

联合编码模式选择的码率控制算法

陈 川, 余松煜

(上海交通大学图像通信与信息处理研究所, 上海 200030)

摘要: 本文提出了一种丢包网络中联合信源信道码率控制算法. 该算法克服了以往码率控制方法和误码复原技术互不关联的弊病, 在统计失真模型的基础上, 联合最优地计算图像级量化参数和寻找最佳宏块编码模式, 在给定的受限码率下能充分利用可用信道带宽, 使得视频信源编码和信道传输总失真最小. 实验结果表明该算法能够获得比传统算法更高的性能增益.

关键词: H1 264/AVC; 率失真; 帧内刷新; 码率控制

中图分类号: TN9191.21 **文献标识码:** A **文章编号:** 0372-2112 (2004) 05-0763-06

Joint Coding Mode Selection and Rate Control Algorithm

CHEN Chuan, YU Songyu

(Inst. of Image Communication & Information Processing, Shanghai Jiaotong University, Shanghai 200030, China)

Abstract: We present a statistics analysis of the picture distortion caused by quantization, channel errors and subsequent error propagation, and propose rate-distortion estimation models which are functions of two parameters, i. e. quantization step and intra refreshing ratio. Based on these RD models, a joint optimal quantization step and intra refreshing ratio selection algorithm is developed by using Lagrange multiplier theory. Simulation results show that our algorithm can make the encoder adapting itself in time to the variation of the channel bandwidth, and significantly improve the video quality over packet loss networks.

Key words: H1 264/AVC; rate-distortion; intra refreshment; rate control

1 引言

率失真分析和码率控制问题一直是基于块变换和运动补偿类视频标准研究的重点^[1-5]. 码率控制策略通常建立在率失真模型的基础上, 通过调整量化参数来达到控制视频码率的目的. 然而传统的码率控制策略大多只考虑了信源编码所造成的失真, 而没有很好地结合压缩视频流在传输过程中发生的信道误码失真, 从而产生出的视频码流在误码环境中鲁棒性较差. 为了有效地抵抗信道误码, 很多误码恢复和掩盖技术应运而生, 其中宏块帧内刷新技术是一种相当简单而有效的误码复原方法, 然而过多的帧内编码宏块势必造成视频码率剧增, 破坏已经建立起来的码率控制规则. 实际上, 误码环境中为了实现最优的编码和抗误码性能, 宏块编码模式选择(帧内/帧间)应当联合码率控制一同被决定, 因为最优的编码模式选择和量化参数均取决于误码环境中率失真函数的最优控制点的选取^[6]. 本文基于此思想提出了一种联合编码模式选择的码率控制算法, 实验表明本算法既能很好地控制编码视频码率又能保持相当好的视频质量.

2 联合信源信道率失真模型

误码环境中图像的失真主要由编码失真(由量化引起)和信道误码失真两部分构成. 假定 $f(n, i)$ 为当前帧 n 像素点 i 上的原始值, 编码端重建此值为 $\hat{f}(n, i)$, 经过信道传输后在解码端的重建值 $\tilde{f}(n, i)$, 假设宏块的编码失真和信道误码失真统计独立, 并采用均方差 MSE(平均到像素级)作为失真衡量标准, 则编码宏块的总失真期望函数为:

$$D(n) = E \{ [f(n, i) - \tilde{f}(n, i)]^2 \} = Dq(n) + Dc(n) \quad (1)$$

其中, $Dq(n) = E \{ [f(n, i) - \hat{f}(n, i)]^2 \}$ 为信源编码失真, $Dc(n) = E \{ [\hat{f}(n, i) - \tilde{f}(n, i)]^2 \}$ 为信道误码失真.

2.1 编码失真模型

编码失真计算相对简单, 我们采用文献[2, 3]中的经典计算模型:

$$Dq(n) = \frac{1}{N} \sum_{j=1}^N X_j^2 \frac{[Q_j(n)]^2}{12} \quad (2)$$

式中, $Q_j(n)$ 为第 j 个宏块的量化阶; X_j 为宏块的加权系数, 用于控制重要的或者感兴趣区域(ROI)^[7] 宏块的编码失

真, 如对前景物体, 可以给较大的 X 权值; N 为一帧中所有宏块的总数. 当图像中每个宏块量化阶取值相同时, 图像编码失真计算模型表达为:

