

基于 SCTP 协议的 PFGS 视频流传输方法

张建龙, 吴成柯, 杜建超, 肖 嵩

(西安电子科技大学 ISN 国家重点实验室, 西安 710071)

摘要:为了适应网络的时变性,提高传输的可靠性同时满足实时性,提出一种新的网络传输结构——RTP/SCTP/IP。系统采用基于 H.264 协议的精简 PFGS 编码结构,利用 SCTP 协议的多流特性,将 PFGS 流分成多个子流分别传输,给出了一种基于网络带宽预测的传输方法。仿真试验表明,该方法较传统网络结构 RTP/UDP/IP 而言,传输的可靠性和网络适应能力有了很大提高,图像平均 PSNR 比传统的 RTP/UDP/IP 网络结构下的 PSNR 高 0.5dB ~ 1.5dB。

关键词:流控制传输协议;多流;渐进的精细可分级

Transport Method of PFGS Video Stream Based on Protocol of SCTP

ZHANG Jian-long, WU Cheng-ke, DU Jian-chao, XIAO Song

(ISN National Key Lab, Xidian University, Xi'an 710071)

【Abstract】 A new transport framework of RTP/SCTP/IP is presented for adapting to the time-variable Internet and enhancing the reliability of transport. An improved simplified PFGS coding framework, using the standard of H.264 in base layer of PFGS, is adopted on source terminal. The stream of PFGS is separated for many sub-flows to transport by the characteristics of SCTP multi-streaming for improving the reliability and flexibility of transmission. A transport method that is a finite state machine(FSM) framework based on network bandwidth prediction over the Internet is presented. Simulation results show that this scheme can highly improve the quality of video and obtain a gain of about 0.5dB ~ 1.5 dB in PSNR compared with the traditional network framework. Ability of adapting to network and reliability of transmission are improved rapidly.

【Key words】 stream control transmission protocol(SCTP); multistreaming; progressive fine granular scalable (PFGS)

目前,多媒体通信已经成为网络通信研究的热点之一,视频通信更是其中的代表,其特点是数据量巨大、允许一定程度的失真、实时性要求高。为了适应时变的IP网络和用户的QoS差异,在视频编码端,MPEG2^[1]提出了可分级编码的思想。为了提高视频流传输的鲁棒性,适应网络的时变特性,FGS被MPEG-4^[2]标准采纳。为了在编码效率和错误恢复能力之间取得一个权衡,人们提出了渐进的精细可分级(progressive fine granular scalable, PFGS)^[3-4]的视频编码方法,该方法在保持了FGS所具有的网络带宽自适应和错误恢复能力的同时,有效地提高了编码效率。传统的视频传输结构采用RTP/UDP/IP,采用UDP协议主要是由于视频流数据量大、实时性要求高,然而PFGS流接收正确的前提是基本层必须正确接收到,UDP是无连接的协议,并不能保证这一点。SCTP^[5]协议是面向连接的传输层协议,主要面向多媒体应用,能够保证基本层的正确到达同时满足系统实时性要求。

为了提高视频传输的鲁棒性和网络传输效率,本文提出了一种基于 SCTP 协议的网络传输结构,在编码端采用基于 H.264 的精简 PFGS 结构,经过 RTP 打包之后通过 SCTP 传输,在 SCTP 连接中采用多流传输,包括基本层子流、多个增强层子流以及 RTCP 的反馈流。根据 RTCP 反馈的网络状况,采用基于网络带宽预测的有限状态机方法传输各个子流,仿真实验证明本文方法能有效改善视频流在网络传输上的鲁棒性,显著提高传输效率。

1 基于 SCTP 传输协议的系统传输框架

本文的系统结构如图 1 所示,其中,信源编码采用基于

H.264 标准的精简 PFGS 结构,传输层上采用面向连接的 SCTP 协议。SCTP 协议比 TCP 协议具有更高的吞吐率和带宽利用率,比 UDP 协议更可靠,更适合传输多媒体业务。另外根据 SCTP 协议的多流特性,将码流分成多个子流传输,互不影响,大大增加了传输的灵活性,更易于根据拥塞状况调整网络传输策略,试验证明该系统的整体性能较传统方法有了很大改善。

