

基于 H. 264 具有抗分组丢失能力的帧内刷新算法

宋 彬, 蒋小兵, 秦 浩, 郭春芳

(西安电子科技大学 ISN 国家重点实验室, 陕西西安 710071)

摘 要: 本文提出一种基于 H. 264 的具有抗分组丢失能力的帧内刷新算法. 该算法首先建立帧内刷新尺度矩阵, 来记录每个宏块被后继帧参考的重要程度; 然后, 基于帧内刷新尺度矩阵提出了丢包环境下的端到端率失真模型 (ME2ERD); 接下来, 利用 H. 264 中补充增强信息域进行信息交互, 结合 ME2ERD 模型自适应进行帧内刷新, 从而降低分组丢失对恢复图像质量的影响. 实验结果表明在高丢包率环境下, 使用本文提出的帧内刷新算法, 与使用随机帧内刷新算法相比, 恢复图像的平均 PSNR (峰值信噪比) 提高 5dB 以上. 算法与 H. 264 标准兼容, 且运算复杂度低, 具有实用价值.

关键词: 抗分组丢失; 帧内刷新; 补充增强信息域

中图分类号: TN919. 8 **文献标识码:** A **文章编号:** 0372-2112 (2009) 01-0221-04

A Packet Loss Resilient Intra Refresh Algorithm Based on H. 264

SONG Bin, JIANG Xiao-bing, QIN Hao, GUO Chun-fang

(National Key Laboratory on Integrate Services Network, Xidian University, Xi'an, Shaanxi 710071, China)

Abstract: An H. 264-based packet loss resilient intra refresh algorithm is presented in this paper. Firstly, the intra-refresh matrix is created to record the dependent relation of the blocks in the current frame to that in the reference frame. Secondly, based on intra-refresh matrix, an end-to-end rate distortion (ME2ERD) model for packet loss network is proposed to estimate the motion mode of a block. Then, by using supplemental enhancement information (SEI) domain, an adaptive intra refresh method is achieved to reduce the quality degradation due to error propagation. Compared with random intra refresh method, the proposed technique can improve the video quality more than 5dB over variable and error-prone channel. Furthermore, this algorithm is compatible with H. 264 without any increasing channel burden and arithmetic complexity, and can be used in practice.

Key words: packet loss resilience; intra refresh; SEI domain

1 引言

基于 IP 网络的视频通信中, 由于“尽力”传输会导致分组丢失, 造成接收端的恢复图像质量急剧下降, 因此必须使用视频抗分组丢失方法来保证图像质量, 常用的算法有: 抗丢包的前向纠错编码^[1], 接收端的误码掩盖技术^[2,3], 用于实时视频传输的混合 ARQ (自动请求重传) 技术^[4], 提高传输可靠性的视频多描述编码^[5]等. 在视频编码标准 H. 264^[6]中, 也使用了参数集、灵活块重排^[7]和 SP 帧 (切换预测帧) 等抗误码工具来保证视频通信质量.

尽管有多种视频抗分组丢失算法, 但对特定视频数据进行帧内编码的帧内刷新算法, 仍然是一种非常简单的、有效的视频抗分组丢失技术. 帧内帧 (I 帧) 只使用当前帧的图像数据信息进行编码, 而不参考相邻帧信

息, 在丢包环境下能够有效的防止误码扩散, 从而保证恢复视频主观质量^[8]. 然而, 编码过多的 I 帧会导致编码比特数大量增加, 不利于带宽受限情况下的实时视频应用. 因此, 对一帧图像中部分数据进行帧内刷新的方法得到广泛使用. 其中随机刷新、定期刷新等算法并不能有效提高丢包环境下的恢复图像质量; 基于感兴趣区域的刷新算法^[9]计算过于复杂, 不利于实现实时视频通信; 率失真优化的帧内刷新算法^[10], 仅对分数像素精度的编码失真进行估计, 而没有充分考虑分组丢失对编码模式选择的影响.