$$Dq(n) = \frac{Q^2(n)}{N} \sum_{i=1}^N \frac{X_i^2}{12} \quad (3)$$

21.2 信道误码失真模型

信道误码失真由不可靠传输信道误码引入, 它涉及到信道误码统计特性, 信源编码方法, 打包策略, 误码恢复和掩盖方法等多方面的影响, 因而对宏块总失真估计的准确程度往往取决于对信道误码失真的估计准确度. 需要注意的是, H.264/AVC 在视频压缩编码中使用了帧内和帧间两种预测技术来减少空域和时域相关性, 从而提高压缩效率, 但却使得信道误码扩散(包括时域和空域扩散)变得更为严重, 如果依然采用传统的误码扩散模型, 失真估计将变得不准确. 为此, 我们重新建立了考虑失真扩散的信道误码失真估计模型. 定义像素点的信道失真函数 $d(n, i) = f(n, i) - \hat{f}(n, i)$, 信道误码失真在统计意义上表达为:

$$\begin{aligned} Dc(n) &= E\{[d(n, i)]^2\} \\ &= E\{[f(n, i) - \hat{f}(n, i)]^2\} \\ &= Dc1(n) + Dc2(n) \end{aligned} \quad (4)$$

其中, $Dc1(n)$ 为当解码端宏块正确接收时的失真度; $Dc2(n)$ 为信道传输过程中发生错误时, 解码端采用误码掩盖而引入的失真. 为了简化分析, 我们假定误码信道中视频丢包率为 p , 每个视频包中含有相同个数的宏块, 则宏块中像素的丢失率就等于丢包率 $p^{[1]}$; 误码发生时, 解码器直接将当前发生误码宏块位置上的像素点值 $f(n, i)$ 用上一解码帧中相同位置宏块的像素点值 $f(n-1, i)$ 做掩盖; 帧间编码时不启用多参考帧, 只以前一解码重建帧作为当前帧间编码的参考帧. 由于不同编码模式下的失真也不完全相同, 故应对帧内和帧间编码模式分别进行失真分析.

21.2.1 帧内编码模式下像素点的失真估计 假定概率 $1-p$ 下, 解码端正确接收宏块数据, 有 $f(n, i) = \hat{f}(n, i)$, 此时解码端和编码端参考图像一致, 不存在信道误码失真, 即: $Dc1(n) = (1-p)E\{[f(n, i) - \hat{f}(n, i)]^2\} = 0$; 如果该宏块在传输过程中有误码发生, 误码发生概率 p , 解码端启用上述误码掩盖技术, 得到 $f(n, i) = \hat{f}(n-1, i)$, 考虑到编码失真和误码掩盖失真的统计独立性, 推导出帧内编码方式下的掩盖失真估计为:

$$\begin{aligned} Dc^I(n) &= Dc2(n) = pE\{[f(n-1, i) - \hat{f}(n, i)]^2\} \\ &= pE\{[f(n-1, i) - \hat{f}(n-1, i) \\ &\quad + \hat{f}(n-1, i) - \hat{f}(n, i)]^2\} \\ &= pE\{[f(n-1, i) - \hat{f}(n, i)]^2\} \\ &\quad + pE\{[d(n-1, i)]^2\} \end{aligned} \quad (5)$$

可以看出, 帧内编码模式下失真主要是由误码概率下当前帧新引入的失真和前一帧误码失真扩散影响所致. 由于 H.264/AVC 编码器对帧内编码使用了帧内预测技术, 误码掩盖失真仍然可以在空域中扩散, 我们称之为帧内预测误码扩散(Intra prediction error diffusion, IPED)失真. 式(5)中并没有考虑该失真扩散的影响. 由于该扩散与信道误码发生概率密切相关, 本文采用近似高斯分布来估计它, 将上述模型乘上一个

误码扩散概率系数 C_{IPED} 来估计 IPED 的影响. 假定与当前丢失宏块空域相邻的宏块中有 M 块采用帧内编码, 则在这 M 个宏块中, 传输错误发生 K 次的概率为: $p_e(k) = \binom{M}{k} p^k (1-p)^{M-k}$, $k=0, 1, 2, \dots, M$. M 个宏块中误码发生总概率为: $\sum_{k=1}^M p_e(k) = [1 - (1-p)^M]$, IPED 扩散系数表达为: $C_{IPED} = \frac{1}{\sqrt{p/2}} \exp[1 - (1-p)^M]$. 于是, 考虑 IPED 后的帧内宏块信道误码失真计算模型为:

$$\begin{aligned} Dc^I(n) &= pC_{IPED}E\{[f(n-1, i) - \hat{f}(n, i)]^2\} \\ &\quad + pC_{IPED}E\{[d(n-1, i)]^2\} \end{aligned} \quad (6)$$