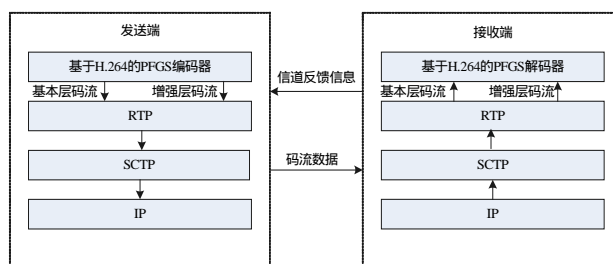


图1 系统结构

本文采用文献[3]中改进的 PFGS 编码器结构,如图 2 所示,该结构认为相邻两帧之间对应中间比特平面的相关性最强,当增强层作为参考帧时,只有中间的比特平面对预测精

基金项目:国家自然科学基金资助项目(60532060, 60372043); 国防重点实验室基金资助项目

作者简介:张建龙(1976 -),男,博士,主研方向:图像和视频压缩编码,Internet 和无线网络视频传输;吴成柯,教授、博士生导师;杜建超,博士、讲师;肖 嵩,博士、副教授

收稿日期:2006-10-29 **E-mail:** zjianl@yahoo.com.cn

度的影响最大,为了节约内存成本、简化编码器结构,只采用 2 个缓冲区存储中间比特平面。本文的基本层采用 H.264 编码标准。选取 H.264 标准主要原因是其具有较高的编码效率,且 H.264 的网络提取层(NAL)面向时变的包交换网络,非常适合在 IP 网络传输。在 NAL 层 DPM 模式下则将其分为 A 区(数据片和宏块的头部信息)、B 区(帧内编码信息和帧内变换系数)和 C 区(帧间编码信息和帧间变换系数)3 部分。

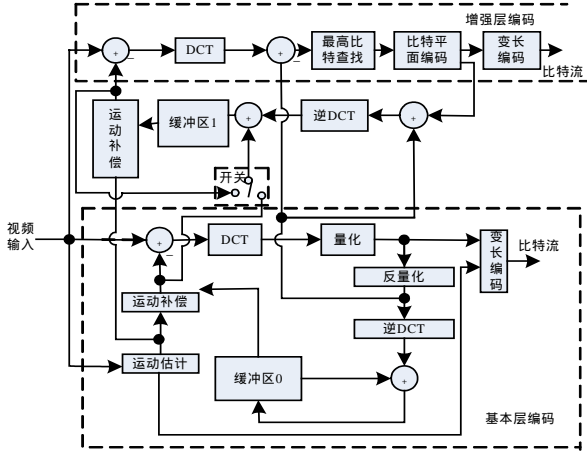


图 2 编码器结构

在网络层上本文选用面向连接的可靠传输协议 SCTP 协议,利用 SCTP 的多流特性,将基本层 A, B, C 区信息分成 3 个子流,将增强层分成相应的子流分别传输,采用不同的保护力度和传输方法,提高传输的可靠性和灵活性,其中, A 区信息始终以最强保护力度传输。

传输网络结构如图 3 所示。

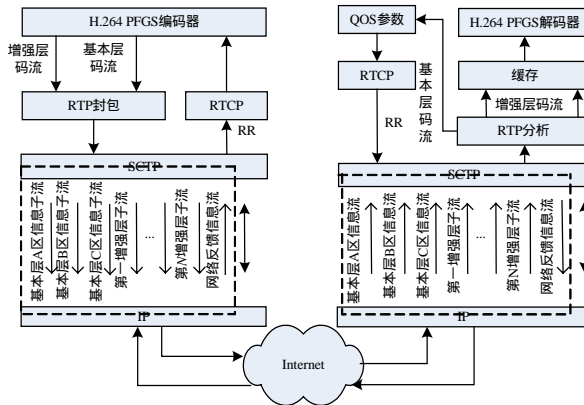


图 3 传输网络结构

图 3 中, SCTP 和 IP 层之间的虚线部分表示 SCTP 连接的 $N+3$ 个子流(stream),分别是:基本层 A 区信息子流,基本层 B 区信息子流,基本层 C 区信息子流,第一增强层子流,……,第 N 增强层子流和解码端反馈的网络状况信息流。由于 SCTP 协议是 TCP 友好协议,可以公平地竞争网络带宽,因此不会对网络整体性能造较大成影响。