本文基于 H. 264 的算法特点, 提出了一种新的自适应帧内刷新算法. 首先建立一个帧内刷新尺度矩阵, 该矩阵中每个元素对应图像中的每个宏块, 元素值为宏块用于后继帧参考区域面积的归一化值; 然后结合帧内刷新尺度矩阵和率失真模型, 在丢包网络环境下提出基于

收稿日期: 2007-07-31; 修回日期: 2008-07-04

基金项目: 陕西省自然科学基金基础研究计划 (No. 2004F07); ISN 国家重点实验室资助课题

帧内刷新尺度的端到端率失真模型 (ME2ERD); 接下来, 利用 H. 264 中补充增强信息 (SEI) 域实现无需专门反馈信道的信息交互, 结合 ME2ERD 模型自适应进行帧内刷新. 实验结果表明, 使用本文提出的帧内刷新算法, 能够大大提高视频抗分组丢失能力, 且算法运算复杂度低, 具有实用价值.

2 基于 H. 264 的帧内刷新尺度矩阵

与目前使用广泛的 H. 263 + 标准相比, H. 264 视频编码算法有许多新颖之处^[6]. 由于 H. 264 采用尺寸可变块编码: $16 \times 16, 16 \times 8, 8 \times 16, 8 \times 8, 4 \times 4, 8 \times 4, 4 \times 8$, 大大提高了宏块运动估值的准确性, 从而提高了编码效率. 在基于块编码的 H. 264 中, 如果某一位置宏块被后继帧运动补偿时参考的像素总数越多, 则该宏块对后继帧图像质量影响越大, 在丢包环境下, 该宏块的丢失将造成严重的误码扩散. 因此, 将这些对后继帧影响大的宏块优先进行帧内模式编码, 将有效提高丢包环境下的恢复图像质量.

基于上述讨论, 可以首先计算各个宏块用于后继帧运动补偿的参考区域面积 (面积值表示该区域包含的像素总数), 并作归一化处理, 以此来建立帧内刷新尺度矩阵, 作为帧内刷新的判决依据. 具体步骤如下:

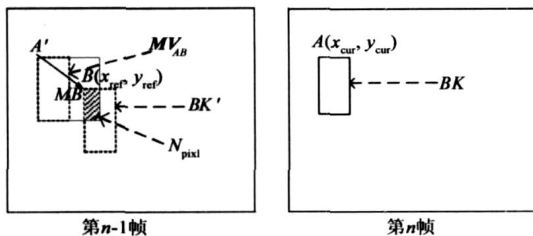


图1 参考区域面积计算示例

步骤1 计算参考宏块坐标

H. 264 预测帧中编码块的大小和尺寸可变, 首先要根据当前帧编码宏块的分割类型确定可变块的宽高, 然后利用该编码块的坐标值和运动矢量得到其参考块在参考帧中的坐标值, 这样就可以计算该编码块的参考块在参考帧各个宏块中的面积分布. 图1以第 n 帧编码块 $BK(8 \times 16)$ 分割模式) 为例计算其在参考帧宏块中的面积. 图中 BK 的左上角像素 A 坐标为 (x_{cur}, y_{cur}) , 在第 $n-1$ 帧的对应位置 A' 坐标 (x_{ref}, y_{ref}) , 通过块 BK 的运动矢量 $MV_{AB}(mv_x, mv_y)$ 可以计算得到第 $n-1$ 帧的参考块 BK 左上角像素 B 坐标 (x_{ref}, y_{ref}) . 如果像素 B 属于第 $n-1$ 帧的宏块 MB , 则宏块 MB 的坐标 (x, y) 为:

$$\begin{cases} x = x_{ref} / 16, & x_{ref} = x_{cur} + mv_x \\ y = y_{ref} / 16, & y_{ref} = y_{cur} + mv_y \end{cases} \quad (1)$$

式中 $/$ 表示整除, H. 264 标准中宏块大小为 16×16 .

