

保证 H. 264 视频通信质量的实时传输协议载荷

宋彬, 秦浩, 郭春芳

(西安电子科技大学综合业务网国家重点实验室, 710071, 西安)

摘要: 在详细研究 H. 264 标准的网络适配层(NAL)和实时传输协议(RTP)的基础上,提出一种可以保证 H. 264 视频通信质量的实时传输协议(ERRTP)载荷结构,载荷结构中使用了 NAL 头字段和前向纠错(FEC)保护字段,在增加开销很小的情况下,支持对 H. 264 视频码流进行 FEC 不等保护,同时可以保证时延在实时视频通信允许的范围内.性能分析和实验结果表明,与 RTP 协议相比,使用 ERRTP 载荷的传输效率基本没有变化,且 ERRTP 支持 FEC 不等保护,与单一 FEC 保护相比,在高丢包率下恢复视频质量的峰值信噪比可以提高 2~9 dB,满足分组丢失环境下实时视频通信的服务质量要求.

关键词: H. 264 标准;实时传输协议;服务质量;前向纠错

中图分类号: TN919.8 **文献标识码:** A **文章编号:** 0253-987X(2007)12-1455-05

Real-Time Transport Protocol Payload for Guaranteeing the QoS of H. 264 Video Delivery

Song Bin, Qin Hao, Guo Chunfang

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

Abstract: On the basis of detailedly researching on real-time transport protocol (RTP) and network adaptation layer (NAL) of video coding standard H. 264, a payload format of real-time transport protocol for guaranteeing QoS of H. 264 video delivery over IP network is proposed, in which the NAL header and the forward erasure correction (FEC) field is employed. With less extra overhead, the proposed payload can support FEC unequal-protection for H. 264 video stream and ensures the delay to be in allowable range for real-time video delivery. Performances analysis and experimental results show that compared with RTP, the transmission efficiency using ERRTP is basically not changing and ERRTP supports FEC unequal-protection, and compared with single FEC protection, the PSNR of recovery video quality can be increased by 2~9 dB, which satisfies QoS demands of real-time video communication under packet drop environment.

Keywords: H. 264 standard; real-time transport protocol; quality of service; forward erasure correction

随着第三代移动通信系统的出现和 IP 网络的迅速发展,视频通信已经成为通信的主要业务之一.继制定了 H. 261、H. 263、H. 263⁺ 等视频压缩标准后,ITU-T 和 ISO/IEC 又联合制定了新的视频压缩编码标准 H. 264^[1],与其他视频编码标准相比,H. 264 标准能更加有效地提高视频编码效率和对网络

的适配性.

然而,IP 网络提供“尽力”传输服务,但其自身并不能保证视频信号传输的服务质量(QoS),特别是对经过高效压缩编码的 H. 264 码流,问题更为突出^[2-3].具体表现在 3 个方面:分组丢失、时延和时延抖动.其中,分组丢失对恢复视频的质量影响最大,

由于 H. 264 压缩编码算法使用运动估值和运动补偿技术,一旦有分组丢失存在,不仅影响当前解码图像,而且会影响后续解码图像,即误码扩散,而时延和时延抖动对实时视频通信的恢复图像质量也会造成影响。

实时传输协议(RTP)^[4]是在 IP 网络上进行视频通信时采用的主要协议,它提供单播或者多播网络上端到端的实时数据传输,RTP 典型的运行于用户数据报协议(UDP)之上.文献[5-7]分别给出了在 RTP 上传输 H. 263、MPEG-4 和 H. 264 等不同媒体格式的方法,但是文献中并未涉及对视频数据进行 FEC 保护的方法,不能保证在丢包环境下视频传输的 QoS.

本文根据 RTP 载荷结构^[4]和 H. 264 网络适配层(NAL)的特点,利用 NAL 单元(NALU)头结构,重新构造 RTP 载荷,提出新的具有抗丢包能力的 RTP 载荷结构(ERRTP).然后,利用 ERRTP 载荷结构,对 H. 264 码流使用 FEC 不等保护,能够保证时延在实时视频通信允许范围内,大大提高 H. 264 视频通信的 QoS.