对于多媒体业务的网络传输,由于解码端要求编码失真曲线尽量平滑,因此对于网络吞吐量应尽量平稳,这就对传输方法提出了较高要求。试验表明, PFGS 增强层除第一增强层外编码比特数都远远高于基本层编码比特数且成倍数关系,随着增强层数的增加倍数增大。根据表 1 中数据分析可以得出,增强层数越高,编码比特数越大,对网络吞吐量影响越大,但对图像质量的影响越小。而在传输策略上需要优先保护重要性高而编码比特数较少的子流,这恰恰符合 PFGS

子流特性。因此,本文采用 AIAD 的传输策略,同样可以达到传统的 AIMD 的效果,同时又保持了图像质量的稳定性。具体做法就是随着网络条件变好,各子流按重要性从高到低顺序开启,随着网络条件变差,各子流按重要性从低到高顺序关闭。

表 1 foreman 序列 PFGS 编码基本层与增强层特性比较

层	各层平均比特率/(bits/帧)	对应码率/(Kb·s ⁻¹)	PSNR/dB
基本层	8 722	256	32.35
增强层 # 1	4 460	131	32.85
增强层 # 2	26 218	768	35.95
增强层 # 3	53 015	1 553	40.51
增强层 # 4	82 326	2 412	45.72
增强层 # 5	124 410	3 645	50.70
增强层 # 6	189 443	5 550	59.04

2 基于网络带宽预测的有限状态机传输策略

在实际应用中,视频流媒体业务对数据的传输延时、包丢失、包失序等要求非常严格,因此,需要 QoS 控制模块来实时检测网络拥塞、调整数据流向。IP 网络上数据包的问题比较复杂,在构造数学模型的基础上可以用 2 个参数来作定性分析:(1)丢包率 P_L (packet loss rate); (2)环回时间 T_{RTT} (round-trip time)。根据通用网络 GE 模型^[6],综合丢包率 P_L 和分组环回时间 T_{RTT} 的作用,可以估算当前可得到的带宽资源为

$$W_{cur} = \frac{C}{T_{RTT} + \sqrt{P_L}} \quad (1)$$

其中, C 为常数,通常依据接收方采取的应答方式取 1.22 或者 1.31^[7]。

对 W_{cur} 关于时间 t 求导可得:

$$\frac{dW_{cur}}{dt} = \left(\frac{\partial W_{cur}}{\partial T_{RTT}} \frac{\partial T_{RTT}}{\partial t} + \frac{\partial W_{cur}}{\partial P_L} \frac{\partial P_L}{\partial t} \right) = \left[-\frac{1}{(T_{RTT} + \sqrt{P_L})^2} \frac{dT_{RTT}}{dt} - \frac{1}{2(T_{RTT} + \sqrt{P_L})^2 \sqrt{P_L}} \frac{dP_L}{dt} \right] \quad (2)$$

根据式(1)可知带宽的变化是由 T_{RTT} 和 P_L 这 2 个因素的变化引起的。由于本文主要研究有线网络,因此可以认为 P_L 为近似常数,即 $\frac{dP_L}{dt} = 0$,此时式(2)简化为

$$\frac{dW_{cur}}{dt} = -\frac{1}{(T_{RTT} + \sqrt{P_L})^2} \frac{dT_{RTT}}{dt} \quad (3)$$

即为有线网络带宽变化模型。对 W_{cur} 求二次导数得:

$$\frac{d^2W_{cur}}{dt^2} = \frac{2}{(T_{RTT} + \sqrt{P_L})^3} \frac{d^2T_{RTT}}{dt^2} \quad (4)$$

每次根据式(1)、式(2)、式(3)求连续 3 个采样时刻 t_1, t_2, t_3 的 W_{cur} 值,之后根据式(4)求得 ΔW_{cur} ,对比 ΔW_{cur} 值符号可以有如表 2 所示的 4 种结果。