21.2.2 帧间编码模式下像素点的失真估计 帧间编码模式下, 编码器采用时域运动预测技术, 对当前帧 n 像素 i 位置上的 $f(n, i)$ 用前一解码重建帧预测像素位置上的 $f(n-1, m)$ 作预测, 得到预测残差 $e(n, i)$, 该残差经过变换编码和量化, 得到 $\hat{e}(n, i) = \hat{f}(n, i) - \hat{f}(n-1, m)$. 为了提高编码效率, 编码器只传送残差 $\hat{e}(n, i)$ 与运动矢量到解码端. 当视频包在传输过程中损坏或丢失时, 解码器同样采用如前方法进行掩盖, 可知误码掩盖失真与帧内编码情况相同; 若解码端在 $1-p$ 概率下正确接收帧间编码数据, 则此时信道误码失真将继承前一帧预测宏块的信道误码失真. 因此帧间编码模式下像素点的信道误码失真表达为:

$$\begin{aligned} Dc^P(n) &= Dc2(n) + Dc1(n) \\ &= pC_{IPED}E\{[f(n-1, i) - \hat{f}(n, i)]^2\} + pC_{IPED} \\ &\quad \#E\{[d(n-1, i)]^2\} + (1-p)E\{[d(n-1, m)]^2\} \end{aligned} \quad (7)$$

21.2.3 连续帧间编码模式下像素点的信道误码失真估计 由于连续(continuous)帧间编码模式的宏块将在误码发生部位局部造成很大的失真^[8], 我们也采用同帧内误码扩散系数类似的方法, 对误码失真的扩散项乘上一个帧间误码扩散系数来模拟, 称该系数为帧间误码扩散系数(Inter prediction error diffusion, PPED). 假设当前帧 n 中即将编码的宏块在第 l 帧时 ($1 < n$) 使用了最近一次帧内编码, 则该宏块已经进行了 $L = n - l - 1$ 次连续的帧间编码, 则 PPED 扩散系数求得为: $C_{PPED} = \frac{1}{\sqrt{p/2}} \exp[1 - (1-p)^L]$, 同时, 帧间编码同样会延续帧内误码扩散导致的失真, 故综合上述失真扩散的影响可得帧间编码模式下信道误码失真计算表达为:

$$\begin{aligned} Dc^P(n) &= pC_{IPED}E\{[f(n-1, i) - \hat{f}(n, i)]^2\} \\ &\quad + pC_{IPED}C_{PPED}E\{[d(n-1, i)]^2\} \\ &\quad + (1-p)C_{PPED}E\{[d(n-1, m)]^2\} \end{aligned} \quad (8)$$

注意到在编码第 n 帧时, 公式(7), (8)中的 $E\{[f(n-1, i) - \hat{f}(n, i)]^2\}$ 项无法得到, 我们利用原始图像数据的 MSE 来估计它: $E\{[f(n-1, i) - \hat{f}(n, i)]^2\} = A \#E\{[f(n-1, i) - \hat{f}(n, i)]^2\} = A \#MSE_{org}(n, n-1)$, A 为一常量. 于是最终平均每个像素点的信道误码失真估计模型为:

$$\begin{cases} Dc^I(n) = pAC_{IPED}MSE_{org}(n, n-1) + pC_{IPED}E\{[d(n-1, i)]^2\} \\ Dc^P(n) = pAC_{IPED}MSE_{org}(n, n-1) + pC_{IPED}C_{PPED} \\ \quad \#E\{[d(n-1, i)]^2\} + (1-p)C_{PPED}E\{[d(n-1, m)]^2\} \end{cases} \quad (9)$$

可以看出, 帧内帧间编码模式下宏块的信道误码失真主要区别在于 $(1-p)C_{PPED}E\{d(n-1, m)\}^2$ 项, 即帧间编码像素点在正确接收条件下从前一帧预测宏块扩散来的失真, 它的大小与一帧图像中帧间编码宏块的多少有关, 也就是说误码信道中, 帧间编码宏块的数量越多则信道误码失真越大. 我们对失真统计分析目的之一就是要找到合适的帧内刷新率以满足在一定受限码率下总失真最小.

2.1.3 联合信源信道图像总失真模型

假设 $N^l(n)$ 为第 n 帧图像中帧内编码的宏块个数, 则帧内刷新率 $B(n) = N^l(n)/N$, 并且由文献[9]知, $E\{d(n-1, m)\}^2 = b\#E\{d(n-1, i)\}^2 = b\#Dc(n-1)$, 根据上述信道误码失真估计模型, 可以得到第 n 帧图像的平均信道失真模型:

$$\begin{aligned} Dc(n) &= B(n)\#Dc^l(n) + [1-B(n)]\#Dc^p(n) \\ &= ApC_{IPED}MSE_{og}(n, n-1) + \{[(1-p)bC_{PPED} \\ &\quad + pC_{IPED}C_{PPED}] - B(n)[pC_{IPED}(1-C_{PPED}) \\ &\quad - (1-p)bC_{PPED}]\}Dc(n-1) \end{aligned} \quad (10)$$

