernoulli 信道模型, Gilbert 信道模型,全局率失真优化, Internet 包交换, H.26L 视频压缩

中图分类号: TN919.8 **文献标识码:** A **文章编号:** 1009-5896(2004)06-0846-07

A Global Rate-Distortion Optimized Approach to H.26L Low Bit Rate Robust Video over the Internet

Yang Hua Yu Song-yu Yang Song-an

(Inst. of Image Comm. & Info. Proc., Shanghai Jiaotong Univ., Shanghai 200030, China)

Abstract A new Global Rate Distortion (GRD) optimal macroblock coding mode decision scheme is introduced for the newly H.26L video compression standard. Based on the Internet packet loss model of Bernoulli and Gilbert, this R-D mode decision approach can result in higher error robust than classical method. Furthermore, our experimental results also demonstrate its superior adaptive error resilient ability and practicability.

Key words Bernoulli channel model, Gilbert channel model, Global rate distortion optimality, Internet packet switch, H.26L video compression

1 引言

近些年来,随着网络技术和无线通讯技术的发展,通过 Internet 和无线信道中传输实时图像已经越来越成为目前图像通讯领域中的研究热点.各种不同的视频流应用环境对图像的压缩和传输也提出了更高的要求. H.26L 作为新一代视频压缩算法^[1],其最终目的不仅仅要提高在无差错电路交换网络上传输的率失真性能,而且还要提高在各种不同网络拓扑结构上具有误码或丢包情况下的错误恢复能力^[2].

虽然 H.26L 开发了针对不同网络传输条件的网络适配层 NAL (Net Adaption Layer),但是与其先进的压缩编码技术相比, H.26L 压缩视频的抗误码技术还比较落后.目前采用的技术主要停留在宏块的帧内刷新和编码端的多解码宏块失真平均的率失真优化方法^[3].其中宏块的帧内刷新由于没有考虑信道的误码分布及图像序列的误码敏感度而得不到较好的抗误码效果;而编码端的多解码宏块平均失真的率失真优化方法的性能依赖于编码端解码器的个数,个数越多,抗误码性能越好,但编码器的复杂度也越高.基于以上两种方法的特点,本文提出了一种根据 Internet 包丢失信道的统计分布特征并结合 H.26L 的多分块匹配和多参考帧预测而得到的宏块全局率失真编码模式判决方法,在计算复杂度不高的前提下有效地提高了视频流在包丢失环

¹ 2002-08-26 收到, 2003-03-27 改回

国家高技术研究发展计划专项经费资助课题 (2002AA103087)

境下的抗误码性能。我们把它称作 GRDO(Global Rate-Distortion Optimality)。此方法以实际的 Internet 信道误码条件下仿真,并在编码端结合 Internet 包丢失的 Bernoulli 模型和 Gilbert 模型,我们分别称作 GRDOBM(GRDO based on Bernoulli Model) 和 GRDOGM(GRDO based on Gilbert Model), 与传统的抗误码算法相比得到了很大的改善。

2 H.26L 基于宏快的 GRDO 编码模式判决方法

信道中包丢失对视频质量产生影响的大小主要由源端的编码方案的选择、网络拥塞的状况及解码端的错误掩盖算法来决定。对于编码端来讲,通常使用较多的帧间编码宏块消除时间上的冗余度来获得更高的编码效率,但是使用较多的帧间编码会使包丢失条件下的误码在时间和空间上进行扩散进而使恢复的视频质量大大降低。采用帧内编码可以有效地终止误码的扩散,但要以较低的压缩效率为代价。因此,我们可以看出,在编码端找到一种合理的 Intra/Inter 编码模式选择方法变得非常重要^[4],它可以在尽量保证编码效率的条件下提高压缩视频的抗误码性能。