步骤2 计算参考块在各宏块中的面积分布

$$\begin{cases} d_x = (x + 1) \times 16 - x_{ref} \\ d_y = (y + 1) \times 16 - y_{ref} \end{cases} \quad (2)$$

式中 d_x 和 d_y 分别表示 $B(x_{ref}, y_{ref})$ 距宏块 MB 右侧和下部边缘的距离. 根据 BK 的编码分割类型, 得到可变块 BK 的宽 (w_{BK}) 和高 (h_{BK}), 由此计算图1中 BK 的参考块 BK 落在宏块 MB 中的像素数为 N_{pixl} (图中阴影部分):

$$N_{pixl} = \min(d_x, w_{BK}) \times \min(d_y, h_{BK}) \quad (3)$$

式中函数 $\min()$ 计算参数间的最小值. 同理, 可以计算块 BK 的剩余部分 (除 N_{pixl} 以外的部分) 落在 MB 相邻宏块中的像素面积.

根据式(3)的计算结果, 第 $n-1$ 帧宏块 MB 被第 n 帧中多个宏块做为运动参考的像素区域总面积 $S_{x,y}$ 是:

$$S_{x,y} = \sum_i^M N_{pixl}^i \quad (4)$$

式(4)表示第 n 帧中有 M 个块参考 MB 的部分或全部像素进行运动补偿, N_{pixl}^i 表示第 i 个块在运动补偿时使用第 $n-1$ 帧宏块 MB 的像素数. 通过式(4)可以计算 $n-1$ 帧中各个宏块用于第 n 帧宏块运动补偿参考的像素面积.

步骤3 计算帧内刷新尺度矩阵

由于视频序列连续多个相邻帧之间有较强的时间相关性, 因此可以通过连续跟踪前驱帧来修正步骤2中的像素面积. 设第 n 帧 (当前编码帧) 的前 j 帧为第 $n-j$ 帧, 则该宏块被第 $(n-j+1)$ 帧中编码块用做参考的区域总面积记做 $S_{x,y}^{n-j}$. 将当前编码帧的前驱 w 帧中 $MB_{x,y}$ (坐标为 (x, y) 的宏块) 的参考区域面积加权求和:

$$\bar{S}_{x,y}^n = \sum_j^w \omega_j S_{x,y}^{n-j}, \quad 0 \leq \omega_j \leq 1 \text{ 且 } \sum_j^w \omega_j = 1 \quad (5)$$

式中 ω_j 随 j 增加而减小, $\bar{S}_{x,y}^n$ 值越大说明 $MB_{x,y}$ 对后继帧编码块的运动补偿影响越大, 即在接收端 $MB_{x,y}$ 解码出错对解码恢复图像质量影响较大, 因此可以利用 $\bar{S}_{x,y}^n$ 值来调整宏块在编码时的帧内/帧间模式, 有效提高丢包环境下的恢复视频质量. 式(6)将 $\bar{S}_{x,y}^n$ 进行归一化处理 (从以像素为单位转换到以宏块为单位), 得到第 n 帧中宏块 $MB_{x,y}$ 的帧内刷新尺度值:

$$\alpha_{x,y}^n = \bar{S}_{x,y}^n / (16 \times 16) \quad (6)$$

利用式(6)对一帧图像每个宏块都计算 $\alpha_{x,y}^n$, 得到第 n 帧的帧内刷新尺度矩阵, 该矩阵每个元素对应着图像中的每个宏块, 元素值为各宏块用于后继帧参考区域面积的归一化值, 用于第 $n+1$ 帧宏块的帧内/帧间编码模式判决. 帧内刷新尺度矩阵给出每个宏块被后继帧宏块参考的重要程度, 在编码模式判决时, 对 $\alpha_{x,y}^n$ 大的宏块提高其帧内编码的可能性, 从而有效减少误码扩散造成的恢复图像质量下降.