表 2 根据 ΔW_{cur} 符号预测的结果

ΔW_{cur} 值符号	结果 1	结果 2	结果 3	结果 4
采样点 t_1 与 t_2	+	+	-	-
采样点 t_2 与 t_3	+	-	+	-

根据表 2 可以预测网络变化,以 W_{cur} 为主要判定标准,本文将网络变化大致分为 3 种情况:

(1)根据结果 1 及式(4)求得 $\frac{d^2W_{cur}}{dt^2}$ (即 ΔW_{cur} 的变化量),

若为正则可以预测网络带宽将进一步增加,定义此时的网络状况为 S_1 ,若为负可以预测网络带宽趋向平稳,定义此时的网络状况为 S_2 。

(2)根据结果 2、结果 3 可以判断网络带宽不稳定,无法进行预测。此时传输策略不宜改变,定义此时的网络状况为 S_3 。

(3)根据结果 4 及式(4)求得 $\frac{d^2W_{cur}}{dt^2}$ (即 ΔW_{cur} 的变化量),若为正可以预测网络带宽进一步减小,定义此时的网络状况为 S_4 ,若为负表示网络带宽减小趋缓,网络趋向稳定,定义此时的网络状况为 S_5 。 S_5 与 S_4 采取同样的措施。

综上所述最终的传输方法如下:

定义基本层子流数目为 l_{base} 个,增强层子流数目为 l_{en} 个,最大可利用带宽 W_{max} ,最小可利用带宽 W_{min} ,完成最基本解码所需带宽 W_{base} (即基本层信息完全解码所需带宽),网络当前带宽为 W_{cur} , l_i^B 为当前状态下传输的基本层子流数目 ($l_i^B \leq l_{base}$), l_i^E 为当前状态下传输的增强层子流数目 ($l_i^E \leq l_{en}$), $G_i(t)$ 为使用 i 力度 FEC 保护的增强层子流数目, $H_j(t)$ 为使用 j 力度 FEC 保护的基本层子流数目, i, j 值越大保护力度越强,相应最大最小值分别为 $i_{max}, i_{min}, j_{max}, j_{min}$ 。传输的原则是:启动传输的子流总是未传输子流中重要性最强的子流,关闭的子流总是传输子流中重要性最小的子流。这样最后的传输方法可通过如图 4 所示的有限状态机来描述。

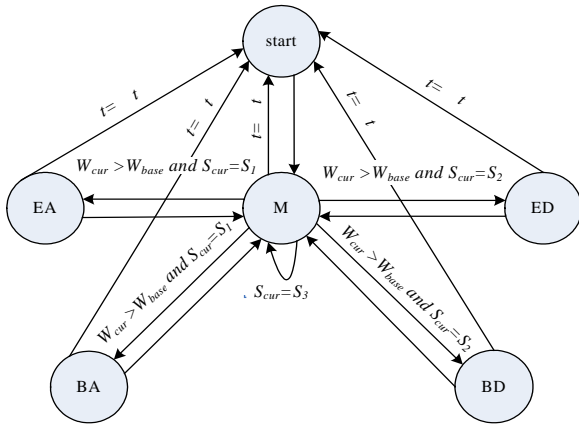


图 4 本文传输方法的有限状态机描述

定义本文网络传输系统为 6 个状态:

- (1)START 状态:计算 $P_L(t), T_{RTT}(t)$ 和 W_{cur} 等基本网络参数,进入 M 状态。
- (2)M 状态:通过对网络参数和设定值比较,进行网络状态的判定,从而转入下一状态。
- (3)EA 状态:此时 $W_{cur} > W_{base}$,且网络带宽趋向于进一步增加,则传输策略为 $l_i^E = l_{i-1}^E + 1, l_i^B = l_{base}, G_i(t) = G_{i-1}(t-1) + 1, H_{j_{max}}(t) = H_{j_{max}}(t-1)$ 。
- (4)ED 状态:此时 $W_{cur} > W_{base}$,但网络带宽趋向于进一步减小,传输策略为 $l_i^E = l_{i-1}^E - 1, l_i^B = l_{base}, G_i(t) = G_{i-1}(t-1) + 1$ 。
- (5)BA 状态:此时 $W_{cur} < W_{base}$,但网络带宽趋向增加,传输策略为 $l_i^B = l_{i-1}^B + 1, H_j(t) = H_{j-1}(t-1), l_i^E = 0$ 。
- (6)BD 状态:此时 $W_{cur} < W_{base}$,且网络带宽趋向于减小,传输策略为 $l_i^B = l_{i-1}^B - 1, H_j(t) = H_{j-1}(t-1), l_i^E = 0$ 。
- (7)任何状态下,间隔 t 时间达到 Δt 时,无条件返回初始状态重新计算网络参数,消除积累误差。