$$\begin{cases} \#_1 = (1-p)bC_{PPED} + pC_{IPED}C_{PPED} \\ \#_2 = pC_{IPED}(1-C_{PPED}) - (1-p)bC_{PPED} \end{cases}, \text{ 则有:}$$

$$Dc(n) = ApC_{IPED}MSE_{og}(n, n-1) + [\#_1 - B(n)\#_2]Dc(n-1) \quad (11)$$

综合编码失真式(3)和信道误码失真式(11), 一帧图像的 MSE 失真计算表达为:

$$\begin{aligned} D(n) &= Dq(n) + Dc(n) = \frac{Q^2(n)}{N} \sum_{j=1}^N \frac{X_j^2}{12} \\ &\quad + ApC_{IPED}MSE_{og}(n, n-1) + (\#_1 - B(n)\#_2)Dc(n-1) \end{aligned} \quad (12)$$

2.1.4 联合信源信道比特率模型

考虑到宏块帧内和帧间编码模式对实际编码比特率有很大的影响, 我们在图像级二项式比特率计算模型^[4,5]的基础上, 引入了帧内宏块刷新参数, 该模型更能真实地反映图像实际的编码比特数:

$$R(n) = MAD^{pre}(n)\# \left[\frac{A_1}{Q(n)} + \frac{A_2}{Q^2(n)} \right] + H(n) \quad (13)$$

$H(n)$ 为图像/宏块头部信息, 运动矢量信息和其它语法结构所花费的比特数; A_1, A_2 为一次和二次系数, 在每帧编码完成时借助于线性回归的计算方法及及时更新^[5]; $R(n)$ 为实际第 n 帧图像编码所花费的比特数; 与文献[4,5]不同的是, 式中 $MAD^{pre}(n)$ 为考虑宏块编码模式后的平均绝对误差值. 通常的 $MAD(n)$ 计算为: $\begin{cases} MAD^l(n) = E[|f(n, i)|], \text{ 帧内编码模式;} \\ MAD^p(n) = E[|e(n, i)|], \text{ 帧间编码模式;} \end{cases}$ 对 H.264/AVC 视频编码而言, 编码器在编码第 n 帧之前, 无法确定当前帧各宏块的编码模式, 因而无法直接得到 $MAD(n)$. 考虑到前后图像间内容的相关性, 我们利用前一帧图像中帧内宏块和帧间宏块的平均 MAD 值, 通过一线性预测来估算当前图像中帧内和帧间宏块的平均 MAD 值, 即:

$$\begin{aligned} \bar{D}(n) &= G\# \frac{1}{A^l(n-1)} \sum_{i=1}^{A^l(n-1)} |f(n-1, i)| + G_2; \\ \bar{D}^p(n) &= G\# \frac{1}{A^p(n-1)} \sum_{i=1}^{A^p(n-1)} |e(n-1, i)| + G_2 \end{aligned}$$

$\bar{D}(n), \bar{D}^p(n)$ 分别表示第 n 帧图像中帧内编码和帧间编码像素的平均 MAD 预测值, G_1, G_2 为两个线性预测系数, $A^l(n-1), A^p(n-1)$ 分别为前一帧编码帧中帧内编码宏块和帧间编码宏块的像素点个数. 于是, 第 n 帧像素级 MAD 值估计为: $MAD^{pre}(n) = B(n)@\bar{D}(n) + [1-B(n)]@\bar{D}^p(n)$.

3 联合编码模式选择的码率控制算法

上述率失真估计模型式(12)和式(13)表明误码信道中, 在给定信道丢包率的条件下, 图像的失真模型和比特率模型均与量化参数 $Q(n)$ 和宏块的帧内刷新率 $B(n)$ 有关. 由此本文提出了一种新的联合编码模式选择的码率控制算法, 该算法在图像级结合了视频抗误码技术和量化控制技术, 在给定的比特率约束条件下使得图像总失真最小. 具体描述为:

$$\begin{aligned} \text{minimize } D(n) &= D[Q(n), B(n)] \\ \text{subject to } R(n) &= R[Q(n), B(n)] \leq R^T(n) \end{aligned} \quad (14)$$