1 RTP 结构和 H. 264 的 NALU

1.1 RTP 结构

RTP 分组由头字段和载荷两部分组成:头字段包括版本号、填充标识、扩展标识位、载荷类型以及序列号、时间戳等信息,使用基本选项时头字段占 12 个字节;载荷中存放有效数据信息,如视频码流等.用户可以自定义载荷的格式,从而传输不同格式的媒体流数据.

1.2 NALU 结构

NAL 是为了提供“网络友好性”而设计的,它有利于 H. 264 中视频编码层(VCL)数据到传输层的适配.NAL 中用 NALU 存放数据,它由 NALU 头和 NALU 载荷两部分组成,具体格式如图 1 所示.



图1 NALU 结构

图 1 中禁用字段占用 1 bit,在 H. 264 编码器中设为 0,如果网络节点识别该 NALU 中存在比特错误,则将其置 1. NALU 参考标识占用 2 bit,其值越大,表示当前 NALU 越重要. NALU 类型占用 5

bit,共 32 种^[1],其中取值 1~12 表示不同类型,取值 13~23 表示保留,其他未定义.

常规方式下,H. 264 视频流以简单的方式打包成 RTP 载荷^[7],共有 3 种不同的打包模式,分别为单个 NALU 分组模式、聚合模式、分片模式.其中,分组模式将一个 NALU 装入一个 RTP 载荷,聚合模式将多个 NALU 装入一个 RTP 载荷,分片模式将一个 NALU 装入多个 RTP 载荷.

上述打包方法没有考虑网络中可能存在丢包,尤其是 H. 264 编码具有非常高的时间相关性,丢包可能导致误码扩散,大大降低了视频通信的 QoS.为了解决该问题,本文提出一种新的具有抗丢包能力的载荷格式.

2 新颖的 ERRTP 载荷结构

为了描述简单,首先说明 FEC 算法的流程:将待保护的数据分割成等长的数据单元,称为数据符号,若干个数据符号组成一个数据块作为 FEC 算法的输入,最终生成若干个校验符号.

本文提出的 ERRTP 载荷结构,如图 2 所示.

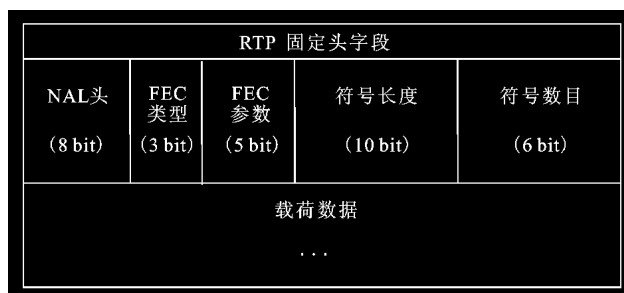


图2 ERRTP 载荷结构

载荷结构中增加的具体内容如下.

(1)NALU 头字段. 规定在一个 ERRTP 载荷中只能存放相同类型的 NALU(即图 1 中 NALU 类型值相同),如果一个 ERRTP 载荷中包含多个 NALU,则这些 NALU 的头字段必定相同,为节省开销,只保留一个 NALU 头存放在 ERRTP 载荷的起始位置,其他 NALU 的头字段则全部删除. 这样设计有两个优点:①额外开销少,尤其是一个 ERRTP 中存放多个 NALU 时,明显节省传输比特数;②不用对分组内的数据解码就可以判别该分组中视频数据的重要性.

(2)FEC 码的类型. 指出此载荷使用的纠删算法,接收端据此决定如何解码收到的 RTP 载荷,例如用 2 标识 Tornado 码^[8],用 3 标识 RS 码,若该字段为 0 则表示载荷不使用 FEC 保护等,这样可标

识 7 种不同的 FEC 码。

(3) FEC 参数域. 指出 FEC 码的生成规则和保护能力. 以 Tornado 码为例, FEC 参数为 2 表示 Tornado(24, 20)(即其中数据节点 20 个, 校验节点 4 个), 参数为 3 表示 Tornado(30, 20)等等. 这个字段表示的内容结合 FEC 类型来对 H. 264 的 NALU 数据进行前向保护。

