

H.264视频鲁棒传输的最优宏块模式选择

杨天武, 彭强, 诸昌铃

(西南交通大学信息科学与技术学院 成都 610031)

【摘要】分析了受损宏块的信道失真及其差错扩散,研究了宏块编码模式对信源失真和信道失真的影响,提出了结合信源信道编码的一种基于率失真优化的快速模式选择算法。通过估算不同宏块模式的信源失真和信道失真,自适应地选择最佳编码模式,使得视频序列的整体传输失真最小。模拟实验显示,相对于H.264参考软件的模式选择算法,该算法显著降低了计算复杂度并提高了视频差错恢复性能。

关键词 视频传输; H.264/AVC; 模式选择; 率失真优化

中图分类号 TN 919.8

文献标识码 A

Optimal Macroblock Mode Selection for H.264 Video Robust Transmission

YANG Tian-wu, PENG Qiang, ZHU Chang-qian

(School of Information Science & Technology, Southwest Jiaotong University Chengdu 610031)

Abstract The channel distortion and error propagation of damaged macroblock are analyzed. The impact of macroblock modes on source distortion and channel distortion is discussed, and thus a fast macroblock mode selection algorithm is proposed based on joint source-channel rate distortion optimization. By evaluating the channel distortion and source distortion of all macroblock modes, the optimal macroblock mode is selected adaptively to minimize transmission distortion. Compared with the mode selection algorithm in H.264 reference software, our simulation results show that the proposed algorithm reduces computational complexity significantly and gets better error resilient.

Key words video transmission; H.264/AVC; mode selection; rate distortion optimization

目前,视频处理技术广泛采用预测编码。当视频在易受干扰的网络传输时,传输差错在时空域扩散,导致视频质量迅速降低。受损宏块的帧内编码更新能阻止差错扩散,提高视频差错恢复能力^[1-2],然而却以降低信源编码效率为代价。最小化视频传输失真的解决策略是联合信源信道的率失真优化(Rate Distortion Optimization, RDO)编码^[3-4]。它包括两项关键技术:传输失真估计和编码模式选择。文献[3]提出了基于整像素精度递推的传输失真精确估计模型,但H.264采用1/4像素精度的运动补偿,使传输失真估计不够准确,严重影响了RDO模式选择的性能。文献[4]提出了传输失真的帧级递推估计模型,通过RDO决定当前帧的帧内宏块更新率。由于随机选择宏块更新,因此可能导致传输失真较小区域连续更新,传输失真较大区域很少更新,而不同活性的宏块需要不同的更新率。H.264测试模型在编码端使用很多解码器仿真接受端解码,能根据较准确的传输失真估计,进行宏块级的有损RDO模式选择。这些方法仅考虑了当前帧的传输失真,缺乏考虑失真扩散和帧内更新对传输失真的影响,难以获得很好的差错恢复性能。

本文分析了H.264视频的传输失真特征,设计出宏块传输失真的快速估计策略,通过研究信道失真的扩散和帧内宏块更新对信源失真的影响,提出了一个联合信源信道编码的RDO宏块模式选择算法,能有效地改善实时视频传输的鲁棒性。

1 宏块传输失真的快速估计

视频传输失真主要包括由量化引起的编码失真和传输差错引起的信道失真两部分。信道失真又包括扩

收稿日期:2005-12-24

基金项目:铁道部科技研究开发项目(2003X040-A)

作者简介:杨天武(1971-),男,博士生,主要从事视频传输技术方面的研究。

散失真和差错掩盖失真。扩散失真指信道失真通过运动补偿路径扩散到当前帧的失真；差错掩盖失真指掩盖当前帧传输出错所导致的失真。假定采用均方差作为失真衡量标准，实验表明编码失真和信道失真以及扩散失真和差错掩盖失真基本上统计独立。

1.1 信道失真与编码失真

假定第 n 帧的第 i 像素 $f_n(i)$ 在编码端和解码端的重建值分别为 $\hat{f}_n(i)$ 和 $\tilde{f}_n(i)$ 。因视频传输差错，它们的值可能不同。定义传输失真 $D(n)$ 、信源失真 $D_S(n)$ 和信道失真 $D_C(n)$ 分别为：

$$D(n) = \sum_{i=0}^T [f_n(i) - \tilde{f}_n(i)]^2, \quad D_S(n) = \sum_{i=0}^T [f_n(i) - \hat{f}_n(i)]^2, \quad D_C(n) = \sum_{i=0}^T [\hat{f}_n(i) - \tilde{f}_n(i)]^2 \quad (1)$$