$R^T(n)$ 为预先分配给第 n 帧图像的目标比特数. 这实际是一个求解二元函数条件极值问题, 为了获得最优的量化参数和帧内刷新率, 我们利用拉格朗日乘数法(Lagrange multiplier)进行求解. 首先构造函数: $J = D[Q(n), B(n)] + K\{R[Q(n), B(n)] - R^T(n)\}$, K 为拉格朗日乘数因子, 然后分别对 $Q(n), B(n)$ 和 K 求偏导并令其为 0, 得到三个联立方程:

$$\begin{cases} \frac{5D[Q(n), B(n)]}{5Q(n)} + K \frac{5R[Q(n), B(n)]}{5Q(n)} = 0 \\ \frac{5D[Q(n), B(n)]}{5B(n)} + K \frac{5R[Q(n), B(n)]}{5B(n)} = 0 \\ R[Q(n), B(n)] = R^T(n) \end{cases} \quad (15)$$

经过一系列的求解步骤, 我们得到在给定信道丢包率 p 条件下, 图像级最优量化参数 $Q(n)$ 和帧内宏块刷新率 $B(n)$ 的计算表达式:

$$\begin{cases} a = A_2^2 - 2A_2F(n) \\ b = 2A_1A_2 - A_1F(n) \\ c = A_1^2 \\ \$ = b^2 - 4@a@c \\ F(n) = \frac{\#_2[R^T(n) - H(n)]@Dc(n-1)}{[D(n) - D^p(n)]@ \frac{2}{N} \sum_{j=1}^N \frac{X_j^2}{12}} \\ Q^{-1}(n) = \begin{cases} -\frac{c}{b}, & \text{if } \$ < 0 \\ -\frac{b+\sqrt{\$}}{2a}, & \text{otherwise} \end{cases} \\ B(n) = \frac{R^T(n) - H(n)}{[\frac{A_1}{Q(n)} + \frac{2A_2}{Q^2(n)}]@ [D(n) - D^p(n)]} - \frac{\bar{D}(n)}{D(n) - D^p(n)} \end{cases}; \quad (16)$$

本文提出的码率控制算法框图如图 1 所示. 该算法集中考虑可变比特率(VBR)情况下编码器在图像组(GOP)级, 图像级(Picture)和宏块(MB)级的比特分配和宏块编码模式选择问题. 该码率控制算法包括如下几个特点:

- (1) 将联合编码模式选择和量化参数控制嵌入到了码率控制策略中.

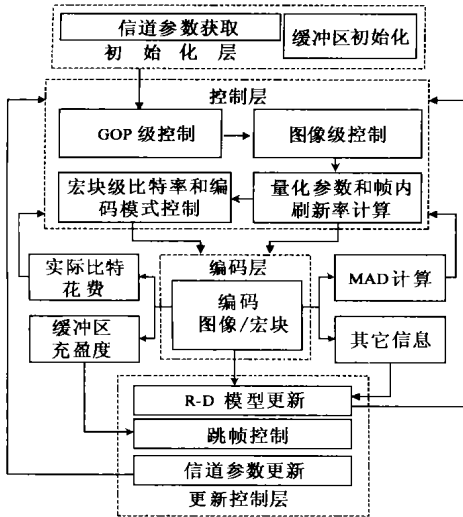


图1 码率控制框架图

(2)提出了一个更加合理的 GOP 级和图像级的 buffer 控制方法,该方法可根据时变信道带宽迅速调整比特分配以防止 buffer 的上溢和下溢,同时保证高效率地利用信道可用带宽。

(3)宏块级模式选择判决既考虑了减小信道误码失真又兼顾了提高 ROI 区域的图像质量。

具体而言,本码率控制策略包括 4 个部分:初始化层;控制层;编码层;参数更新层。

311 初始化层

编码器在初始化层完成两个任务:(1)信道参数获取,如当前可用信道带宽和信道丢包率;(2)编码器虚拟 buffer 大小的确定和初始化。本文采用一虚拟 buffer 来存储编码产生的压缩视频比特流,并根据时变信道带宽不断地更新 buffer 的充盈度。假定视频序列由多个 GOP 组成,每个 GOP 的结构为 IPP, P, I 表示帧内编码帧, P 表示帧间编码帧,每个 GOP 中包括 N_G 帧图像,用 $f_{k,l}$ ($k=1,2,\dots;l=1,2,\dots,N_G$) 代表第 k 个 GOP 中的第 l 帧图像,则在即将编码 $f_{k,l}$ 时刻的虚拟 buffer 的充盈度为:

$$B_v(k,l) = \min \left\{ \max \left\{ 0, B_v(k,l-1) + A(k,l-1) - \frac{R(k,l-1)}{F_r} \right\}, B_s \right\} \quad (17)$$