(4) 符号长度. 在使用 FEC 算法时, 每个数据符号的长度, 单位为字(每字等于 4 B)。

(5) 载包含的符号数目. S 个同一类型的 NALU 经过 FEC 校验后, 在一个 ERRTP 载荷中存放的符号个数。

在 ERRTP 载荷中加入 FEC 保护字段后, 网络节点或接收端可以根据这些字段给出的 FEC 码类型和参数对接收到的数据分组进行解码, 并恢复丢失的数据包, 提高分组丢失环境下恢复视频的质量。

3 用 ERRTP 载荷保证实时视频通信的 QoS

3.1 使用 FEC 码保护 H. 264 码流

利用 ERRTP 载荷结构, 对 H. 264 码流进行 FEC 分组保护, 基本原理是: 首先将 H. 264 中同一类型但不等长的 S 个 NALU 分为一组; 再将 S 个 NALU 载荷等分为 M 个数据符号, 其中每个数据符号的长度为 K 个字节; 然后对 M 个数据符号使用 FEC 码编码生成 N 个校验符号; 最后将 $M+N$ 个符号在 P 个 ERRTP 载荷中进行传输, 而这些相同类型 NALU 的头信息存放在 ERRTP 载荷头位置, 如图 3 所示. 接收端在收到第 $P+1$ 分组后开始检测前面 P 个分组中是否有数据包丢失, 如果有就

进行 FEC 解码, 恢复丢失的数据符号。

结合图 3 给出的具体方法, ERRTP 载荷结构中主要字段设置如下: FEC 类型字段取值 2, FEC 参数字段取值 2, 合起来表明使用的 FEC 码是 Tornado(24, 20)码; 符号长度字段取值为 $K/4$, 符号数目字段对应取值为 $(M+N)/P$ 。

H. 264 使用 FEC 码保护时, 需要考虑数据符号的大小、FEC 输入数据块中数据符号的个数以及校验数据的冗余度. 其中, 数据符号长度要小于网络的最大传输单元(MTU), 使得数据符号不会被网络传输协议再次分割, 保证 FEC 码的有效保护作用. 另外, 在视频通信延时允许的范围内, 可增加数据符号数 M , 从而增强 FEC 码的抗分组丢失能力. 然而, 当增加 FEC 码校验符号数 N 时, 其抗丢包的性能也会增强, 当然冗余度也随之增加。

3.2 使用 FEC 不等保护

H. 264 码流基于 NALU 进行传输或存储, 由图 1 可知, NALU 由 NAL 头信息和 NAL 载荷组成, 且不同 NALU 参考标识和 NALU 类型对解码恢复图像的影响不同. 因此, 在对 H. 264 的码流进行分组保护时, 可以根据 NALU 头字段中参考标识字段或 NALU 类型字段的取值将 H. 264 的数据分为两类, 一类为相对重要的图像数据(例如参考标识等于 1), 另一类为次要的图像数据(例如参考标识等于 0). 然后, 对重要的图像数据使用冗余度较大、抗丢包能力强的 FEC1 码进行保护, 而次要的图像数据可以使用冗余度较小、抗丢包能力较弱的 FEC2 码进行保护. FEC1 码和 FEC2 码表示任意两种子类型, 这两种子类型可以属于同一 FEC 类型, 也可以属于不同的 FEC 类型。