率失真优化的模式选择方法就是一种常用且有效的编码模式选择方法^[5]。在 TML8.0^[3],由于此方法中所有的失真函数只考虑编码的量化失真而不考虑实际信道的包丢失的统计分布特征,我们把它称作 NCCRDO(RDO of No Considering of Channel)。在 NCCRDO 中,编码块的最优 Intra/Inter 预测方法选择、参考帧选择及宏块的帧内 / 帧间编码模式选择都是采用率失真代价函数最小的原则。虽然编码效率较高,但算法的复杂度较大,并且抗误码特性较差。本文对上述的方法进行了一些简化,编码块的最优帧内 / 帧间预测方法选择和参考帧选择仍采用基于 Hadamard 正交变换的 SAD 判决,而仅对宏块的帧内 / 帧间编码模式选择采用率失真优化方法,并且,失真函数的计算要结合信道的统计特征,包括量化失真和信道误码失真。

若原始图像与编码端重建图像差的平方和用符号 $\sum \| \cdot \|^2$ 表示,用 f_n^i 表示当前编码帧 n 上像素点 i 的原始亮度值, \tilde{f}_n^i 表示当前帧 n 编码端解码重建像素点 i 的亮度值, \hat{f}_n^i 表示当前帧 n 解码重建像素点 i 的亮度值, MB(MacroBlock) 表示当前宏块, MODE 表示编码模式, QP(QuantizationParameter) 表示量化参数, $R_{MB}(s, \text{MODE}/\text{QP})$ 表示宏块在当前编码模式和相应的量化参数条件下亮度信号 s 的编码比特数。那么最小的代价函数如式 (1):

$$J_{\min} = \sum_{i \in \text{MB}} \|f_n^i - \tilde{f}_n^i\|^2 + \sum_{i \in \text{MB}} \|\hat{f}_n^i - \tilde{f}_n^i\|^2 + \lambda_{\text{MODE}} R_{\text{MB}}(s, \text{MODE}/\text{QP}) \quad (1)$$

$$\text{cd}_n = \sum_{i \in \text{MB}} \|\hat{f}_n^i - \tilde{f}_n^i\|^2 \quad (2)$$

其中式 (2) 的 cd_n (channel distortion) 表示由于 Internet 信道包丢失而产生的第 n 帧的失真函数,简称为信道失真函数,如果编码序列帧格式为 IPPP...PP,那么拉格朗日乘子 $\lambda_{\text{MODE}} = 5 \cdot e^{\text{QP}/10} \cdot [(\text{QP}+5)/(34-\text{QP})]$ 。

3 面向 Internet 信道的 H.26L 信道失真函数的计算

根据式 (3) 可得第 n 帧某宏块像素点 i 由于信道包丢失而产生的失真函数:

$$\text{cd}_n^i = \|\hat{f}_n^i - \tilde{f}_n^i\|^2 \quad (3)$$

考虑到信道误码的影响,在编码端是无法正确知道解码重建的像素值 \hat{f}_n^i ,显然误码重建值 \tilde{f}_n^i 为一个随机变量,为了便于分析信道失真的统计特性,本文做出如下假设:

(1) 考虑到 RTP/UDP/IP 传输方案下, Quarter Common Intermediate Format 格式 H.26L 视频流的码率比较低, 远低于 1400 字节的 Maximum Transmission Unit 上限^[6], 所以 H.26L 视频流采用每帧一个数据包的打包策略^[7];

(2) 编码端已知 Internet 信道的平均包丢失的统计特性;

(3) H.26L 解码器采用比较简单的误码掩盖策略. 如果当前数据包误码丢失, 直接复制上一个完整解码帧数据进行误码掩盖.

下面基于 Bernoulli 和 Gilbert 包丢失信道模型来讨论信道失真函数的特性.

3.1 GRDOGM 方法中信道失真函数的计算

通过对 Internet 上包丢失的大量测量显示, 包丢失的行为用两状态的马尔可夫链描述更为合理, 即包丢失的 Gilbert 模型^[8-9]. 如图 1 所示, R 为包被正确接收的状态; L 为网络拥塞

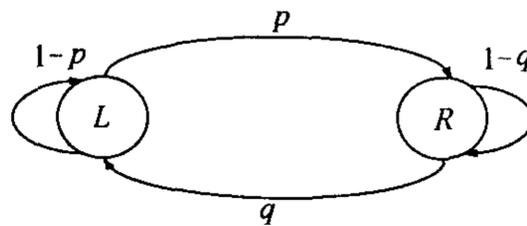


图 1 包丢失的 Gilbert 模型

或超过最大延时阈值而造成包丢失的状态, 参数 p, q 为 R 和 L 之间的转移概率. 可得 Gilbert 模型的包平均丢失概率为 $P_{Lav} = q/(p+q)$. 如果我们用 A 矩阵表示两状态的马尔可夫过程的转移矩阵, 用 p_{xy} 表示从 x 状态到 y 状态发生的概率, 那么 $A = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix} = \begin{bmatrix} p_{LL} & p_{LR} \\ p_{RL} & p_{RR} \end{bmatrix}$.