$A(k,l-1)$ 为第 $l-1$ 帧图像实际产生的比特数; $R(k,l-1)$ 为即将编码第 $l-1$ 帧时刻的信道可用带宽; F_r 为预先设定的编码帧率; B_s 为整个 buffer 的大小。虚拟 buffer 初始状态设定为 $B_v(1,1) = B_s/5$,这是考虑到编码器需要留有一定的空间来对付突发的编码码率剧变。相邻两个 GOP 间过渡时的 buffer 充盈度可表示为: $B_v(k+1,1) = B_v(k, N_G + 1)$ 。

312 GOP 级控制

考虑到信道带宽的时变性和前后两个 GOP 间的影响,每个 GOP 的比特分配为:

$$T_g(k) = \frac{R(k,1) \# N_G + B_v(k-1, N_G + 1) - B_s}{F_r} \quad (18)$$

每编码完成一帧图像,分配给当前 GOP 中剩余图像帧的比

数调整为:

$$T_g(k,l+1) = T_g(k,l) + \frac{R(k,l+1) - R(k,l)}{F_r} \# (N_g - 1) - A(k,l) \quad (19)$$

这样使得分配给 GOP 中各图像的比特数可随当前信道带宽的变化而及时调整,防止 buffer 的溢出。

313 图像级控制

对第一个 I 帧,采用作为输入参数的初始量化阶 Q_0 进行量化,后续 P 帧的比特分配初始计算为:

$$T_f(k,l) = (1-S) @ \frac{T_g(k,l)}{N_g - (l-1)} + S @ [T_f(k,l-1) - A(k,l-1)] \quad (20)$$

S 为加权因子,用于控制前一帧图像的比特数对当前帧目标比特数的影响。本文中 S 取值为 0.05。最终的图像比特分配须参考当前 buffer 的充盈度来决定:若 $B_v(k,l) + T_f(k,l) - \frac{R(k,l)}{F_r} < (1-H) \# B_s$,认为由公式(20)计算得到的目标比特数是合适的,因而将此目标比特数分配给即将编码的图像;若这个条件不满足,则初始计算得到的目标比特数将作减小,直到满足该条件为止。这里的 H 称为编码器 buffer 的安全空间 (safety margin),本文中 H 取值为 0.2。分配好一帧图像的目标比特数后,编码器采用公式(17)计算得到当前误码信道中的最佳量化参数 $Q(n)$ 和帧内宏块刷新率 $B(n)$ 。

314 宏块级控制

宏块级控制的任务包括两个方面:(1)宏块的编码模式判决;(2)宏块级比特分配。

由于在图像级控制中得到了第 n 帧图像的帧内宏块刷新率 $B(n)$,于是有 $N_l = B(n) \# N$ 个宏块将进行帧内编码。对一个宏块 x ,本文预先定义了一个视觉敏感度因子 $A(x)$,将它作为编码模式判决的依据,判决方法为:在 N 个宏块中选择 N_l 个 $A(x)$ 值最大的宏块作为帧内编码宏块。视觉敏感度因子定义为:

$$A(x) = A_r \# R(x) + A_p \# P(x) + A_l \# L(x) \quad (21)$$

$R(x)$ 为宏块原始亮度值的标准差,它能很好地反映宏块的复杂度; $P(x)$ 是宏块处于图像中的位置系数,它的采用考虑了 H1264/AVC 中 ROI 的概念^[7]; $L(x)$ 为宏块 x 作最近一次帧内刷新后的连续帧间编码帧数,这是为防止信道误码失真扩散过大而采用。 $R(x)$, $P(x)$ 和 $L(x)$ 均是对整帧图像所有宏块作规范化 (Normalized) 后的数值。三个加权因子须满足: $\begin{cases} \{A_r, A_p, A_l\} \in [0,1] \\ A_r + A_p + A_l = 1 \end{cases}$,其具体取值可以根据不同应用的侧重点不同而定。

由于本码率控制策略在图像级 buffer 控制中预留了一定的空间,故在宏块级的比特分配中可以适当地调整各宏块的量化参数,具体步骤如下:

314.1 宏块 x 的目标比特数

$$T_m(x) = \frac{X(x) \# A(x)}{\sum_{x=1}^N X(x) \# A(x)} \# T_f(n) \quad (22)$$

$T_f(n)$ 为分配给第 n 帧所有宏块的目标比特数; $X(x)$ 为宏块

x 的 MAD 值; A(x) 为第 x 个宏块的视觉敏感度因子.