3 自适应帧内刷新算法

3.1 基于帧内刷新尺度的端到端率失真模型

在 H. 264 预测帧中,使用率失真(RD)优化方法来选择宏块的编码模式,即将预测误差和信息编码码率结合起来选择编码模式^[11]:

$$J_{\text{mode}}(MB_{x,y}, I_{x,y}) = D_{\text{REC}}(MB_{x,y}, I_{x,y}) + \lambda * R_{\text{REC}}(MB_{x,y}, I_{x,y}) \quad (7)$$

式中 $I_{x,y}$ 表示 $MB_{x,y}$ 的可选预测模式,包括帧内编码模式和帧间编码模式, Lagrangian 参数 λ 的值为 $0.85 \cdot 2^{(QP-12)/3}$ (QP 是量化因子)^[11], R_{REC} 和 D_{REC} 分别表示对应预测模式下的编码比特数和由量化造成的解码恢复图像失真. 在 IP 网络下,考虑分组丢失的影响^[12],利用上一节给出的帧内刷新尺度矩阵,本文提出基于帧内刷新尺度的端到端率失真模型(ME2ERD):

$$J_{\text{mode}}(MB_{x,y}, I_{x,y}) = D_{\text{REC}}(MB_{x,y}, I_{x,y}) + \beta(\alpha_{x,y}^n, I_{x,y}, P_{\text{loss}}) * D_{\text{REC}}(MB_{x,y}, I_{x,y}) + \lambda * R_{\text{REC}}(MB_{x,y}, I_{x,y}) \quad (8)$$

式中 D_{REC} 表示解码端出现丢包时由于误码扩散造成的图像失真,参数 $\beta(\alpha_{x,y}^n, I_{x,y}, P_{\text{loss}})$ 取值与宏块 $MB_{x,y}$ 的帧内刷新尺度 $\alpha_{x,y}^n$ 、帧内/帧间编码模式 $I_{x,y}$ 、丢包影响系数 P_{loss} 相关,其中丢包影响系数是网络丢包对解码端的影响程度,我们在实验中使用网络丢包率表示 P_{loss} . $\beta(\alpha_{x,y}^n, I_{x,y}, P_{\text{loss}})$ 的具体计算方法:

$$\beta(\alpha_{x,y}^n, I_{x,y}, P_{\text{loss}}) = \begin{cases} 1, & (I_{x,y} \text{ intra}) \\ 1/(1 - P_{\text{loss}})^2, & (\alpha_{x,y}^n \leq 1, I_{x,y} \text{ inter}) \\ \sqrt{\alpha_{x,y}^n}/(1 - P_{\text{loss}})^2, & (\alpha_{x,y}^n > 1, I_{x,y} \text{ inter}) \end{cases} \quad (9)$$

式(9)表明在丢包环境下,编码帧间帧时对那些易造成误码扩散的宏块($\alpha_{x,y}^n$ 值较大),其 β 值也较大,将 β 值带入式(8)就可以提高这些宏块帧内编码的可能性.

ME2ERD 模型充分考虑到由于网络丢包造成误码扩散所带来的恢复图像失真,在进行帧内/帧间编码模式判决时可以提高 $\alpha_{x,y}^n$ 值较大宏块的帧内编码可能性,有效地降低误码扩散造成的恢复图像质量失真. 另外,由式(9)可知,在实际视频通信系统中,需要实时反馈丢包影响系数,下面利用 H. 264 中 SEI 域给出无需专门反馈信道的信息交互方式,实现自适应帧内刷新算法.

3.2 基于 SEI 域的自适应帧内刷新

H. 264 中定义了补充增强信息(SEI)^[6],它的数据表示区域与视频编码数据独立. 在 SEI 域内可以记入有用的自定义信息,这些自定义信息称之为 SEI 消息. SEI 域消息包括固定格式的 SEI 头信息和变长的 SEI 有效载荷,SEI 头信息又由 SEI 类型和 SEI 载荷长度组成. 本文规定 SEI 类型取值为 19,表示 SEI 有效载荷存放接收

端统计的网络丢包率.