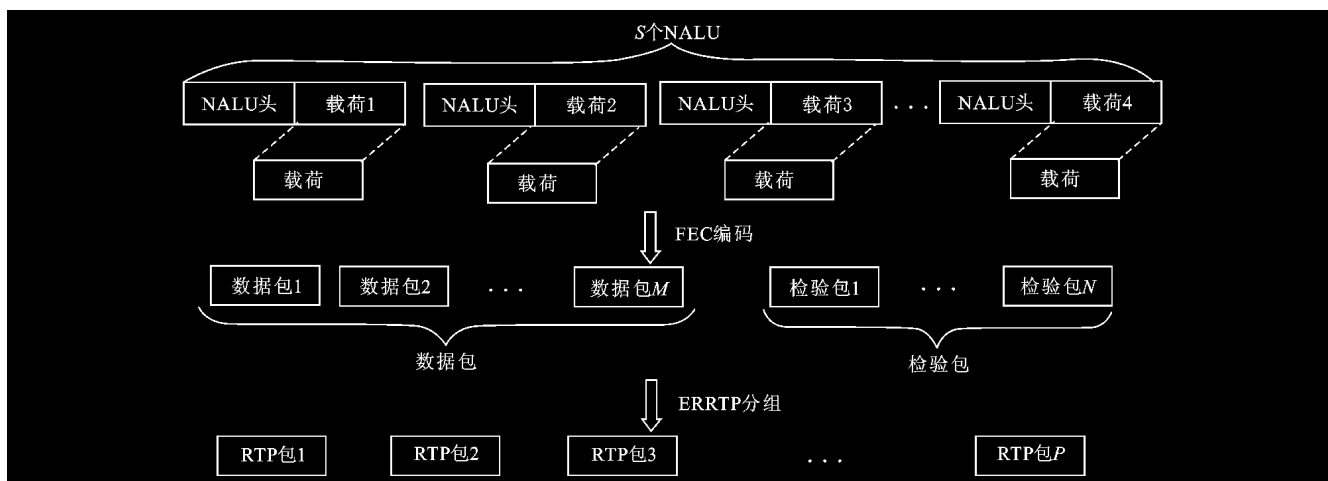


图 3 使用 FEC 码保护 H. 264 码流的方法

在 ERRTP 载荷中,由于 NALU 的头信息存放在载荷头位置,因此可以在解析参考标识和 NALU 类型的定义内容后,非常方便地对 H. 264 中具有不同重要性的 NALU 进行不等保护,而不用对码流进行解码。

3.3 FEC 保护的实时性

使用 FEC 码保护码流会引入时延,时延的大小与图像数据分组的大小相关。在图 3 中,将 S 个同一类型 NALU 分为一组,其中一个 NALU 包含一个分片(Slice)的码流数据。若一帧图像只划分为一个 Slice,则一帧图像在编解码端都会有 $S/2$ 帧的 FEC 处理时延。设 T_f 为使用 FEC 保护后引入的时延, T_c 和 T_t 分别为 H. 264 编解码器处理时延和网络传输时延,则在 IP 网络进行实时视频通信时,一帧图像的时延为

$$T_T = T_f + T_c + T_t \quad (1)$$

本文假定 T_c 能够满足实时性要求,即

$$T_c \leq 2T_{th}, \quad T_{th} = 1/F_t \quad (2)$$

式中: F_t 为解码目标帧率(单位 帧/s),则式(1)可改为

$$T_T \leq 2T_{th}(S/2) + 2T_{th} + T_t = (S+2)T_{th} + T_t \quad (3)$$

由式(3)可知,与 RTP 协议相比,使用 ERRTP 协议后时延 T_T 增加了 ST_{th} ,为了减小实时视频通信系统的时延,则 S 取值不应太大。从图 3 可以看出, S 个 NALU 载荷的总字节数等于数据符号的个数 M 与符号长度 K 的乘积。在使用 FEC 保护数据符号时,符号长度 K 一般取定值^[8],因此 S 的值正比于 M ,也就是说为了保证视频通信的实时性, M 的取值不应太大。

在 M 受限的情况下,本文使用文献[9]提出的高效 Tornado 码保护算法。该算法避免了采用多级偶图的编码方式,而是只设计使用一级偶图的编码方式,并且去掉了最后一级的传统编码方式,算法非常灵活,可针对任意数目的数据符号进行纠删保护,生成任意数目的校验符号。高效 Tornado 码较之传统 Tornado 编解码时间减少了 50%。当使用高效 Tornado 算法时,恢复图像的 PSNR 可提高 2 dB 以上。因此,高效 Tornado 算法可以在改善抗丢包能力的同时尽量减少 FEC 引入的时延,非常适于在 ERRTP 载荷结构中使用。

4 性能分析和实验结果

下面给出在仿真的 IP 网络环境下,使用 RTP

载荷结构和 ERRTP 载荷结构传输 H. 264 码流数据的性能比较。

首先分析使用 RTP 和 ERRTP 载荷的传输效率。以图 3 为例, S 个 NALU 在分组传输后,不同协议下每个分组的字节数为

$$\begin{aligned} L_{RTP} &= 12 + (S/P)(\bar{L}_{NALU} + 1) \\ L_{ERRTP} &= 12 + 4 + (S/P)\bar{L}_{NALU} \end{aligned} \quad (4)$$