式中 T 为一帧的总像素数。传输失真基本上等于信道失真与编码失真之和^[4]，即：

$$D(n) = 2 \sum_{i=0}^T [f_n(i) - \hat{f}_n(i)] [\hat{f}_n(i) - \tilde{f}_n(i)] + D_S(n) + D_C(n) \approx D_S(n) + D_C(n) \quad (2)$$

H.264在进行宏块模式选择时，编码失真已经计算得到，因此只需估算信道失真。

1.2 扩散失真与掩盖失真

当正确接受第 n 帧数据时，假定像素 i 采用参考帧的像素 j 预测。当传输出错时，像素 i 采用前一帧的像素 k 掩盖，此时掩盖失真 $D_Y(n)$ 和扩散失真 $D_K(n)$ 分别定义为：

$$D_Y(n) = \sum_{i=1}^T [\tilde{f}_{n-1}(k) - (\tilde{f}_{n-1}(j) + e_n^i)]^2, \quad D_K(n) = \sum_{i=1}^T [\tilde{f}_{n-1}(j) + e_n^i - \hat{f}_n(i)]^2 \quad (3)$$

式中 e_n^i 为像素 i 的编码残差。当传输出错时解码端的重建值为参考帧像素 j 的值：

$$D_C(n) = \sum_{i=0}^T [\hat{f}_{n-1}(k) - \tilde{f}_n(i)]^2 = D_Y(n) + D_K(n) + 2 \sum_{i=1}^n [\tilde{f}_{n-1}(k) - \tilde{f}_{n-1}(j) - e_n^i] [\tilde{f}_{n-1}(j) + e_n^i - \hat{f}_n(i)] \approx D_Y(n) + D_K(n) \quad (4)$$

为考察扩散失真与掩盖失真的统计独立性，采用Internet视频差错模板^[5]在不同丢包率下进行模拟解码，强行丢失第30和60帧，计算这两帧信道失真与扩散失真之差的均值，即掩盖失真估计值，结果如表1所示。从表中可知，掩盖失真估计值与掩盖失真实际值近似，误差基本在3%以内；丢包率为0%时，掩盖失真估计值为实际值。

表1 不同丢包率下的信道掩盖失真估计值

测试序列	丢包率/(%)	第30帧	第60帧
"Mother&Daughter"	0	17.9	78.0
	5	18.6	80.3
	10	18.0	80.8
"Carphone"	0	88.3	52.6
	5	86.8	52.6
	10	87.8	53.6

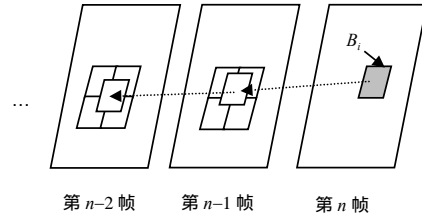


图1 受损子宏块的反向运动跟踪

1.3 编码宏块的信道失真快速估计

每帧编码完成后，编码端模拟解码器掩盖当前帧，并计算出每个子宏块(8×8像素)的掩盖失真。如图1所示，将子宏块看作运动整体，并假定其运动向量为其中心像素的运动向量，因为子宏块内基本不存在细节运动，所以假设很合理，其信道失真采用后向运动跟踪方法近似估计。当子宏块 B_i 反向运动到第 f 帧时，会覆盖该帧的4个子宏块，记第 k 个被覆盖的子宏块为 B_f^k ，被覆盖的像素个数为 T_f^k 。考虑编码失真和信道失真以及扩散失真与差错掩盖失真的统计独立性。 B_i 的信道失真可快速估算为：

$$D_C(B_i) = p D_Y(B_i) + \prod_{j=n-6}^n \theta_j \sum_{k=1}^4 \left(\frac{T_f^{n-6}}{64} D_C(B_f^{n-6}) \right) + p \sum_{f=n-6}^{n-1} \left(\left(\prod_{j=f+1}^n \theta_j \right) \sum_{k=1}^4 \left(\frac{T_f^k}{64} D_Y(B_f^k) \right) \right) \quad (5)$$