假设视频流的第 1 帧 (I 帧) 已经被正确的接收到, 在传输了 n 个包以后, 也就是 n 帧后, Gilbert 信道模型的转移矩阵就变成 A^n , 即 $A^n = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}^n = \begin{bmatrix} p_{LL}^{(n)} & p_{LR}^{(n)} \\ p_{RL}^{(n)} & p_{RR}^{(n)} \end{bmatrix}$. 其中 $p_{xy}^{(n)}$ 表示

输 n 个包后, 从状态 x 到状态 y 的转移概率. 我们在进行 Intra/Inter 编码模式判决时是以宏块为单元的, 所以我们首先必须求解当前编码模块能被正确接收和丢失的概率. 根据前面的假设, 我们采用的是 1 帧 1 包的打包策略, 如果当前帧已被正确接收, 那么该帧中所有的编码宏块的信息也被正确接收, 在第 1 帧 (I 帧) 也被正确接收的前提下, 根据 Gilbert 转移矩阵, 可得当前帧即当前帧所有宏块被正确接收的概率 $p_R^{(n)} = p_{RR}^{(n)}$; 相反地, 如果第 n 帧的包丢失, 即第 n 帧编码宏块也丢失, 在第 1 帧 (I 帧) 被正确接收的前提下, 可得当前宏块丢失的概率 $p_L^{(n)} = p_{RL}^{(n)}$.

3.1.1 Intra 编码宏块像素点信道失真函数的求解 在当前数据包以概率 $p_R^{(n)}$ 被正确接收的条件下, 此时解码端无信道误码失真: $\tilde{f}_n^i = \hat{f}_n^i$; 在当前数据包以概率 $p_L^{(n)}$ 丢包的条件下, 此时当前帧像素点使用上一解码帧中的相同位置像素点值 \tilde{f}_{n-1}^i 做掩盖. 相应的信道失真函数为 $\|\tilde{f}_n^i - \hat{f}_{n-1}^i\|^2$, 考虑到 H.26L 编码失真和信道误码失真的独立性, 故帧内编码方式下像素点的亮度误码失真函数期望为

$$\begin{aligned} cd_n^i &= p_R^{(n)} \cdot 0 + p_L^{(n)} \cdot \|\tilde{f}_n^i - \hat{f}_{n-1}^i\|^2 = p_L^{(n)} \cdot \|\tilde{f}_n^i - \hat{f}_{n-1}^i\|^2 \\ &= p_L^{(n)} \cdot \|(\tilde{f}_n^i - \tilde{f}_{n-1}^i) + \tilde{f}_{n-1}^i - \hat{f}_{n-1}^i\|^2 = p_{RL}^{(n)} \cdot (\|\tilde{f}_n^i - \tilde{f}_{n-1}^i\|^2 + cd_{n-1}^i) \end{aligned} \quad (4)$$

从式 (4) 不难发现: 帧内编码像素点的亮度误码失真主要是由误码概率下的当前帧误码掩盖失真而导致.

3.1.2 Inter 编码宏块像素点信道失真函数的求解 H.26L 采用的是多参考帧帧间预测模式, 设当前像素点的最佳运动匹配参考点为 \tilde{f}_{n-l}^j , 这里 $n-l$ 表示参考像素点在前第 l 个参考编码帧上, 同时设当前像素点运动估计后的残值为 \tilde{e}_n^i , 则基于信道误码的解码端的像素重建值如下: 当前包被正确接收时, $\hat{f}_n^i = \tilde{f}_{n-l}^j + \tilde{e}_n^i$; 当前包被丢失时, $\hat{f}_n^i = \tilde{f}_{n-l}^j$. 接收端当前像素的误码失真函数期望可表示为