式中: $\bar{L}_{NALU}$ 表示 NALU 载荷的平均大小(不包括头信息,见图 1)。由式(4)可知,与 RTP 协议相比,使用 ERRTP 后每个分组大小的变化是 $(4-S/P)$ 字节,即在传输效率上与 RTP 协议基本相同。

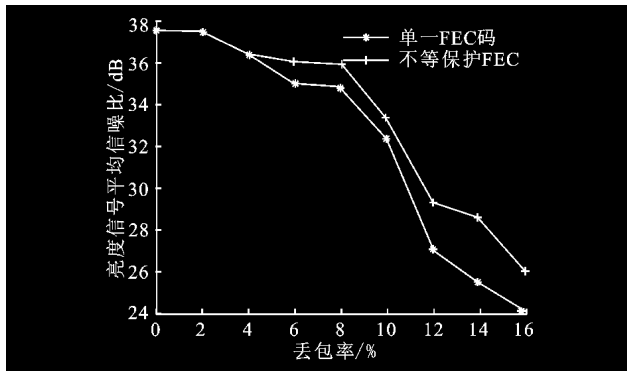
然后,分析使用 ERRTP 载荷对视频码流进行 FEC 保护后,引入时延对实时视频通信的影响。由式(3)可知,与 RTP 相比,使用具有抗误码能力的 ERRTP 载荷增加的时延为 ST_{th} 。例如,实验中 F_t 为 30 帧/s,设 T_t 取值为 60 ms,当 S 取值为 5 时,由式(3)计算 T_T 为 293.3 ms,基本可以满足 IP 网络上时延小于 300 ms 的实时视频业务。

接下来比较两种载荷结构的抗分组丢失能力: RTP 没有抗分组丢失能力;ERRTP 提供 FEC 类型字段,接收端或网络节点就可以结合 FEC 类型和 FEC 参数字段进行前向纠错。另外,利用 ERRTP 载荷结构,本文使用的 FEC 抗分组丢失方法能够在不对 H. 264 码流解码的情况下,根据 NALU 的重要程度进行不等保护。

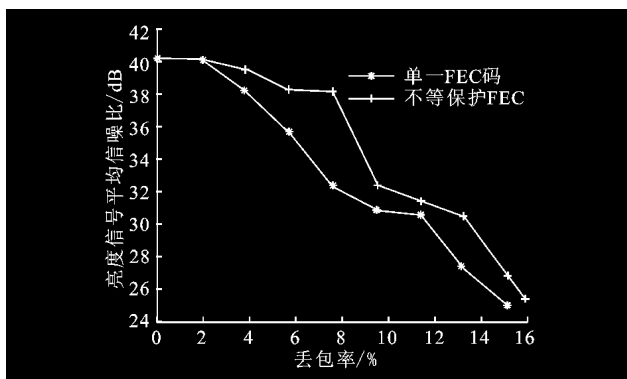
实验使用多个标准图像序列各 300 帧(重复 20 次)进行研究,它们均为 CIF 格式(352 像素×288 像素),Y:U:V 是 4:2:0。实验中设定信道速率为 384 kb/s,目标帧频为 30 帧/s,用亮度信号的平均 PSNR 作为衡量恢复图像客观质量的测度,FEC 保护算法使用 Tornado 码。图 4 给出了在 ERRTP 载荷支持下使用单一 FEC 码和不等保护 FEC 的实验结果,其中单一 FEC 码使用 TN(28,20)码,而 FEC 不等保护中,FEC1 码选用 TN(30,20)码、FEC2 码选用 TN(26,20)码。由图 4 可以看出,对于不同的序列,与单一 FEC 码保护相比,在较高的信道丢包率($>5\%$)环境下,使用高效 Tornado 码实施不等保护,可以显著提高恢复视频质量,尤其是对于 Foreman 序列,在信道丢包率达 10%的情况下,可将视频恢复质量提高 9 dB 以上。

