

# 局域网环境下视频实时传输系统的软件实现方法研究

王文义<sup>1</sup> 李尊锋<sup>1</sup> 周 兵<sup>2</sup>

<sup>1</sup> (郑州大学信息工程学院, 郑州 450052)

<sup>2</sup> (北京航空航天大学计算机科学系, 北京 100083)

E-mail zgjsjwwy@263.net

**摘 要** 该文介绍了一种基于软件编/解码的数字视频实时传输系统的实现方法,能够在普通局域网环境中实现点到点的视频实时采集、编码、传输、解码和显示。该方法采用一种新的快速运动估计技术和改进的 TM5 位率控制算法以提高编码器的编码效率和编码图像质量,并通过选择合适网络协议和发送策略保证传输的可靠性。实验表明,所用方法不但能够保证图像传输质量,而且能够满足网络实时性的要求。

**关键词** MPEG2 实时传输系统 视频编码器 打包算法 运动估计 位率控制

文章编号 1002-8331- (2002)11-0154-03 文献标识码 A 中图分类号 TP393

## The Research of Video Real-time Transmission System Using Software Implementation under the Environment of LAN

Wang Wenyi<sup>1</sup> Li Zunfeng<sup>1</sup> Zhou Bing<sup>2</sup>

<sup>1</sup> (Information Engineering College of Zhengzhou University Zhengzhou 450052)

<sup>2</sup> (Dept. of Computer Science, Beijing University of Aeronautics & Astronautics, Beijing 100083)

**Abstract:** This paper presents a method of digital video real-time transmission system implemented by using software to accomplish encoder and decoder. It can real-timely achieve sampling, encoding, transmission, decoding and displaying from end to end on a common LAN. The method adopts a new technology of fast motion estimation and the improved TM5 bit rate control algorithm to raise the encoding efficiency and the image quality of encoder and chooses the suitable network protocol and strategy of sending to insure transmission reliability. Experiments show that the method not only can insure the transmission quality of image but also can satisfy the demand of network real-time feature.

**Keywords:** MPEG2 real-time transmission system, video encoder, packetization algorithm, motion estimation, bit rate control

### 1 引言

随着多媒体技术和互联网的发展,数字视频的实时传输成为计算机领域一个研究的热点。传统的模拟视频信息由于抗干扰能力差而不利于传输,而数字视频信息却能够克服这个缺点,但数字视频在传输业务中也存在着视频信息数据量巨大,延迟敏感性强等特点。所以,必须对视频数据进行压缩,在实时传输中保证低延迟。目前流行的编码器基本上都是采用硬件实现,存在价格高、灵活性差、不易修改等缺点。随着计算机处理速度的不断提高和内存空间的不断扩大,通过软件方法实现数字视频的实时传输已成为可能。而且软件实现编/解码,具有很大的灵活性,能随时根据用户的不同需求修改编/解码器,这是硬件编码器无法做到的,但它也存在着编码效率问题。因此,在以软件方法实现数字视频的实时传输中,最大限度地提高编码器编码效率,提高传输可靠性,减少传输延迟,成为该系统设计的关键。该文提出一种用软件实现视频信息编码的方案,它既能很好地克服用硬件实现的不足之处,又能满足对不同类型视

频信息处理的要求。

### 2 系统实现

#### 2.1 硬件系统

基于局域网中的两台 PC 机。其中一台作为编码服务端,通过 USB 接口连接一个摄像头,用于采集、编码、发送,机器为 P III 800MHZ,外频 133MHZ,内存 256MB;另一台作为解码客户端,用于接收、解码、显示,机器为 P II 450MHZ,外频 100MHZ,内存 128MB。系统处理流程如图 1 所示。

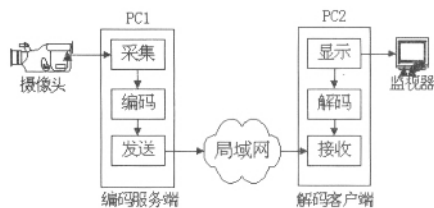


图 1 系统的处理流程图

基金项目:国家自然科学基金(编号:69935010),教育部优秀青年教师资助计划项目

作者简介:王文义,男,1947年生,教授,主要研究领域为并行处理技术和软件工程。李尊锋,男,1977年生,硕士研究生,研究方向为图像处理技术。

周兵,男,1964年生,博士研究生,研究方向为图像处理及视频编码技术。

2.2 软件系统

系统软件平台为 Windows 2000 ,开发工具采用 MS Visual C++ 6.0。根据图 1 中的系统处理流程图 ,软件系统的总体结构设计如图 2。

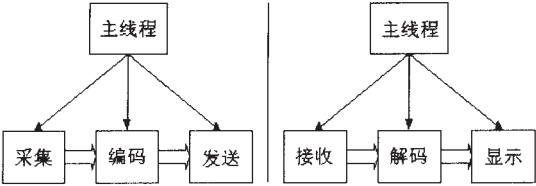


图 2 总体结构图

图中 ,编码服务端的采集、编码和发送三个模块 ,以及解码客户端的接收、解码和显示三个模块 ,分别以一个子线程来实现 :