$$\begin{aligned} cd_n^i &= p_R^{(n)} \cdot \|\tilde{f}_n^j - \hat{f}_n^j\|^2 + p_L^{(n)} \cdot \|\tilde{f}_n^i - \hat{f}_{n-1}^i\|^2 \\ &= p_R^{(n)} \cdot \|(\tilde{f}_{n-l}^j + \tilde{e}_n^i) - (\hat{f}_{n-l}^j + \tilde{e}_n^i)\|^2 + p_L^{(n)} \cdot \|\tilde{f}_n^i - \hat{f}_{n-1}^i\|^2 \\ &= p_R^{(n)} \cdot \|\tilde{f}_{n-l}^j - \hat{f}_{n-l}^j\|^2 + p_L^{(n)} \cdot \|\tilde{f}_n^i - \hat{f}_{n-1}^i\|^2 \\ &= p_R^{(n)} \cdot cd_{n-l}^j + p_L^{(n)} \cdot (\|\tilde{f}_n^i - \hat{f}_{n-1}^i\|^2 + cd_{n-1}^i) \\ &= [p_{RR}^{(n)} \cdot cd_{n-l}^j + p_{RL}^{(n)} \cdot cd_{n-1}^i] + p_{RL}^{(n)} \cdot \|\tilde{f}_n^i - \hat{f}_{n-1}^i\|^2 \end{aligned} \quad (5)$$

从式 (5) 可以看出: 帧间编码像素点的亮度误码失真函数主要由前面帧的误码失真扩散传递导致; 同时在较高误码概率下, 当前帧误码掩盖导致的失真贡献也很大.

3.2 GRDOBM 方法中信道失真函数的计算

Bernoulli 信道模型比较简单, 它只是 Gilbert 模型的一个特例, 当 Gilbert 模型中 $p+q=1$ 时即为 Bernoulli 模型. 描述包丢失的 Bernoulli 过程只有一个参数包丢失平均概率 p_L , 相应包正确接收的概率为 $1-p_L$. 信道失真函数的计算与 Gilbert 模型同理可得.

3.2.1 Intra 编码宏块像素点信道失真函数的求解

$$cd_n^i = p_L \left[\|\tilde{f}_n^i - \tilde{f}_{n-1}^i\|^2 + cd_{n-1}^i \right] \quad (6)$$

3.2.2 Inter 编码宏块像素点信道失真函数的求解

$$cd_n^i = [(1-p_L) \cdot cd_{n-l}^j + p_L \cdot cd_{n-1}^i] + p_L \cdot \|\tilde{f}_n^i - \tilde{f}_{n-1}^i\|^2 \quad (7)$$

4 测试结果比较与性能分析

4.1 仿真条件的设定

本仿真所使用的测试序列为 Foreman_qcif.yuv 和 Silent_qcif.yuv, 图像大小为 176×144 , 帧率设为 10 fps. H.26L 视频编码器的 QP 在整个过程中保持不变, 最大的参考帧数为 5, 最大的运动搜索范围设为 ± 12 个像素点. 实验中每个序列循环编码 4000 帧, 且每次序列循环的编码帧格式都是 IPP...P, 在解码端使用相应的错误掩盖策略.

Internet 信道误码条件根据 ITU 推荐的典型的误码分布为 3%, 5%, 10% 和 20% 的误码错误文件^[10], 若不考虑信道的延时信息, 其 Gilbert 模型和 Bernoulli 模型的参数如表 1 所示.

表 1 Gilbert 模型和 Bernoulli 模型的参数

| 包丢失概率 | | 3% | 5% | 10% | 20% |
|-----------|-------|--------|--------|---------|--------|
| Gilbert | p | 97.04% | 88.35% | 86.201% | 80.9% |
| | q | 3.125% | 5.43% | 10.65% | 2.204% |
| Bernoulli | p_L | 3.3% | 5.3% | 11.5% | 20.8% |

4.2 全局率失真方法在 Internet 信道上抗误码性能的测试

4.2.1 理想信道下的性能比较 在这里理想信道指的是视频流在信道中传输无任何包丢失. 我们比较下面 3 种方法在理想信道下的性能: (1) 基于 SAD Hadamard 变换的编码模式判决方