314I 2 宏块量化参数调整 首先由宏块级二次比特率模型:

$$T_m(x) = X(x) \# \left[\frac{E_1}{Q(x)} + \frac{E_2}{Q^2(x)} \right],$$
求得第 x 个宏块的初步量化参数 Q(x), 然后确定宏块编码量化参数为:

$$Q(x) = \begin{cases} \max\{Q(x-1) - 2, Q(x)\}, & \text{if } Q(x) < Q(x-1) \\ \min\{Q(x-1) + 2, Q(x)\}, & \text{otherwise} \end{cases} \quad (23)$$

最终将上述得到的量化参数限制在 [0, 51] 范围内:

$$\text{if } (Q(x) < 0), Q(x) = 0$$

$$\text{if } (Q(x) > 51), Q(x) = 51$$

314I 3 宏块比特率模型更新 每个宏块编码完成后二次比特率模型中 E₁, E₂ 需要更新, 为下一个宏块获得更准确的比特率模型参数, 计算方法同图像级比特率参数更新.

315 跳帧控制和图像级参数更新

每帧图像编码前, 编码器首先需要检查 buffer 的充盈度和目标比特数的大小, 如果当前 buffer 的充盈度加上分配给当前图像的目标比特数大于 (1-H)#B_s, 则下一帧图像将被跳过. 跳帧后的 buffer 充盈度调整为:

$$B_v(k, l+2) = \max \left\{ B_v(k, l) - \frac{R(k, l+1)}{F_r}, \frac{B_s}{5} \right\} \quad (24)$$

每帧图像编码完成后, buffer 的充盈度更新为: B_v(k, l) = max { B_v(k, l-1) + A(k, l) - R(k, l-1)/F_r, B_s/5 }; 图像级失真通过公式 (12) 迭代计算更新; 所有宏块的实际量化参数取平均后得到本帧图像的平均量化阶, 并利用线性回归的计算方法更新参数 A₁, A₂, 为下一帧图像获得更为准确的比特率模型参数.

4 实验

本节给出了联合编码模式选择的码率控制算法在修改的

JVT/M2.0 软件上的实验结果. 信道丢包和误码分布使用 ITU 推荐的平均丢包率为 3%、5%、10% 和 20% 的错误图样文件来模拟实现.

(1) 首先测试了联合信源信道的图像总失真估计模型 (12) 的准确程度. 测试序列采用 QCIF, 亮色比 4:2:0 格式的 Foreman 和 News, 信道丢包率设定为 5%. 图 2 给出了采用估计模型计算和实际失真的 MSE 值对编码帧数的比较结果. 从图中可以看出, 采用联合信源信道图像失真估计模型得到的 MSE 值相当接近于实际失真的 MSE 值, 由此可知本文建立的失真模型较准确地估计了丢包网络中图像的编码失真和信道误码失真.

(2) 比较了任意帧内刷新 (RBIC) 方法^[10]同本文提出的基于率失真分析的宏块级编码模式选择算法在不同丢包率下的 PSNR 性能. 图 3 给出了 Foreman 编码码率 64kbps, 帧率 10fps 和 News 序列编码码率 64kbps, 帧率 15fps 的比较结果. 从图中不难看出, 信道丢包现象对压缩流解码后的视频质量影响相当大, 随着信道丢包率的上升, 解码端视频解码后的 PSNR 值急剧下降, 本文的算法由于准确地估计了编码失真, 信道误码失真和宏块连续帧间编码所产生的局部误码扩散的影响, 并综合考虑了宏块的编码复杂度, 对失真较大和感兴趣区域的宏块采用了帧内编码模式, 因而典型丢包情况下的视频质量明显好于 RBIC 方法. 图中 PSNR 曲线还表明, 信道丢包率越高, 本文提出的算法效果越好, 尤其对于运动较大的测试序列效果更为明显.

(3) 对本文提出的联合信源信道码率控制策略同 TM5 码率控制算法在比特分配波动方面的性能作了比较. 图 4 给出了 Foreman 编码码率 74kbps, 帧率 10fps, Hall Monitor 编码码率 64kbps, 帧率 15fps 的实际比特花费比较曲线. 从中可以看出本文的算法实际产生的比特分布较为平坦, 波动情况明显小