(1)实时采集。利用 video for windows (VFW) 中提供的 API 接口函数 ,通过摄像头实时采集视频数据 ,并将采集到的数据放入指定的缓冲区中。

(2)实时编码。利用标准 C 实现不包括 B 帧的 mpeg2 的编码器 ,生成的 mpeg2 码流写入 FIFO 缓冲队列 (实际就是编码器位率控制的 VBV buffer)。

(3)实时传输。利用 Windows sockets 进行网络编程。采用客户机/服务器模式进行点到点的通信 ,编码器端集成服务器 ,解码器端集成客户机。通过指定服务器的 IP 地址和约定端口号实现连接。服务器端采用并发机制 ,能够接受多个请求 ,满足多路视频的实时显示。编码服务端从 FIFO 缓冲队列中利用定时器定期发送固定尺寸的数据包。解码客户端从网络中接收数据并放入客户端的 FIFO 缓冲队列中。

(4)实时解码。利用标准 C 实现标准的 mpeg2 解码器 ,解码线程从客户端 FIFO 缓冲队列取出数据进行解码。

(5)显示。采用 DirectDraw 的几种通用的显示模式进行回放。支持图像拉伸和多种显示格式。

2.3 相关技术的解决方法

(1)线程同步。该系统采用多线程机制 ,所以存在数据的同步问题 ,可采用关键代码段来限制多个线程访问同一代码段。

(2)队列的上、下溢。当出现队列上溢时 ,导致视频信息出错 ,传统的解决方法是对出错帧做跳帧处理。而该系统则采用暂时挂起编码线程 ,强制发送 ,发送完再唤醒编码线程来解决。当队列下溢时 ,采用挂起解码线程 ,接收到数据后唤醒解码线程的方法来解决。

(3)定时发送。采用多媒体定时器以准确定时。考虑到某一个发送周期不一定能保证队列中有发送尺寸数据的情况 ,这里采用的解决方法是挂起发送线程 ,等到下一个发送周期再发送。因此 ,选择一个合适的发送周期和发送尺寸 ,保证每次发送都有数据 ,就显得尤为重要。

3 系统实现的关键技术

编码器和传输部分的设计水准对视频数据传输系统的实时性至关重要。因此 ,文章将着重讨论网络协议、发送模块和编码器等三部分的设计。同时 ,对某些影响系统实时性的关键问题进行必要的分析 ,并给出相应的有效算法。

3.1 网络协议的选择

Internet 多媒体信息传输一般都采用 RTP 协议 ,但由于

RTP 本身不提供及时传输和 QoS 保证 ,因此尚需要额外定义一些传输控制策略来保证服务质量。考虑到这种情况 ,该系统拟选择 TCP 作为传输控制协议。TCP 是面向连接的 ,能够保证传输的可靠性。由于基于局域网的系统因传输错误数据重发带来的延迟开销较小 ,因此能够满足视频实时传输对延迟的要求。同时 ,也由于增加了传输的可靠性 ,对信源无须再进行纠错编码 ,这就能较好地保证图像传输质量 ,提高编码器效率。

3.2 发送模块的设计

发送模块主要实现模拟恒定位率 ,用定时器定期发送固定尺寸的数据包的打包算法<sup>[1]</sup>。该算法不仅实现方法比较简单 ,而且还可以在实验中通过调节定时器周期  $T$  和发送包尺寸  $N$  ,来模拟各种位率的编码 ,以获取一定图像质量下合适的编码参数 ,满足 :

$$T \times N = R$$

其中  $T$  表示定时器的周期 , $N$  表示数据包的固定尺寸 , $R$  表示恒定位率。这样 ,通过反馈 ,编码器就能够进行有效的位率控制。

3.3 编码器的设计

3.3.1 编码帧类型的选择

MPEG2 视频有 I、P、B 三种帧类型。其中 I 帧为帧内编码 ,P 帧为前向预测编码 ,而 B 帧为双向预测编码 ,所以其运算量将远远大于前两帧。由于预测时 B 帧必须保留前向和后向两帧的参考帧图像 ,所以它占去了大量的缓冲区资源。鉴于此 ,该系统为了提高编码器的效率 ,将不对 B 帧进行编码 ,这样做虽然会导致生成码字一定量的增加 ,但由于局域网带宽相对较高 ( $\geq 10M$ ) ,完全可以在满足视频传输要求的前提下 ,大大减少编码器的运算量。