H. 264 中定义的 SEI 域可根据需求存放足够多的用户自定义信息. 对于不支持解析这些用户自定义信息的 H. 264 解码器,会自动丢弃 SEI 域中的数据. 因此,在 SEI 域内记入有用的自定义信息不会影响基于 H. 264 视频通信系统的兼容性. 在双向视频通信时,通信系统的各个终端都既有编码器,又有解码器. 基于这一特性,接收端可以将丢包信息记入本端编码码流的 SEI 域中,发往对端. 这样既不占用专门的反馈信道,又能加快信息反馈的速度. 基于 SEI 域自适应帧内刷新算法的具体步骤:

(1) 接收端解码器将分组丢失信息发送给本端编码器. 由于 H. 264 在 IP 网络上传输使用 RTP 协议封装,因此可以利用 RTP 包的接收和 RTP 头信息解析计算网络丢包率,作为解码端的丢包影响系数.

(2) 将(1)中信息在接收端编码器编码为 SEI 消息,然后进行网络适配层(NAL)的封装,通过 IP 网络发往发送端解码器.

(3) 发送端解码器解码 SEI 消息,将 SEI 消息内容通知给本端编码器或者丢弃.

(4) 如果发送端编码器能够解析 SEI 消息,可以由 SEI 消息解析网络丢包率,利用式(8)给出的 ME2ERD 模型,在发送端编码当前帧时,做为宏块帧内/帧间编码模式判决的参数. 这样可以根据丢包率的变化自适应进行帧内刷新,来提高丢包网络环境下恢复图像的质量,保证实时视频通信的服务质量.

4 实验结果

本文使用两个 CIF 格式(352 × 288)标准图像序列:“Stefan”(运动较为剧烈)和“Coastguard”(运动较为平缓),各取 300 帧,重复 20 次进行实验研究. 实验中设定目标帧率为 25frames/s,编码速率为 512kbps,用 PSNR 作为衡量恢复图像客观质量的测度. 另外,式(5)中取当前帧的前两帧数据进行加权求和, ω_j 值分别为 0.7, 0.3 ($j=1, 2$). 在不同丢包率环境下,图 2 给出下面四种方法性能比较的实验结果:

- * 没有使用任何抗分组丢失方法(NoER);
- * 每帧随机选定 10 个宏块强制进行帧内编码,即使用随机帧内刷新算法(RIR);
- * 使用文献[10]提出的基于率失真优化的帧内刷新算法(RDerr);
- * 利用 ME2ERD 模型和 SEI 域,使用本文提出自适应帧内刷新算法(MAIR).

由图 2 可以看出,无论运动剧烈的序列,还是运动较为平缓的序列,与 NoER 方法相比,使用 RIR 算法解码恢复图像的 PSNR 提高小于 1dB;而使用 MAIR 算法,

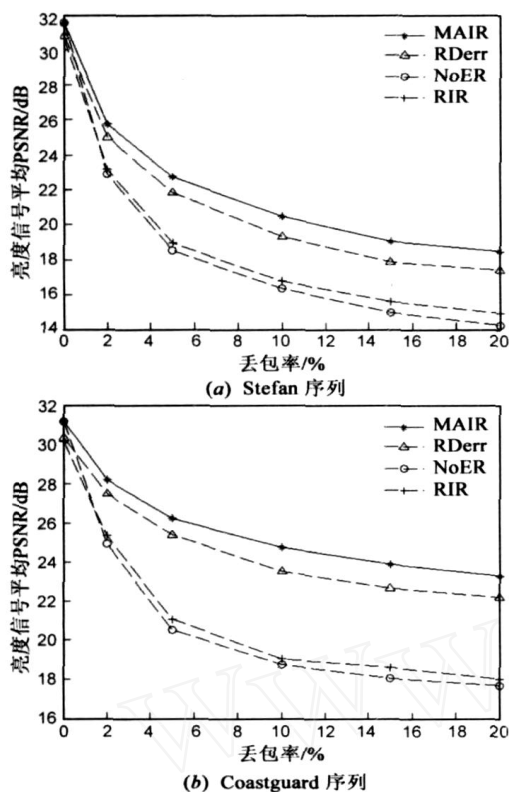


图2 四种方法在不同丢包率下的性能比较

在高丢包率下,恢复图像平均 PSNR 比 RIR 算法能够提高 5dB 以上,且与 RDerr 算法相比也可以提高近 1dB. 另外,使用本文提出的 MAIR 算法,可以显著提高恢复图像的主观质量.