5 结束语

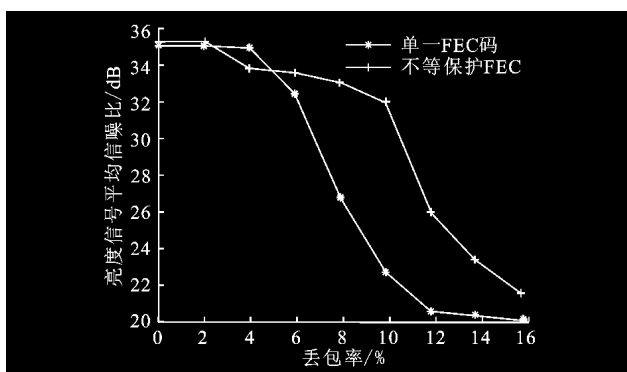
本文提出一种新的 ERRTP 载荷结构,它基本没有增加信道负担,且支持对 H. 264 视频码流进行



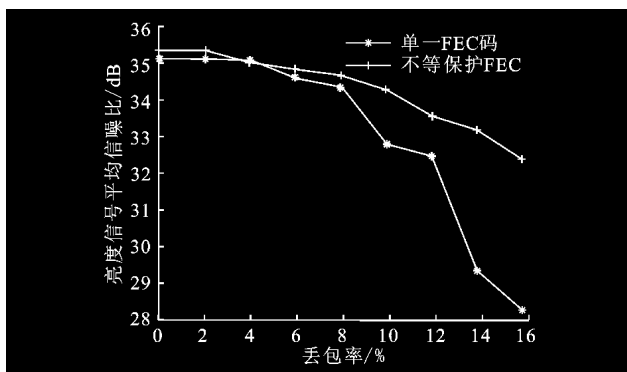
(a) Hall Monitor CIF 图像序列



(b) Mother CIF 图像序列



(c) Foreman CIF 图像序列



(d) Container CIF 图像序列

图4 利用 ERRTP 进行 FEC 抗分组丢失保护的实验结果

前向纠错保护,利用 ERRTP 载荷结构,使用高效 Tornado 码实施不等保护,能够保证纠错能力并满足实时视频通信时延的要求.实验结果表明,使用 ERRTP 和高效 Tornado 码结合的不等保护,可以在高丢包率环境下大大提高 H. 264 实时视频通信的 QoS.

参考文献:

- [1] Telecommunication Standardization Sector. ITU-T recommendation H. 264 Advanced video coding for generic audiovisual services [S]. Geneva, Switzerland: Telecommunication Standardization Sector, 2003.
 - [2] Wenger S. H. 264/AVC over IP [J]. IEEE Trans on Circuits and Systems for Video Tech, 2003, 13(7): 645-656.
 - [3] Mohamed S, Rubino G. A study of real-time packet video quality using random neural networks [J]. IEEE Trans on Circuits and Systems for Video Tech, 2002, 12(12): 1071-1083.
 - [4] Schulzrinne H, Casner S, Frederick R, et al. RFC 3550 RTP: a transport protocol for real-time applications [S/OL]. (2003-07-01)[2007-02-08]. <http://www.ietf.org/rfc/rfc3550.txt>.
 - [5] Ott J, Bormann C, Sullivan G, et al. RFC4629 RTP payload format for ITU-T Rec H. 263 video [S/OL]. (2007-01-05)[2007-02-08]. <http://www.ietf.org/rfc/rfc4629.txt>.
 - [6] Kikuchi Y, Nomura T, Fukunaga S, et al. RFC3016 RTP payload format for MPEG-4 audio/visual streams [S/OL]. (2000-11-01) [2007-02-08]. <http://www.ietf.org/rfc/rfc3016.txt>.
 - [7] Wenger S, Hannuksela M M, Stockhammer T, et al. RFC 3984. RTP payload format for H. 264 video [S/OL]. (2005-02-01)[2007-02-09]. <http://www.ietf.org/rfc/rfc3984.txt>.
 - [8] Luby M, Mitzenmacher M, Shokrollahi A, et al. Practical loss-resilient codes [C]// Proc of the 29th ACM Symposium on Theory of Computing. New York: ACM Press, 1997: 150-159.
 - [9] 周宁兆, 宋彬, 常义林. 基于 H. 264/AVC 的视频通信抗分组丢失方法研究 [J]. 计算机学报, 2006, 29(1): 1-7.
- Zhou Ningzhao, Song Bin, Chang Yilin. Error resilience techniques for video communication based on H. 264/AVC [J]. Chinese Journal of Computers, 2006, 29(1): 1-7.

(编辑 刘杨)