3.3.2 基于块匹配的快速运动估计技术

在基于运动估计的帧间压缩方案的编码过程中 ,运动估计占很大计算量 (如 H.263 编码器中 ,就占了 42% 的计算量<sup>[2,3]</sup>) ,而基于块匹配的搜索策略又是运动估计的关键。因此 ,为提高编码器的效率 ,该系统编码器采用了菱形的自适应快速运动搜索策略<sup>[3]</sup>。

搜索起点预测 利用运动的相关性 ,以空间相邻块和时间相邻块 (前一帧图像相同位置块) 的运动矢量来预测当前块搜索的起点位置。

搜索模式 搜索路线采用基于梯度式的菱形搜索 (DS) 模式 ,因为菱形符合编码运动矢量位数的分布<sup>[4]</sup>。搜索方向采用梯度式搜索 ,该方式搜索点的分布面广 ,加上其 “跨越式” 的搜索 ,因此能较快到达搜索的终点 ,搜索步长是自适应调节的 ,根据不同的运动类型 ,采用大菱形搜索 (LDS) 和小菱形搜索 (SDS) ,如图 3 ,LDS 用于大范围的快速定位 ,SDS 则用于其小范围的 “微调”。

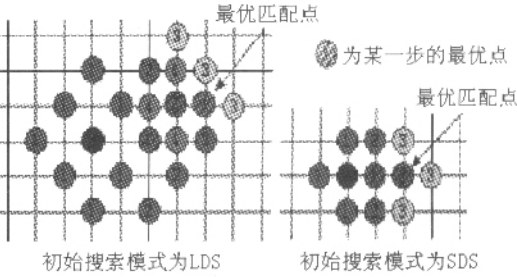


图 3 菱形的自适应快速运动搜索示例

搜索中止准则 在保证搜索精度的前提下及时终止,这样能大大减少搜索时间。该算法的中止条件:当 SDS 模式搜索到的最优点在中心时(如图 3);SAD (匹配块和当前块的绝对差之和)满足某一阈值时;若  $SAD < T_1$ ,表明搜索结果已足够精确,应立即终止搜索;若  $T_1 \leq SAD < T_2$ ,表明搜索结果尚可,只需在小范围内进行微调,于是将搜索模式从 LDS 转为 SDS。其中,  $T_1 = \alpha \times Prev\_SAD$ ,  $T_2 = \beta \times Prev\_SAD$ ,  $Prev\_SAD$  为参考帧相同位置块的 SAD。若当前帧为第一个 P 帧,不存在参考  $Prev\_SAD$ ,则令  $T_1 = 500$ ,  $T_2 = 750$ 。这里  $0.9 < \alpha < 1.4$ ,  $1.2 < \beta < 2.5$ ,且  $\alpha < \beta$  ( $\alpha$ ,  $\beta$  属于经验值)。

用该搜索算法可以使运动估计的速度比使用全搜索算法提高 44.9 倍(如表 1,该表是对 cif 格式 352×288 图像,下同)。

3.3.3 基于 MMX 指令级的快速 DCT 和反 DCT 算法

该算法充分利用 MMX (多媒体扩展指令集)技术,利用汇编指令实现 DCT 和反 DCT 变换。由于采用了 MMX 技术,相对于标准的 DCT 和反 DCT 变换来说,使得变换和反变换的速度分别提高了 17.3 倍和 13.8 倍(表 1)。

表 1 平均每一帧处理该部分所需的时间比较(单位:ms)

全搜索算法的运动估计	快速搜索的运动估计	标准的 DCT 变换	MMX 技术 DCT 变换	标准的反 DCT 变换	MMX 技术反 DCT 变换
404	9	69	4	55	4
提高了 44.9 倍		提高了 17.3 倍		提高了 13.8 倍	

3.3.4 改进的 TM5 位率控制算法

由于 TM5 算法<sup>[5]</sup>存在以下两个局限:第一,分配目标位数据量时没有充分考虑图像的相对复杂度;第二,算法不能很好处理场景变换的情况。所以该系统采用编码帧图像平均复杂度计算算法,通过比较当前帧图像复杂度与所有已编码帧图像平均复杂度,合理调整当前帧目标位数据量,保证图像各帧质量的平滑度。同时,采用文献[6]中的场景变换的检测方法,动态调整场景转换帧目标位数据量,实验结果如表 2 所示。表中数据是在 0.5Mbps 和 1Mbps 位率下得到的图像的平均 PSNR 值,这些数据充分表明该系统编码器已很好地解决了因场景变换带来的质量下降问题。