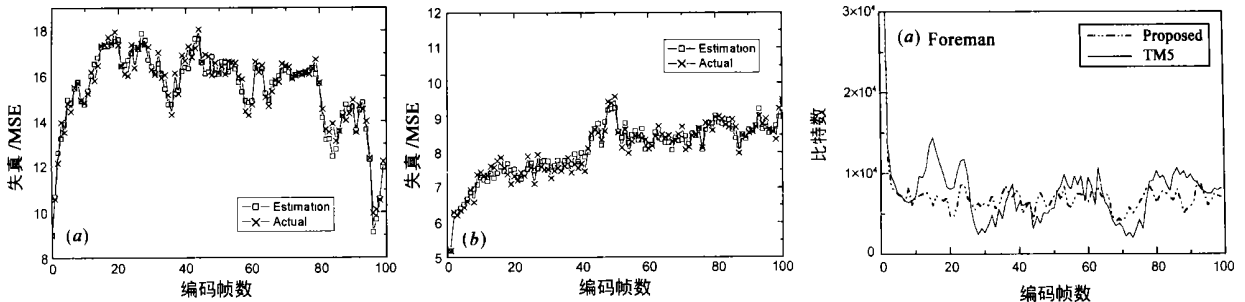


图 2 失真估计模型比较结果

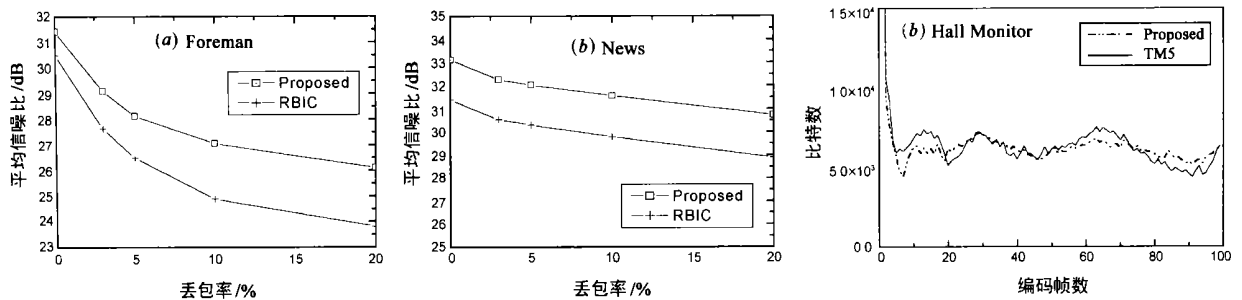


图 3 典型丢包率下的 PSNR 比较结果

图 4 编码比特分布比较结果

于 TM5 算法. 这说明该方法在比特分配自适应调整方面比 TM5 算法更为有效.

5 结论

本文在对图像编码失真和信道误码失真的统计分析基础上建立了误码信道中的图像失真估算模型; 同时考虑到宏块的编码模式对比特率的影响, 改进了二次比特率模型. 基于构建的率失真模型, 提出了一种联合宏块模式选择和量化参数控制的码率控制算法. 实验表明该算法在丢包网络环境中能够根据网络状况及时自适应地调整比特分配和量化参数, 并且具有很好地抵抗信道误码的性能.

参考文献:

- [1] Zhang R, et al. Video coding with optimal inter/ intra2mode switching for packet loss resilience [J]. IEEE J Select. Areas Commun, 2000, 18 (6): 966- 976.
- [2] Hang H M, et al. Source model for transform video coder and its application2Part I: Fundamental theory [J]. IEEE Trans. On CSVT, 1997, 7 (2): 287- 298.
- [3] Corbera J R, et al. Rate control in DCT video coding for low\$ delay communications [J]. IEEE Trans. on CSVT, 1999, 9(1): 172- 185.
- [4] Chiang T, et al. A new rate control scheme using a nov rate2distortion model [J]. IEEE Trans. on CSVT, 1997, 7(1): 246- 250.
- [5] Lee H J, et al. Scalable rate control for MPEG24 video [J]. IEEE Trans. on CSVT, 2000, 10(6): 878- 894.
- [6] Sun F F, et al. MPEG coding performance improvement by jointly opt2mizing coding mode decision and rate control [J]. IEEE Trans. on CSVT, 1997, 7(3): 449- 458.
- [7] Hannuksela M M. New image segmentation method [Z]. ITUT VCEG Document, 2001.
- [8] Kim C W, et al. High complexity mode decision for error prone environment [Z]. JVT Document, 2002.
- [9] He Z, et al. Joint source channel rat2distortion analysis for adaptive mode selection and rate control in wireless video coding [J]. IEEE Trans. on CSVT, 2002, 12(6): 511- 523.
- [10] Cote G, et al. Optimal intra coding of blocks for robust video communication over the Internet [J]. Image Commun, 1999, 15(12): 25- 34.

作者简介:



陈川男, 1974 年出生于四川隆昌, 博士研究生, 现在上海交通大学图像通信与信息处理研究所攻读博士学位, 目前研究方向为视频压缩编码, 易错信道视频鲁棒传输, 流媒体技术.



余松焜男, 1941 年出生于上海, 1963 年毕业于上海交通大学电子工程系, 现为上海交通大学图像通信与信息处理研究所副所长, 教授, 博士生导师, 研究领域包括数字电视, 图像通信, 图像处理等, 已经发表 5 数字图像处理 6 5 现代图像压缩编码技术 6 等著作和论文多篇, 并获得多项科研奖励.