法^[11]。(2) 基于 NCCRDO 的编码模式判决方法^[12]。(3) 基于亮度信号的 GRDO 的 Intra/Inter 判决方法。

在 GRDO 的 Gilbert 模型和 Bernoulli 模型中, $p = q = p_L = 0$; QP 分别为 16, 18, 20, 22 下的平均码率和峰值信噪比的关系如图 2 所示。

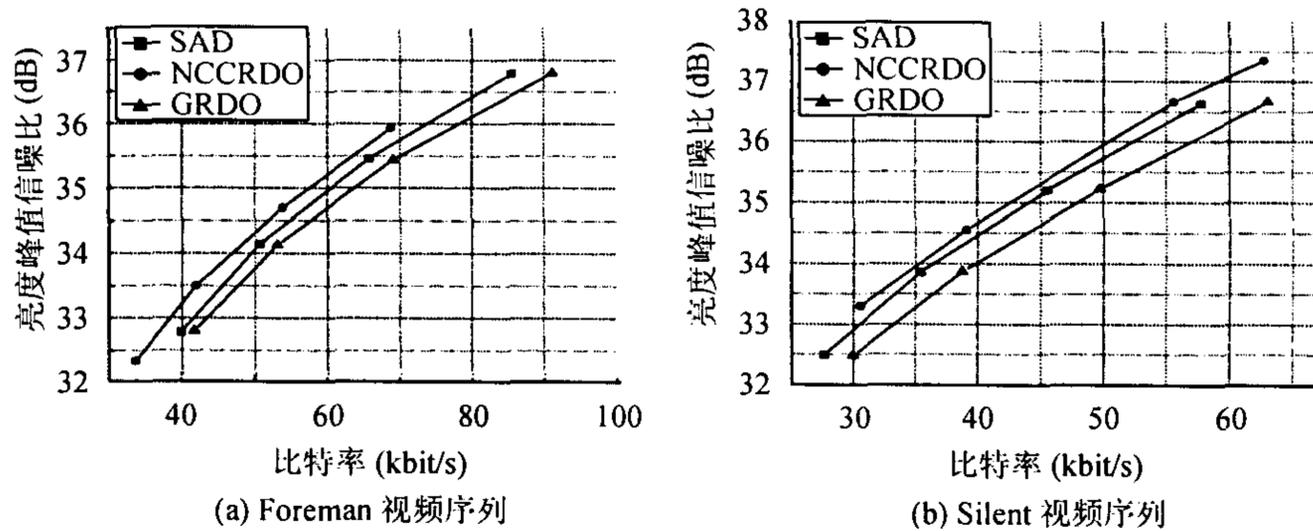


图 2 理想信道下几种方法的性能对比

从图 2 中我们看出: 在相同比特率的条件下, NCCRDO 的 PSNR 最高, 其次为 SAD, 最后为 GRDO。最大值与最小值之间的差为 0.6 dB 左右。

4.4.2 误码信道下 GRDOBM 和 GRDOGM 方法的抗误码性能测试 在实验中, 比较不同抗误码算法的性能要在相同或尽量相近的比特率下进行, 我们在各种方法下选取一定的 QP 值使码率约在 64 kbps 左右, 如果我们把常用的 GOB(Group Of Blocks) 刷新方法称为 GOBIU(GOB Intra Update), 那么 NCCRDO, GOBIU, GRDOBM 和 GRDOGM 的抗误码性能测试结果如图 3。与 NCCRDO 相比, 各种方法的 PSNR 在包丢失信道下的增益如表 2 所示。