表 2 TM5 位率控制与改进后的位率控制实验结果比较

序列名	分辨率	帧数	位率	TM5 SNR	改进 TM5 PSNR	注:该序列
Foreman	352×288	150	500kbps	38.12dB	38.71dB	运动平缓,有场景变换
			1Mbps	40.34dB	40.82dB	

(上接 93 页)

今后 SDSE 需要增加连续语音音节分割和句法、语义分析功能来满足连续语音识别的需要,提高语音识别引擎的应用范围。

语音是人类最自然的通信方式,语音识别技术的研究也已经进行了较长的时间,但目前真正能够实际应用的语音识别系统并不多。我国的语音处理和识别研究尽管表面上看来同先进国家之间比较接近,但实际上有较大的差距,一是跟踪研究较多,开创性、系统性研究不够,二是理论与研究与应用研究缺少衔接。今后,应该更加注重面向应用的实用语音处理与识别研究,真正提高语音处理和识别技术的研究与应用水平。

(收稿日期:2001 年 5 月)

参考文献

4 系统实验结果分析

该系统对 qcif 格式(176×144)的图像能够完全实时传输;对 cif 格式(352×288)的图像可以基本实时传输,并能很好地保证图像质量;而对于 640×480 大小的图像,则实时传输的效果较差,这是由于受软件编码能力的限制,所采集的数据不能完全得到编码,导致采集帧率大于编码帧率。另外,当编码位率增加时,虽然可使编码图像质量增强,但却使编码帧率减少,延迟增加。因此,选择图像质量和编码帧率的最佳折衷点也成为用软件方法实现实时视频传输系统所面临的一个主要问题。实验结果如表 3 所示。

表 3 不同图像尺寸和编码位率下的 PSNR 值和延迟比较

图像尺寸	位率 (mbps)	采集帧率 (f/s)	编码帧率 (f/s)	PSNR (db)	延迟
176×144	1.098	30	24	47.37	感觉不到
176×144	1.716	30	20	49.51	很小
352×288	1.098	30	15	44.11	小
352×288	1.716	30	12	45.50	较小
640×480	1.098	15	7	42.28	较大
640×480	1.716	15	4	43.70	大

5 结束语

该文提出了一个用软件方法实现视频实时传输系统的设计方案,针对系统设计中的关键问题进行了分析并给出了有效的解决方法。该系统中的某些技术已成功地应用于家庭监控和某些军事领域。当然,该系统也还存在着某些不足之处,比如打包算法,因为是按固定尺寸发送,并未考虑视频流的特点;此外,系统还未加入错误控制机制,当出现不能恢复的数据丢失时,解码器就会显得无能为力,这些都有待于进一步改进和完善。(收稿日期:2001 年 8 月)

参考文献

1.F le leannee C M Guillemot.Error resilient video over the internet[C]. SPIE proc Visual Communications and Image Processing (VCIP),1999 : 271~280  
2.Yi Liu.Software Codecs on PC.AMD 2000  
3.涂亚明,李波.一种基于自适应搜索的快速运动估算法[J].北京航空航天大学学报,2000-04  
4.A M Tourapis,O C Au,M L Liou et al.An Advanced Zonal Block Based Algorithm for Motion Estimation[C].ICIP,1999 :610~614  
5.Test Mode 5.ISO/IEC JTC1/SC29/WG11/N0400,1993  
6.余越,周健等.一种场景切换检测及自适应码率控制方法[J].通信学报,1999-05  
1.L Rabiner,B H Juang.Fundamentals of Speech Recognition[M].Pren-tice hall,1993 :97~121  
2.俞一彪,袁保宗.汉语语音理解系统 BJD-I 的声学语音处理[J].信号处理,1987 3 (4) :206~213  
3.M J E Gales.Cluster Adaptive Training of Hidden Markov Models[J]. IEEE speech audio processing,2000 8 (4) :417~428  
4.Lin Dong,Baozhong Yuan,Jicheng Zhou.A Chinese Speech Recognition Method Based on Neural Network[C].In Proceeding of CSSP,1999-10 :200~203  
5.Yibiao Yu,Heming Zhao.Speech Recognition Based on Estimation of Mutual Information[C].In Proceeding of ICSLP PII,2000-10 :1046~1049  
6.张贤达.现代信号处理[M].北京:清华大学出版社,1995 :101~103  
7.李博轩.Visual C++6.0 多媒体开发指南[M].北京:清华大学出版社,2000