3 仿真试验

本文采用支持 SCTP 协议的 NS-2 8.0A 版本作为网络仿真平台。方法 1:传输层使用 UDP 传输协议的传统网络传输结构,信源结构同本文方法;方法 2:采用本文网络结构但纠错采用文献[6]的不等保护纠错方法。网络拓扑图如图 5 所示。信源端采用 foreman 和 mother 序列、CIF 格式,帧频 30 帧/s,PFGS 编码的增强层最大为 6 层,各参数取值为 $L_{max} = 6, L_{normal} = 4, W_{min} = 30k, W_{base} = 150k, W_{normal} = 600k, W_{max} = 1.5M$,FEC 纠错采用 RS 码。 Δt 取为 3,每隔 3s 重新计算一次 W_{cur} ,及时消除积累误差同时将计算量控制在一定范围内,节省计算成本。

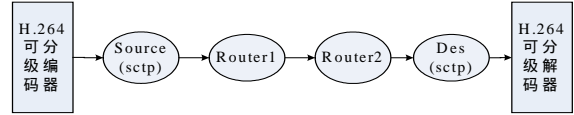


图 5 仿真试验网络拓扑

图 6、图 7 表示信源分别为 foreman 序列和 mother 序列下采用本文传输方法和采用文献[6]的不等重保护方法及普通 PFGS 视频流等错误保护方法的视频流 PSNR 比较。

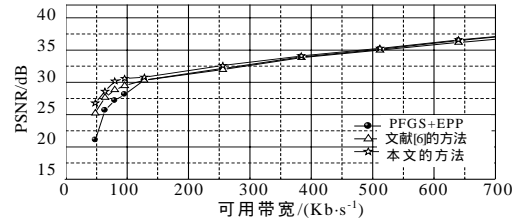


图 6 foreman 序列各传输方法 PSNR 值比较

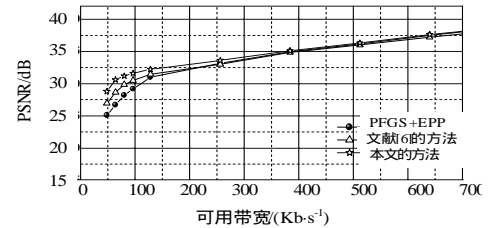


图 7 mother 序列各传输方法 PSNR 值比较

由于不等重保护方法,使得在低带宽网络条件下,重要层数据可靠到达,很大程度上改善了视频流的抗误码性能。因此本文方法和文献[6]方法在低带宽条件下比 PFGS 视频流等保护方法的系统性能均有较大改善。本文方法根据网络带宽的二阶导数来预测网络带宽的短时期变化,使得码率输出更加稳定,曲线更加平滑,因此,本文方法具有更强的网络适应能力。每隔 3s 计算一次 W_{cur} ,更新预测起始值,消除了预测带来的积累误差。从图中可以看出,带宽越低,本文方法 PSNR 值提高越大,这是因为本文采用 SCTP 作为传输基本层的协议。foreman 序列,在带宽 100Kb/s 条件下,本文方法的平均 PSNR 值比等保护方法提高大约 1dB,比文献[6]方法提高 0.5dB;极低带宽 70Kb/s 条件下,本文方法的平均 PSNR 值比等保护方法提高大约 1.5dB,比文献[6]方法提高 1dB 左右。对于 mother 序列,由于图像场景变化不大,因此,本文方法的 PSNR 值提高没有 foreman 序列幅度大。图 8 为 2 种方法下 foreman 序列的码率输出曲线比较。

测试序列共 50 帧,可以看出本文方法曲线明显平滑,因此,在网络适应性上本文的方法更强,这进一步证明了该方法的优越性。

(下转第 33 页)