表 2 与 NCCRDO 相比的各种方法 PSNR 在包丢失信道下的增益

| | | | 3% | 5% | 10% | 20% |
|---------------|--------------|--------------|---------------|--------------|--------------|--------------|
| Foreman 序列 | GOBIU 方法 | 码率 (kbps) | 59.14 | 59.41 | 60.39 | 62.03 |
| | | 量化参数 QP | 19 | 19 | 19 | 19 |
| | | PSNR 增益 (dB) | 0.635 | 0.196 | 0.46 | 0.294 |
| | GRDOBM 方法 | 码率 (kbps) | 60.78 | 58.14 | 60.15 | 67.93 |
| | | 量化参数 QP | 22 | 23 | 24 | 24 |
| | | PSNR 增益 (dB) | -0.175 | 1.396 | 4.11 | 5.134 |
| | GRDOGM 方法 | 码率 (kbps) | 60.26 | 57.97 | 64.63 | 63.0 |
| | | 量化参数 QP | 23 | 23 | 24 | 25 |
| | | PSNR 增益 (dB) | 0.125 | 1.776 | 4.51 | 5.584 |
| Silent 序列 | GOBIU 方法 | 码率 (kbps) | 63.8 | 59.75 | 61.14 | 64.69 |
| | | 量化参数 QP | 15 | 16 | 16 | 16 |
| | | PSNR 增益 (dB) | -0.164 | 0.086 | 0.334 | 0.362 |
| | GRDOBM 方法 | 码率 (kbps) | 63.005 | 59.805 | 66.27 | 63.84 |
| | | 量化参数 QP | 17 | 18 | 18 | 19 |
| | | PSNR 增益 (dB) | -0.51 | 0.156 | 1.304 | 2.86 |
| | GRDOGM 方法 | 码率 (kbps) | 65.88 | 62.217 | 64.2 | 60.50 |
| | | 量化参数 QP | 18 | 18 | 19 | 20 |
| | | PSNR 增益 (dB) | -0.16 | 0.456 | 2.434 | 4.16 |

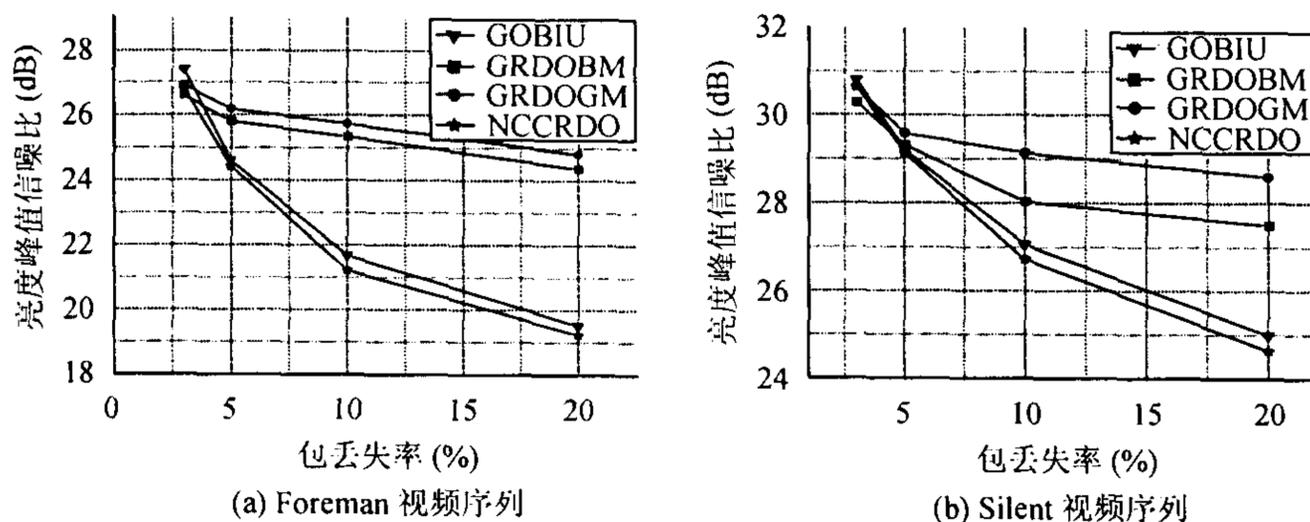


图3 误码信道下几种方法的性能对比

从表2和图3我们可以看出:

(a) Internet 信道上的包丢失对 H.26L 视频流的影响非常大, 并且 PSNR 随着平均丢失概率的增加急剧下降。

(b) 相对于 NCCRDO, 常用的 GOBIU 抗误码方法可以改善视频流的抗误码能力, 但改善的幅度不大, 一般不超过 0.6dB。这是因为该方法只是简单地根据包丢失概率来确定 GOBIU 刷新的频率并且不考虑图像的纹理特征及运动特征。