式中 p 为丢包率； θ_j 为扩散失真的影响因子。由于帧内更新和空间滤波导致，其值为：

$$\theta_j = \sum_{k=1}^4 T_{j+1}^k \beta_j \beta_m / 64 \quad (6)$$

式中 β_i 为空间滤波器对扩散失真的影响因子, 相关于具体视频特征, 取值在 0.85~1.09 之间, 可取固定值, 也可动态适应; β_m 为编码模式对扩散失真的影响因子, 当帧内编码时扩散失真只能以概率 p 扩散到后续帧, 其他情况为 1。当前宏块的信道失真等于它所包含的 4 个子宏块的信道失真之和。

2 编码模式的选择

2.1 帧内更新对编码失真的影响

如果当前帧可用比特数为 R_1 , 假定 3 种编码: (1) 用步长 Q_1 量化, 编码比特数为 R_1 , 编码失真为 $D_{S_1}(n)$ 。(2) 用步长 Q_2 量化, 并且帧内编码更新部分宏块, 编码比特数也为 R_1 , 编码失真 $D_{S_2}(n)$ 。(3) 用步长 Q_2 量化, 不进行帧内宏块更新, 编码比特数 R_2 。第(3)种和第(2)种编码的量化步长相同, 编码失真差异极小, 等于 $D_{S_2}(n)$ 。第(2)种编码比第(3)种编码增加的冗余编码比特数为 $R_A = R_1 - R_2$, 冗余比特使得差错恢复能力增强, 但比第(1)种编码的编码失真增加 $D_{S_A}(n) = D_{S_2}(n) - D_{S_1}(n)$ 。

对于分级量化编码, 编码失真与编码数据率存在关系^[6]:

$$D_S(n) = a2^{-2R} \sigma \quad (7)$$

式中 a 为参数; R 为单位像素的平均编码比特数; σ 为运动补偿残差失真, 近似等于两部分失真之和, 其中, 第一部分为原始图像的运动补偿残差失真, 第二部分为由编码失真引起的运动补偿失真。假定第(1)种和第(3)种编码的 σ 值分别为 σ_1 和 σ_2 , 则下面的关系成立:

$$\sigma_1 = \sigma_2 - \beta D_{S_A}(n) \quad (8)$$

式中 β 为常数^[4], 与运动补偿和空间滤波器相关。由式(7)、(8)可得:

$$D_{S_A}(n) = D_{S_2}(n) - D_{S_1}(n) = a \left(\sigma_2 2^{\frac{-2}{T}(R_1 - R_2)} - (\sigma_2 - \beta D_{S_A}(n)) 2^{\frac{-2}{T}R_1} \right) \quad (9)$$

式中 T 为像素总数。对于第(2)种编码, 单位冗余编码比特导致的编码失真增量为:

$$\phi(n) = \frac{d(D_{S_A}(n))}{d(R_A/T)} = 2 \ln 2 a \sigma_2 2^{\frac{-2}{T}R_2} \left/ \left(1 - a\beta 2^{\frac{-2}{T}R_1} \right) \right. \quad (10)$$

在运动搜索中, 能获得宏块所有编码模式的运动补偿残差失真。率失真优化过程中, 能够获得宏块所有编码模式的编码比特和编码失真及其编码器实际采用的编码模式。因此编码完成后, 就可得到 R_1 、 R_A 和 σ , 并计算出 $\phi(n)$ 。实验观察各帧的 ϕ 值非常接近, 当前帧采用前一帧 ϕ 值编码。

2.2 信道失真的扩散因子

当受损宏块被帧内更新时, 降低了宏块在当前帧的信道失真, 同时也降低了信道失真的扩散。假定第 i 宏块被帧内更新的间隔为 n 帧, 即第 i 宏块每 n 帧被帧内编码一次。如果第 i 宏块的上下左右宏块的更新间隔为 l_U 、 l_D 、 l_L 和 l_R , 则第 i 宏块的扩散因子为:

$$\varphi_i = (1 - \beta_i^{(4l_U + l_D + l_L + l_R)/8}) / (1 - \beta_i) \quad (11)$$

2.3 编码模式的确定

在 H.264 中, 数据率受限的有损率失真优化的模式选择是最小化下的拉格朗日代价函数:

$$m_o = \arg \min_{m \in M} [D_S(MB_i, m) + D_C(MB_i, m) + \lambda_m R(MB_i, m)] \quad (12)$$