5 结论

本文根据每个宏块被后继帧参考的重要程度不同建立帧内刷新尺度矩阵,并在率失真模型的基础上提出 ME2ERD 模型,接下来利用 H.264 中 SEI 域给出自适应帧内刷新算法. 实验结果表明,在高丢包率环境下,使用本文提出的算法可以大大提高恢复图像的主客观质量. 且该算法没有过多增加运算复杂度,保证实时视频通信的服务质量.

参考文献:

- [1] Qian Xu, Vladimir Stanković, Zixiang Xiong. Distributed joint source-channel coding of video using raptor codes [J]. IEEE Journal on Selected Areas in Communications, 2007, 25 (4): 851 - 861.
- [2] Min-Cheol Hwang, Jun-Hyung Kim, Hae-Yong Yang, et al. Frame error concealment technique using adaptive inter-mode estimation for H.264/AVC [J]. IEEE Transactions on Consumer Electronics, 2008, 54 (1): 163 - 170.
- [3] 周智恒, 谢胜利. 新的自适应 MRF-MAP 视频序列误差掩盖 [J]. 电子学报, 2006, 34 (4): 628 - 633.

Zhou Zhi-heng, XIE Sheng-li. New adaptive MRF-MAP error concealment of video sequences [J]. Acta Electronica Sinica, 2006, 34 (4): 628 - 633. (in Chinese)

- [4] Ivan V. Bajic. Efficient cross-layer error control for wireless video multicast [J]. IEEE Transactions on Broadcasting, 2007, 53 (1): 276 - 285.
- [5] Ottavio Campana, Roberto Contiero, Gian Antonio Mian. An H.264/AVC video coder based on a multiple description scalar quantizer [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2008, 18 (2): 268 - 272.
- [6] ITU-T Recommendation H.264, Advanced video coding for generic audiovisual services [S].
- [7] 姜东, 李波, 李炜等. 基于 ZIG-ZAG 交织的 H.264/AVC 容错编码算法 [J]. 电子学报, 2006, 34 (11): 1941 - 1945.
Jiang Dong, LI Bo, LI Wei, et al. An error resilient H.264/AVC video coding algorithm using ZIG-ZAG interleaved FMO [J]. Acta Electronica Sinica, 2006, 34 (11): 1941 - 1945. (in Chinese)
- [8] 李世平, 蒋刚毅, 郁梅. 快速帧内预测模式选择新方法 [J]. 电子学报, 2006, 34 (1): 141 - 146.
Li Shi-ping, Jiang Gang-yi, Yu Mei. New fast mode selection for intra prediction [J]. Acta Electronica Sinica, 2006, 34 (1): 141 - 146. (in Chinese)
- [9] Quqing Chen, Zhibo Chen, Xiaodong Gu, et al. Attention-based adaptive intra refresh for error-prone video transmission [J]. IEEE Communications Magazine, 2007, 45 (1): 52 - 60.
- [10] Ma Zhonghua, Yu Songyu, Zeng Zhaoping. Error resilient intra refresh scheme for H.26L stream [J]. Electronics letters, 2002, 38 (19): 1153 - 1154.
- [11] Thomas Wiegand, Heiko Schwarz, Anthony Joch, et al. Rate-constrained coder control and comparison of video coding standards [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2003, 13 (7): 688 - 703.
- [12] G Cote, F Kossentini. Optimal intra coding of blocks for robust video communication over the internet [J]. Signal Processing: Image Communication, 1999, 15 (1 - 2): 25 - 34.

作者简介:



宋彬男, 1973 年 11 月出生于河南郑州. 2002 年在西安电子科技大学获工学博士学位, 现为西安电子科技大学副教授, 主要从事图像压缩和视频通信等领域的研究, 发表论文 20 多篇.
E-mail: bsong@mail.xidian.edu.cn

蒋小兵男, 1982 年 9 月出生于河南郑州. 2005 年毕业于西安电子科技大学通信工程学院, 现在西安电子科技大学攻读硕士学位, 主要研究方向为视频编码与通信.