(c) 本文提出的基于信道特征的 GRDO 方法大大提高了流视频的鲁棒性, 并且平均包丢失概率越大, 性能提高的越明显。相对于 NCCRDO, Foreman 序列, 在 20% 丢包的情况下, PSNR 可提高 5.584dB, Silent 序列丢包 20% 的情况下, PSNR 可提高 4.16dB。

(4) 本文使用的信道包丢失模型有两种: Bernoulli 模型和 Gilbert 模型。由于 Gilbert 模型是一个二阶的马尔可夫过程, 这样的模型与实际的信道更加吻合, 更加有效地阻止了时间和空间误码的扩散。对于 Foreman 序列, 在 5% 丢包概率以上, GRDOGM 要比 GRDOBM 提高 0.4dB 左右, 对于 Silent 序列, 20% 丢包情况下, GRDOGM 比 GRDOBM 提高达 1.3dB。

5 结论

本文从较新的 H.26L 视频压缩编码标准出发, 并结合 Internet 信道的 Bernoulli 和 Gilbert 的包丢失模型, 提出了全局率失真优化的编码模式判决的 GRDOBM 和 GRDOGM 方法。由于 Gilbert 信道模型更加与实际的信道特性相匹配, 利用它的转移概率矩阵使编码的视频流具有更强的错误恢复能力。经实验证明, 在实际的 Internet 信道误码环境下, 本文提出的方法具有更强的抗误码性能和实用性, 得到了较好的效果。

参 考 文 献

- [1] ITU-T VCEG-N18, Performance evaluation of H. 26L, TML8 vs. H.263++ and MPEG-4, VCEG(SG16/Q6), Fifteenth Meeting, Santa Barbara, CA, September 2001.
- [2] ITU-T Q15-J-22, Video coding requirements for network robustness, VCEG(SG16/Q15), Tenth meeting, Osaka, 16-18 May, 2000.
- [3] Gisle Bjontegaard, H. 26L test model long term number(TML-8) draft0, VCEG (SG16/Q6), Fourteenth Meeting, Santa Barbara, CA, September 2001.
- [4] Zhang R, Shankar Regunathan L, Rose K. Optimal Intra/Inter mode switching for robust video communication over Internet, 33rd Asilomar Conference on Singals, Systems and Computers, Pacific Grove CA, USA, Oct. 1999: 332-336.
- [5] Sullivan G J, Wiegend T. Rate-distortion optimization for video compression. *IEEE Signal Processing Magazine*, 1998, 5(6): 74-90.

- [6] Casner S H, Frederick S. RTP: a transport protocol for real time applications, RFC1889, January 1996.
- [7] ITU-T VCEG-N52, H.26L over IP and H.324 framework, VCEG(SG16/Q6), Fifteenth Meeting, Santa Barbara, CA, USA, September 2001: 21-24.
- [8] Bolot J C, Turletti T. Adaptive error control for packet video in the Internet, in Proc. IEEE ICIP, Lausanne, September 1996, vol.1: 25-28.
- [9] Girod B, Stuhmueller K W, Link M, Horn U. Packet loss resilient Internet video streaming, SPIE Proc. Visual Commun. Image Processing, San Jose, California, Jan. 1999, vol.3653: 833-844.
- [10] ITU-T Document Q15-I-16, Proposed error patterns for Internet experiments, VCEG Q.15/SG16, Ninth Meeting, Red Bank, New Jersey, October 1999: 19-22.
- [11] H. 26L Test Model 5.2 software packet, ftp. pictel. com/Videosite/H.26L/, 2000.
- [12] H. 26L Test Model 8.0 software packet, ftp. pictel. com/Videosite/H.26L/, 2001.

杨 华: 女, 1976 年生, 博士生, 主要研究方向为视频压缩及图像通信.

余松煜: 男, 1942 年生, 教授, 博士生导师, 主要研究方向为视频压缩及图像通信.

杨松岸: 男, 1974 年生, 博士生, 主要研究方向为宽带网络和基于 IP 的 QoS.