式中 M 为宏块 MB_i 的所有有效编码模式; λ_m 为拉格朗日乘数。式(12)没有考虑帧内更新和差错扩散对传输失真的影响。考虑到这些因素, 并结合式(5)、(9)和(10), 使得传输失真最小的宏块模式为:

$$m_o = \arg \min_{m \in M} [D_S(MB_i, m) + \lambda_m \varphi_i D_C(MB_i, m)] / [\phi(n-1) + \lambda_m R(MB_i, m)] \quad (13)$$

3 模拟实验

模拟实验在 H.264 测试模型 JM9.8 上进行, 仿真文献[5]实验 4 的 GOB 交织打包传输技术, 使用 JM9.8 的差错掩盖和流控技术。采用 IPPP 的编码模式编码 100 帧, 编码比特率为 128 Kb/s。设置快速运动搜索, 并开启所有的编码模式。采用 Internet 数据包丢失模板文件“3”、“5”、“10”和“20”, 实际丢包率分别为 3.3%、5.6%、

11.5%和20.8%^[5]。测试计算机配备256 MB内存和P 1.2 GHz处理器。

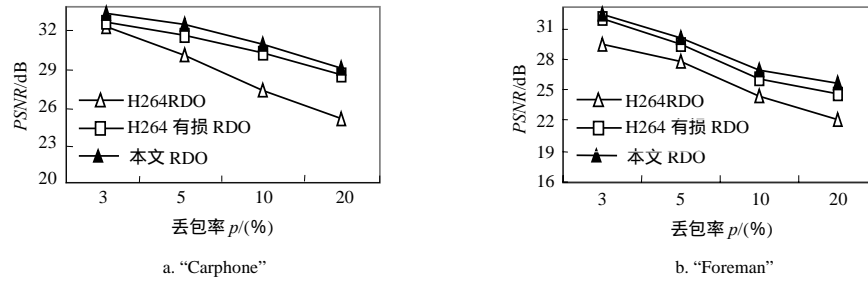


图2 三种率失真优化编码在不同丢包率下的性能对比

图2显示了三种率失真优化的模式选择算法在“Carphone”和“Foreman”序列中的差错恢复性能对比。H.264RDO指H.264的正常率失真优化编码，差错恢复能力很弱，丢包率为20%时PSNR性能低于本文RDO算法4 dB左右。H.264有损RDO指H.264测试模型的针对丢包网络的率失真优化编码，它在编码端使用30个解码器仿真客户端解码，能精确估算信道失真，但是需要大量缓存和计算资源，如表2所示。同时因为没有H.264RDO从整个序列失真最小角度优化，差错恢复性能低于本文RDO算法0.3~1.1 dB。

表2 不同编码方法估计信道失真所占用的时间和缓存

编码方法	“Foreman”		“Carphone”	
	每帧计算时间/ms	占用缓存/帧	每帧计算时间/ms	占用缓存/帧
本文RDO	11.1	1	10.9	1
H.264有损RDO	2 760.0	60	2 537.0	60

4 结论

本文根据传输失真的统计特性，提出了基于子宏块后向跟踪的快速信道失真估计算法，该算法仅仅通过一次解码器的差错掩盖仿真，就能比较精确地估计信道失真，与H.264测试模型的信道估计算法相比，计算复杂度显著降低。此外，还分析了帧内更新对编码失真的影响和信道失真的传播特性，从传输失真最小角度提出了最优宏块编码模式的选择算法，与H.264有损RDO相比，差错恢复性能有明显提高。

参考文献

- [1] Judy Y L, John V. Adaptive Intra block update for robust transmission of H.263[J]. IEEE Trans. on Circuits and System for Video Technology, 2000, 10(1): 30-35.
- [2] Wang Y, Zhu Q F. Error control and concealment for video communication: A review[J]. Proc. of the IEEE, 1998, 86(5): 974-997.
- [3] Zhang R, Regunathan S L, Rose K. Video coding with optimal intra/inter mode switching for packet loss resilience[J]. IEEE Journal of Selected Areas in Communications, 2000, 18(6): 966-976.
- [4] He Zhihai, Cai Jianfei, Chen Changwen. Joint source channel rate-distortion analysis for adaptive mode selection and rate control in wireless video coding[J]. IEEE Trans. on Circuits and System for Video Technology, 2002, 12(6): 511-523.
- [5] Wenger S. H.264/AVC over IP [J]. IEEE Trans. on Circuits and System for Video Technology, 2003, 13(7): 645-656.
- [6] Sullivan G J, Wiegand T. Rate-distortion optimization for video compression[J]. IEEE Signal Processing Mag., 1998, 15(11): 74-90.

编辑 漆蓉